UNIVERSITÉ DE NEUCHÂTEL

The Multidimensional Nature of Poverty in Upper-Middle-Income Countries

PART I

Ph.D. Thesis/Dissertation

Author: A. Tugce BEYCAN

Supervisor: Christian Suter

Department of Sociology, University of Neuchâtel

Faculty of Letters and Human Sciences

December 18, 2020

UNIVERSITÉ DE
NEUCHÂTEL

Faculté des Lettres et Sciences Humaines
Espace Tilo-Frey 1
CH - 2000 Neuchâtel
Tél : +41 (0)32 718 17 00
Fax : +41 (0)32 718 17 71
secretariat.lettres@unine.ch

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de M. Christian Suter, directeur de thèse, professeur, Université de Neuchâtel ; Mme Filomena Maggino, professore, Dipartimento di Scienze Statistiche, Sapienza Università di Roma ; Mme Valerie Møller, Professor Emeritus, Quality of Life Studies, Institute of Social and Economic Research, Rhodes University, Grahamstown, South Africa ; M. Serge Paugam, Directeur de recherche au CNRS, Directeur d'études à l'EHESS, Directeur du Centre Maurice Halbwachs (CNRS/EHESS/ENS), Paris autorise l'impression de la thèse présentée par Madame Aysegül Tugce Beycan en laissant à l'autrice la responsabilité des opinions énoncées.

Neuchâtel, le 18 décembre 2020

Le doyen
Pierre Alain Mariaux

P.O.

Acknowledgment

I would keep quite short acknowledgements as I do always prefer keeping private my private life but of course, I would like to thank first of all my family especially my dad and mom for being simply the most democratic and best parents in the world, for their unconditional love, kindness, endless support, and great respect towards my choices since my childhood; all my friends for just being the most respectful, supportive, faithful, and true friends through years- they know who they are!, I know most of them for years and years since our childhood and university years, they have always been by my side and I know they will always be, all of them have a special place in my heart so I won't do an acknowledgment list for respect to each of them; but also, my colleagues I have found a chance of meeting and/or working with them- notably I have learnt a lot from M. Hakki Ozel, Elisa Benes, Alina Matei, Christine Maurel, Sandrine Casanova, Rainer Bruggemann, Serge Paugam, Salvatore Babones, Samuel Cohn, Filomena Maggino, and Vani, I cannot thank them enough, and all I can do is wishing them endless luck, success, health during their life time and I hope stars will always be with them. I would like to thank also all the members of the jury of my dissertation, I am so grateful to Filomena Maggino, Valerie Møller, and Serge Paugam for having accepted to be part of the jury. Many thanks also to the national statistical offices of Brazil, Mexico, South Africa, and Turkey for giving me the data for this dissertation. And top of that, I would like to thank my excellent and greatest supervisor in the entire world, Prof. Christian Suter, for his endless kindness, patience, understanding, support, authenticity, humanity and believing in me. I cannot truly find enough worlds to describe how I feel lucky to have found a chance of working with him, he taught me how to respect and love researching and teaching.

Finally, I dedicate my work to my late great uncle and great aunt who are now shining in the sky.

Summary

This thesis deals with the multidimensional nature of poverty in upper-middle-income countries with a specific eye on Brazil, Mexico, South Africa and Turkey by employing national household surveys. I conceptualize poverty based on Amartya Sen's *capabilities* approach by categorizing *functionings* in terms of micro and macro aspects. This typology shows the importance of social bonds (e.g., relationships with family members, community, colleagues at work, friends, people in daily life) in determining capabilities of people. Multidimensional poverty based on capabilities is often measured via indices such as the Multidimensional Nature of Poverty developed by Alkire and Foster, which is used by the United Nations in annual human development reports since 2010. However, such indices do not show the combinations of functionings and provide the number-oriented studies (such as headcount ratio of poverty, intensity of poverty, severity of poverty rates, contribution of each indicator to the index). Sen defines capabilities as the combinations of functionings and I argue that for understanding the real liberty of people we need to capture comprehensively living conditions. Accordingly, I propose a new poverty measurement methodology by employing classification trees (Breiman et al.'s CART) and profiles extraction method. Through the 4 case studies (Brazil, Mexico, South Africa, Turkey), I demonstrate the most relevant combinations of functionings ('multi-configurations of poverty'), distinguishing poverty groups. The results show the major inequalities between poverty groups, from the most prosperous to the poorest. Moreover, I explain how this methodology can be used by policy makers to reduce poverty, and I present in an extensive way the poverty approaches from the literature such as relative deprivation, social exclusion, social disqualification, underclass, basic needs, marginalidad, absolute poverty, monetary poverty, culture of poverty, subjective poverty.
Key words: Multidimensional poverty; Capabilities; Upper-Middle-Income countries; Brazil; Mexico; South Africa; Turkey

Résumé

Cette thèse explique la nature multidimensionnelle de la pauvreté dans les pays à revenu intermédiaire de la tranche supérieure avec un œil particulier sur le Brésil, le Mexique, l'Afrique du Sud et la Turquie en utilisant des enquêtes nationales sur les ménages. Je conceptualise la pauvreté sur la base de l'approche des *capabilitiés* d'Amartya Sen en catégorisant les *functionings* en termes d'aspects micro et macro. Cette typologie montre l'importance des liens sociaux (par exemple, les relations avec les membres de la famille, la communauté, les collègues de travail, les amis, les personnes dans la vie quotidienne) dans la détermination des capabilités des personnes. La pauvreté multidimensionnelle basée sur les capabilités est souvent mesurée à l'aide d'indices tels que l'indice de la pauvreté multidimensionnelle développée par Alkire et Foster, qui est utilisée par les Nations Unies dans les rapports annuels sur le développement humain depuis 2010. Cependant, ces indices ne montrent pas les combinaisons de functionings et fournissent les études axées sur le nombre de pauvres (telles que le taux de pauvreté, l'intensité de la pauvreté, la contribution de chaque indicateur à l'indice). Sen définit les capabilités comme des combinaisons de functionings et pour comprendre la liberté réelle des personnes, nous devons saisir les conditions de vie de manière complète. En conséquence, je propose une nouvelle méthodologie de mesure de la pauvreté en utilisant des arbres de classification (Breiman et al.) et la méthode d'extraction de profils. À travers les 4 études de cas (Brésil, Mexique, Afrique du Sud, Turquie), je démontre les combinaisons de functionings les plus pertinentes ('multi-configurations de la pauvreté'), en distinguant les groupes de pauvreté. Les résultats montrent les inégalités majeures entre les groupes de pauvreté, des plus prospères aux plus pauvres. De plus, j'explique comment cette méthodologie peut être utilisée par les décideurs politiques pour réduire la pauvreté, et je présente de manière détaillée les approches de la pauvreté à partir de la littérature telles que la privation relative, l'exclusion sociale, disqualification sociale, *underclass*, *basic needs*, *marginalidad*, la pauvreté absolue, la pauvreté monétaire, la culture de la pauvreté, la pauvreté subjective.
Mots-clés : Pauvreté multidimensionnelle ; Capabilités; Pays à revenu intermédiaire de la tranche supérieure; Brésil; Mexique; Afrique du Sud; Turquie

Table of Contents

1. Introduction 17
1.1 Goal of research and scientific pertinence 17
1.2 World poverty 27
1.3 Structure of thesis 34
2. Theoretical framework 37
2.1 Capabilities 39
2.2 Capabilities: absolute or relative approach on poverty? 92
Absolute poverty 93
Basic needs 103
Relative deprivation 108
Sen's view on absolute versus relative approaches 119
2.3 Capabilities and social bonds 124
Social exclusion, social disqualification, marginalidad 125
How social relationships define capabilities of the person 139
2.4 Multidimensional poverty 145
3. Poverty in upper-middle-income countries 173
3.1 Socio-economic background of select upper-middle-income countries 173
3.2 National poverty measurement practices in select upper-middle-income countries 179
3.3 National poor relief and poverty eradication programs in South Africa, Turkey, Mexico, and Brazil 193
4. Methodology 217
4.1 Introduction 217
Indicators, dimensions 217
Aggregation technique 219
Decomposition analysis 222
4.2 Data and indicators 226
4.3 Steps to follow in the analysis 282
5. South Africa: an exemplary case study 285
5.1 Introduction 285
Quality report of variables 285
5.2 First step (step1 analysis): detecting multidimensional poverty patterns 286
5.3 Second step (step2 analysis): detecting dimensional poverty profiles of poverty groups 401
5.4 Third step (step3 analysis): detecting various profiles of poverty groups 410
5.5 Fourth step (step4 analysis): detecting multidimensional poverty profiles of specific groups . 435

Table of Illustrations

Table 1. Annie's functionings, an example for putting in context our typology of functionings..... 78
Table 2. Kate's capability set and functionings, an example for putting in context our typology of functionings and capability set in a dynamic aspect. 81
Table 3. Kate's capability set and personal goals, an example for putting in context our typology of functionings and capability set in a dynamic aspect. 82
Table 4. Dietaries by men, women, children, drawn by Rowntree (1901, 99-102) 98
Table 5. The weekly cost of a standard dietary for men, women, and children, computed by Rowntree (1901, 105) 99
Table 6. Minimum necessary expenditure per week for families of different size, called the primary (weekly) poverty line drawn by Rowntree (for a full coverage see Rowntree (1901, 110)) 99
Table 7. The list of basic needs and their core indicators, source: Hicks and Streeten (1979, 578) ... 105
Table 9. Example 1 of the computation of the MPI, source: Author 153
Table 10. MPI Indicators and their description, source: UNDP Technical notes of the Human Development Report 2016, 9-10 154
Table 11. Example 2 of the computation of the MPI, source: Author 155
Table 12. Dimensions and Indicators used in the computation of the MPI across European Union countries by Alkire and Apablaza (2016, 13)..... 157
Table 13. Dimensions, indicators, first cut-off developed by Suppa (2017) for the German MPI, source: Suppa (2015, 30)..... 160
Table 14. SDG Indicators (dashboard of indicators) for Goal 1. End poverty in all its forms everywhere. 167
Table 15. OECD Better life index dashboard..... 169
Table 16. Some of Sustainable Development Goals (SDGs) indicators for select 19 upper-middle-income countries based on UN (United Nations) data base 176
Table 17. Some socio-economic indicators for select 19 upper-middle-income countries based on UN (United Nations) data base 177
Table 18. Some of World Development indicators for select 19 upper-middle-income countries based on Word Bank data base..... 179
Table 19. Indicators and deprivation thresholds of the social deprivation index of the MMPM 184
Table 20. Dimension and indicators of the Colombian Multidimensional Poverty Index (CMPI)..... 187
Table 21. Dimensions and indicators of the Costa Rican Multidimensional Poverty Index (CRMPI)188
Table 22. Ecuador's Unmeet Basic Needs (NBI) dimensions and indicators 189
Table 23. Ecuador Multidimensional Poverty Index (ECMPI) dimensions and indicators 190
Table 24. Indicators and deprivation thresholds of the SAMPI 192
Table 25. Features of selected variables from National General Household Survey 2011, South Africa 230
Table 26. Features of selected variables from Statistics on Income and Living Conditions Survey (SILC) 2011, Turkey 251
Table 27. Features of selected variables from Modulo de Condiciones Socioeconomicas (MCS) 2012, Mexico..... 263
Table 28. Features of selected variables from National Household Sample Survey (PNAD) 2011, Brazil 272
Table 29. Deprivation lines and composition of variables for step 1 of the methodology: overall deprivation index and multidimensional poverty patterns, data: South Africa GHS 2011 289
Table 30. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GR and 33 variables), data: South Africa GHS 2011 299
Table 31. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GR and 49 variables), data: South Africa GHS 2011 300

Table 32. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GRS and 33 variables), data: South Africa GHS 2011 301
Table 33. Per centage of groups in the samples of target groups, data: South Africa GHS 2011 301
Table 34. Confusion matrix for showing the utility of the loss parameter in rpart 303
Table 35. Training vs. test samples in CAP111 analysis, data: South Africa GHS 2011 306
Table 36. cp table- summary information on complexity parameter, split, and error rates of the saturated CAP111 tree, data: South Africa GHS 2011..... 310
Table 37. Confusion matrix of the CAP111's predicted results on the test data (based on the fitting model with default parameters), data: South Africa GHS 2011..... 320
Table 38. An example of confusion matrix for explaining accuracy 326
Table 39. Coverage of GR1's patterns among the population classified as GR1 (in per centage), among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP111S2..... 331
Table 40. Coverage of GR2's patterns among the population classified as GR2 (in per centage), among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP111S2..... 332
Table 41. Training vs. test samples in CAP112 analysis, data: South Africa GHS 2011 335
Table 42. Coverage of GR1's patterns among the population classified as GR1 (in per centage), among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP112, data: South Africa GHS 2011..... 341
Table 43. Coverage of GR3's patterns among the population classified as GR3 (in per centage), among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP112, data: South Africa GHS 2011..... 341
Table 44. Training vs. test samples in CAP113 analysis, data: South Africa GHS 2011 345
Table 45. Coverage of GR1's patterns among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP113 pruned tree, data: South Africa GHS 2011..... 347
Table 46. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP113 pruned tree, data: South Africa GHS 2011..... 349
Table 47. Training vs. test samples in CAP114 analysis, data: South Africa GHS 2011 350
Table 48. Coverage of GR1's patterns among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP114 tree, data: South Africa GHS 2011..... 352
Table 49. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP114 tree, data: South Africa GHS 2011..... 353
Table 50. Training vs. test samples in CAP115 analysis, data: South Africa GHS 2011 354
Table 51. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP115 pruned tree, data: South Africa GHS 2011..... 358
Table 52. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP115 pruned tree, data: South Africa GHS 2011..... 358
Table 53. Training vs. test samples in CAP116 analysis, data: South Africa GHS 2011 359
Table 54. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP116 pruned tree, data: South Africa GHS 2011..... 362
Table 55. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP116 pruned tree, data: South Africa GHS 2011..... 362

Table 56. Training vs. test samples in CAP117 analysis, data: South Africa GHS 2011 363
Table 57. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP117 pruned tree, data: South Africa GHS 2011..... 365
Table 58. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP117 pruned tree, data: South Africa GHS 2011..... 366
Table 59. Training vs. test samples in CAP118 analysis, data: South Africa GHS 2011 367
Table 60. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP118 pruned tree, data: South Africa GHS 2011..... 370
Table 61. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP118 pruned tree, data: South Africa GHS 2011..... 370
Table 62. Training vs. test samples in CAP119 analysis, data: South Africa GHS 2011 372
Table 63. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP119 pruned tree, data: South Africa GHS 2011..... 375
Table 64. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP119 pruned tree, data: South Africa GHS 2011..... 375
Table 65. Training vs. test samples in CAP1110 analysis, data: South Africa GHS 2011 376
Table 66. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP1110 pruned tree, data: South Africa GHS 2011..... 380
Table 67. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP1110 pruned tree, data: South Africa GHS 2011..... 380
Table 68. Main confusion matrix statistics per model, data: South Africa GHS 2011 393
Table 69. Dimensional profiles of GR3 based on the CAP112 model, data: South Africa GHS 2011405
Table 70. Dimensional profiles of GR4 based on the CAP113 model, data: South Africa GHS 2011408
Table 71. Dimensional profiles of GR5 based on the CAP114 model, data: South Africa GHS 2011410
Table 72. Sub-group profiles of GR3 based on the CAP112 model, data: South Africa GHS 2011 .. 413
Table 73. Sub-group profiles of GR4 based on the CAP113 model, data: South Africa GHS 2011 .. 422
Table 74. Sub-group profiles of GR5 based on the CAP114 model, data: South Africa GHS 2011 .. 430

Table of Figures

Figure 1. Actually achieved functionings (AAF) and valued functionings to be achieved (VFTA) between the two periods t and t+1. 61
Figure 2. The micro level functionings (MIF) vs. the macro level functionings (MAF). 66
Figure 3. The outputs of four types of functionings (MIF, IF, MAF, EF) in terms of macro vs. micro level, achievement vs. non-achievement, advantage vs. disadvantage factor. Note: * MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5. 71
Figure 4. Functionings and capabilities in a dynamic aspect. 77
Figure 5. Multidimensional Poverty analysis based on capabilities in a dynamic prospect, see Figure 3 for the terminology of functionings. Notes: (i) * MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5; (ii) HF: harmful, NHF: not harmful. 85
Figure 6. South Africa provinces. 228
Figure 7. South Africa cities. 229
Figure 8. Regions of Turkey based on NUTS1 classification. 250
Figure 9. 26 states and 1 federal district of Brazil. 272
Figure 10. Histogram of the overall deprivation index based on 33 variables, data: South Africa GHS 2011 295
Figure 11. The saturated tree obtained by fitting the model CAP111, data: South Africa GHS 2011 310
Figure 12. The CAP111 saturated tree's cross-validated error rate by the cp and size of tree, data: South Africa GHS 2011 311
Figure 13. The CAP111 saturated tree's resubstitution and cross-validated error rates by the tree number, data: South Africa GHS 2011 312
Figure 14. The CAP111 saturated tree's resubstitution, cross-validated error rates and min(CV Error + 1SE Rule), data: South Africa GHS 2011 312
Figure 15. Cross-validated error rate by the size of tree and complexity parameter (cp), based on the model with the default parameters (CAP111S), data: South Africa GHS 2011 315
Figure 16. Cross-validated error rate by the size of tree, complexity parameter (cp) and min (CV Error+1SE) rule based on the model with default parameters (CAP111S), data: South Africa GHS 2011 316
Figure 17. Pruned classification tree of the model CAP111S (with default parameters), data: South Africa GHS 2011 316
Figure 18. Pruned classification tree of the model CAP111S (with default parameters), by employing the package rpart.plot, data: South Africa GHS 2011 317
Figure 19. Pruned classification tree of the model CAP111S (with default parameters), by employing the package partykit, data: South Africa GHS 2011 319
Figure 20. CAP111S2 saturated tree. We add S2 next to CAP111 as this is the second attempt to obtain a better classification tree for an optimal interpretation, data: South Africa GHS 2011 323
Figure 21. Resubstitution and cross-validated error rates for the CAP111S2 classification tree, data: South Africa GHS 2011 323
Figure 22. Resubstitution, cross-validated error rates and the 1-SE Rule (min (CV Error + 1 SE)) to find out the optimum cp for pruning process, data: South Africa GHS 2011 324
Figure 23. Pruned classification tree CAP111S2, data: South Africa GHS 2011 333
Figure 24. Pruned2 CAP111S2 tree, data: South Africa GHS 2011 336
Figure 25. Resubstitution, cross-validation error rates by cp and size of tree and 1-SE Rule for CAP112 saturated tree, data: South Africa GHS 2011 337
Figure 26. CAP112 tree, data: South Africa GHS 2011 340
Figure 27. CAP112 pruned tree, data: South Africa GHS 2011 344
Figure 28. Resubstitution, cross-validated error rates by cp and size of tree and 1-SE Rule for CAP113 saturated tree, data: South Africa GHS 2011 346
Figure 29. CAP113 pruned tree, data: South Africa GHS 2011 348

Figure 30. Resubstitution and cross-validation error rates by cp and size of tree and 1-SE Rule for CAP114 saturated tree, data: South Africa GHS 2011 351
Figure 31. CAP114 tree, data: South Africa GHS 2011 352
Figure 32. Resubstitution and cross validated error rates by cp and size of tree and 1-SE Rule for CAP115 saturated tree, data: South Africa GHS 2011 355
Figure 33. CAP115 pruned tree, data: South Africa GHS 2011 357
Figure 34. Resubstitution error and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP116 saturated tree, data: South Africa GHS 2011 360
Figure 35. CAP116 pruned tree, data: South Africa GHS 2011 361
Figure 36. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP117 saturated tree, data: South Africa GHS 2011 364
Figure 37. CAP117 pruned tree, data: South Africa GHS 2011 367
Figure 38. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP118 saturated tree, data: South Africa GHS 2011 369
Figure 39. CAP118 pruned tree, data: South Africa GHS 2011 372
Figure 40. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP119 saturated tree, data: South Africa GHS 2011 374
Figure 41. CAP119 pruned tree, data: South Africa GHS 2011 377
Figure 42. Resubstitution and cross-validated error rate by cp and size of tree and 1-SE Rule for CAP1110 saturated tree, data: South Africa GHS 2011 378
Figure 43. CAP1110 pruned tree, data: South Africa GHS 2011 381
Figure 44. Deprivation per centages represented by different colors in terms of variables emerging in the patterns by group within each model, data: South Africa GHS 2011 395
Figure 45. Frequency and composition of classification patterns per model, data: South Africa GHS 2011 400
Figure 46. CAP112DEC pruned tree, data: South Africa GHS 2011 416
Figure 47. CAP112DEC2 pruned tree, data: South Africa GHS 2011 419
Figure 48. CAP113DEC pruned tree, data: South Africa GHS 2011 426
Figure 49. CAP113DEC2 tree, data: South Africa GHS2011 427
Figure 50. CAP114DEC tree, data: South Africa GHS 2011 432
Figure 51. CAP114DEC2 tree, data: South Africa GHS 2011 434

1. Introduction

First of all, we will explain the goal and scientific pertinence of this thesis. Second, we will briefly introduce the situation of poverty in the world accompanied by some recent statistics published by the World Bank and the United Nations, and the approach of modern world-system in global poverty studies. This brief introduction will give context to the scope of this thesis. Third, we will show the structure of thesis in terms of chapters and sections.

1.1 Goal of research and scientific pertinence

In the classical sense, people consider poverty as the situation of not being able to afford good living standards and to meet their needs due to lack of income. However, diverse social and socioeconomic studies show that rather than being one-dimensional, monetary phenomenon, poverty manifests itself in manifold aspects of life (Sen 1981, 1984, 1987a, 1992, 1999, 2009; Alkire 2002; Fukuda-Parr 2015). The prevalent narrative of poverty as the lack of money (for monetary approach to poverty, see Orshansky 1965, Hagenaars and van Praag 1985, Hagenaars 1986, Haughton and Khandker 2009) is quite stimulated by the dominance of economics and the lack of interest in social phenomena among many economists and politicians. Money remains powerless in many aspects of life; for instance, it is not possible to buy affection, happiness, dignity, or human rights or to overcome social pressure with money or to have access to everything (e.g., the person may have money but many goods and services may not be available on the market or the person may have money but health and educational services may not be available in the place where he or she lives) and most importantly money does not necessarily provide freedom to live the life he or she values. The lack of income is just one of the financial aspects of life. That the person (aged 18 or over) does not have sufficient income to pay bills is due to being unemployed or to working in not adequate conditions including low payment, or to irregularity in the market in terms of earnings, or to incapacity of the system to provide people the rights they entitle to, or to the violation of social and economic rights of citizens. In this perspective, we can say that detecting people who do not have enough money and that labelling these people as poor won't necessarily solve the hardships these people witness because such monetary aspect remains as an indirect tool and does not demonstrate neither the origin of problem nor the state of real liberty of people. The focus should not be on whether the person has or not money but instead the focus should be on what are the opportunities, sources and means the person has actually in his or her life.

The vision of multidimensional poverty intrinsically brings us to accept the reality that because of multifarious living circumstances, even though people in the same country experience poverty from various angles. In the international political arena, the tendency is to comprehend poverty as being deprived of subsistence, basic needs[1] (i.e., lack of access to the fundamental elements of life including diet, shelter, sanitation, clothes, health and educational services) in the Global South; on the other part as being deprived of well-being (e.g., various psychological issues due to different types of social relationships, precarious labour market standards, material deprivation, social life) in Northern countries. Nevertheless, the results of assorted academic and other research-based studies (e.g., international, non-governmental organizations) show that people in many developing countries witness not only hunger, starvation, diseases, non-fulfillment of basic needs, but they do witness also many other drivers of nonhuman living conditions such as varied types of enforcements, social pressure, unsafety, violation of dignity

[1] See Streeten et al. (1981) for basic needs and see Rowntree (1901) for subsistence.

and violation of freedom. Furthermore, we do also know that some problems encountered in the daily life of people living in well-off countries are in the agenda of people living in the Global South; for instance, low-paid work, dependency on social allowances, feeling alone, feeling empty. Herewith, we can say that poverty distinguishes as a complex, more sophisticated phenomenon with multiple configurations in opposition to its popular monetarily based perception. In other words, there is no singular form of poverty in neither developing countries nor developed countries. In this sense, it becomes crucial to consider both diversity and frequency of poverty patterns for demonstrating the nature and extent of poverty.

In a multidimensional perspective, scholars have proposed several approaches to define and measure poverty. In this regard, the first initiative was to endorse one particular list of *absolute* or *basic needs* (Rowntree 1901, Booth 1902-03, Streeten et al. 1981) such as shelter, clothes, nutrition, access to educational and health services. Later, more refined approaches such as *relative deprivation* (Townsend 1979, Mack and Lansley 1985) focusing on the determination of social necessities in a given society, or *social exclusion* (Lenoir 1974; Evans 1998: Duffy 1995; Silver 1994; Paugam (ed.) 1996; Commission of the European Communities 2003, 9; United Nations 2016, 18) dealing with the process of not being able to participate effectively in social, economic, political, cultural life in a given society, are emerged. Concerning *absolute* and *relative deprivations*, the common practices consist of converting the values of items/materials into monetary metric to settle specific poverty thresholds such as 1 USD or 2 USD per day or 60 per cent of median household equivalent income (Haughton and Khandker 2009; Ravallion 1998, 2012b) and to define 'the poor' that remain under the threshold. However, the monetary indicators do not show how people suffer from different forms of hardships in their daily life. On the other hand, as *social exclusion* is a multidimensional concept per se, several debates on the link between social exclusion and poverty are flourished in the literature. Social exclusion studies generally concentrate on specific target groups such as long term unemployed, single mothers, lone parents, assisted people. At this point, the 1998 Economics Nobel Prize Laureate Amartya Sen's philosophy and economics-based 'the *capabilities approach'* (Sen 1981, 1984, 1987a, 1999, 2009) has brought a new perspective on poverty analysis.

Human development and well-being-oriented the approach of capabilities takes into account the plurality of different lifestyles as the foundation of conceptualization of poverty. Sen explains poverty as the deficit of real liberty/freedom, or in other words as the situation of not being able to have opportunities, means and sources to achieve various actions, activities, states the person values in life (i.e., opportunities, means, sources are called as *functionings* or *beings and doings* in the capabilities framework). This way of perceiving poverty is well acclaimed in the scientific community, but is also criticized for being vast, philosophic and for not being a precise concept when it comes to measuring poverty. Sen has suggested the concept of basic capabilities as the essential items the person needs including the items mentioned in the basic needs and subsistence approaches (i.e., access to drinking water, food/for overcoming hunger and starvation, shelter). This list has not been particularly approved by relativists who prefer the 'socially perceived necessities'[2] (Mack and Lansley 1985) or any other list defined on the basis of society's major cultural/social/economic practices or by the society itself via public opinion polls. But Sen has pointed out that the capabilities approach is open to consider societal

[2] For another application of 'socially perceived necessities', see also Wright (2011). The author describes the socially perceived necessities in South Africa across different social groups.

context in defining the functionings list and in fact the concept of real liberty/freedom is the absolute term of the capabilities. He just simply recommends giving priority to access to food and drinking water to eliminate hunger and starvation in the Global South.

The amount of works concerning the place of capabilities in the sociological theory, or the link between capabilities and sociology is quite limited in the literature. Relative deprivation and social exclusion approaches are generally preferred than the capabilities approach to measure poverty from the (quantitative) sociological perspective. The first effort to measure capabilities (or measure multidimensional poverty based on the capabilities approach) was the Human Development Index (HDI), which is elaborated by Amartya Sen and Mahbub ul-Haq in collaboration with the United Nations Development Program in 1990 (UNDP 1990). Nowadays the most popular way of operationalizing capabilities is the Multidimensional Poverty Index (MPI), which is based on the counting approach and is developed by Alkire and Foster (2007, rev. 2009b, 2011a) in 2007. This latter initiative is the turning point in contemporary poverty studies and notably in international poverty comparisons; the MPI has significantly influenced and promoted the passage from monetary to multidimensional poverty. The MPI has been recognized as one of the global tools for measuring poverty worldwide through both academic and non-academic research. Since 2010, the MPI has become the official poverty measure of the United Nations in the annual human development reports (UNDP 2010). The (composite) index methodology such as the MPI aggregates different dimensions and indicators and generates the overall (aggregate) poverty measure. For instance, the traditional headcount ratio yields the proportion of poor people in society. Other indices such as intensity and severity of poverty bring additional information to the headcount ratio (e.g., MPI corresponds to the headcount ratio adjusted by the intensity of poverty). Quantitative studies centered on composite indices are often number oriented and hereby yield rough information on the magnitude of poverty in a given society. Notably, overall (aggregate) poverty statistics appear at first glance interesting and easy to understand for public and policy makers who are not necessarily specialized in statistics. However, such practice is in fact not very rich regarding description and prediction of poverty shapes. For determining adequate and efficient social policies of reduction and elimination of poverty, detailed sub-analyses as a function of different types of capabilities and by different social groups in a given society is the crucial step. The MPI based poverty studies produce some sub-group analyses (e.g., by urban vs. rural, age groups), inform about the contribution of each indicator and of each dimension to the overall poverty rate. But the MPI methodology partitions society into two groups via the poverty line (i.e., the poor who are below the poverty line vs. the non-poor who are above the poverty line), and the head count ratio, intensity of poverty, and severity of poverty are computed based on the censored data that include only people below the poverty line. On the other hand, other poverty studies based on fuzzy set, partial order theory do not only produce headcount ratio but compute also poverty degrees for both above and below the poverty line.

In this thesis we propose a new way (methodology) of operationalizing capabilities for measuring multidimensional poverty in upper-middle-income countries. For improving the capabilities-based quantitative poverty analyses, we will employ Breiman et al.'s (1984) classification trees method, namely the *pattern recognition*[3] approach from the field of data mining and machine learning. We propose this application as a new alternative in the field of

[3] We also note that pattern recognition approach is rarely used in social sciences despite its successful employment in medical, ecological, financial analysis (see the chapter of methodology).

capabilities to measure multidimensional poverty. We assume that there is no single way of experiencing poverty and that poverty is not a specific phenomenon happening only for who remain under the poverty line. A person meets different types of poverty depending on circumstances in life and the combination of poverty and non-poverty situations is most likely to happen in the life of individuals (e.g., a person may be poor regarding access to education and be not poor regarding other aspects of life at a given time). To get a comprehensive picture of living conditions and to evaluate capabilities of people (i.e., whether the person is able or not to live the life he or she values and he or she has right to value), we argue that it is crucial to take in consideration simultaneously different aspects of life including achievements vs. non-achievements as a result of personal (micro-level) activities and actions, and advantages vs. disadvantages the person gets because of social environment, nature conditions, innate conditions. Differently from other existing multidimensional poverty analyses in the literature, we don't use the poverty line to divide society into two groups (poor vs. non-poor). In addition, we don't prepare a particular list of basic needs/basic capabilities before starting the analysis[4]. The national household surveys, which we will use in the analysis, include many aspects of life, and thus we are able to create multiple indicators to reflect different living conditions emerging in the society. The current multidimensional poverty studies select and define a particular list of indicators based on arbitrarily criterion, or on expertise such as human rights framework, or on the results of former poverty studies, or on correlations between indicators such as principal component analysis (Alkire et al. 2015; Alkire 2007a, 2007b). The selection of indicators in the sense that each single indicator should reflect a particular aspect of life for avoiding conceptual overlapping (i.e., not including the indicators meaning/addressing the same thing) is of paramount importance, but it is also crucial to maintain the diversity of indicators. Poverty is intangible; we cannot smell, we cannot touch, we cannot taste it. We can see some aspects of poverty such as dwelling roof or exterior wall conditions, street conditions, but we cannot see every aspect of poverty in one picture. There are also some hidden situations, which cannot be captured directly at first glance. Thereby indicators become momentously significant for showing poverty situations. In other words, for reflecting multifold aspects of life, we need to incorporate many indicators as much as possible in the analysis by respecting the condition of non-overlapping.

Given that (i) an indicator reflects a functioning (i.e., doings and beings); (ii) the combinations of functionings correspond to the actual capabilities of the person (i.e., the set of possibilities, means, sources, states at a given time *t* and the results of previous actions, activities); (iii) the outputs of functionings are achievements, non-achievements, advantages, and disadvantages, we aim to discover the combinations of achievements, non-achievements, advantages, disadvantages that people have at a given time. These combinations will reflect the different living conditions existing in society. We note that non-achievements and disadvantages are supposed to reduce the capabilities of the person. Better living conditions signify greater capabilities for living the life the person values. Greater capabilities mean greater real liberty and greater real liberty means less poverty according to Sen's framework. In other words, the discovery of actual capabilities patterns will demonstrate different *poverty configurations* existing in society at a given time. However, because of many indicators it becomes very complex and almost impossible to interpret many combinations. Thus, we need a statistical method capable of dealing with complexity, and able to show *the most relevant* combinations.

[4] We note that as a function of the steps in our methodology, we will use the indicators.

In line with this purpose, we will measure multidimensional poverty by employing the classification trees method of Breiman et al. (1984)[5]. By combination, we don't mean mathematical aggregation of indicators scores (e.g., sum) but we mean taking in consideration at the same time the indicators for instance, the person has access to medical services, is going to school, is able to buy the books he or she needs, lives in a family of five people, lives in a dwelling having access to drinking water, sanitation facilities, electricity, and experiences air and noise pollution. Such combination in fact describes the life situation or the ensemble of living conditions of the person at a given time as a function of the indicators included in the analysis. This type of description is also known as the profile of the person at a given time.

For obtaining a comprehensive look at poverty, we will detect the most relevant combinations for different poverty groups, namely from the least poor to the poorest people at a given time. Different poverty groups will be simply based on the total number of deprivations (i.e., non-achievements and disadvantages) experienced by people through all indicators from different aspects/dimensions of life. We will therefore show the most relevant configurations of poverty manifesting in society among different poverty groups and the results will demonstrate how the least poor group differ from the poorest group or how the other in-between groups differ from the least poor and from the poorest groups in terms of capabilities. We will see if there are some common deprivations, common capabilities patterns between the groups and what are the major discrepancies regarding non-achievements and disadvantages between the groups. Looking at poverty configurations in the entire society will also inform about the inequalities between the groups in society. As we have mentioned, current poverty statistics aim at computing the overall aggregate headcount ratio such as the multidimensional poverty index (i.e., the number of poor people 'people below the poverty line' in society) and do not show the nature of poverty. But in this thesis, we will provide *the multidimensional poverty configurations index (MPCI)*, which will show both the content and the frequency of each relevant configuration of poverty by different poverty groups.

Our methodology to measure multidimensional poverty consists of four steps. At first step, as we have mentioned we will apply the classification trees method to detect the most relevant configurations of poverty in terms of capabilities. The other steps are for enriching the analysis: the second step focuses on the dimensional analysis in the sense that the dimensional profiles (e.g., educational profile, health profile, financial capacity profile) are extracted within each poverty group; the third step extracts for each poverty group the profiles composed of the indicators that are not used at first step (i.e., the indicators that are addressed to specific target groups such as economic activity, grant receivers, and the indicators of geographical situation, and of household composition); and the fourth step interests in the profiles of specific groups such as elderly, disabled, unemployed. To extract profiles (see Fattore et al. 2011b, 2012; Fattore and Maggino 2014, Fattore 2016 for the meaning of profile) at second, third, fourth steps, we will use a simpler method compared to classification trees because there are less indicators in these steps compared to the first step. The full information on these steps will be given in the methodology chapter.

[5] Via the package Rpart in statistical software R, see Therneau and Atkinson (1997).

As the target group, in this thesis, we specifically focus on *upper-middle-income* countries[6] to measure poverty as a multidimensional phenomenon. By working at multidimensional level, we also want to emphasize the limits of monetarily approach to poverty. There are three main reasons behind the choice of the target group 'upper-middle-income countries' in our study:

First, upper-middle-income countries are more heterogeneous compared to other developing countries in terms of different types of deprivation. This heterogeneity is a suitable way of showing the complexity and the diversity of poverty. Upper-middle-income countries, notably those in the G20 group[7] (e.g., Brazil, Turkey), are closer to high-income countries regarding some spheres of life. High-income countries are more likely characterized with diverse well-being and high-tech materials related deprivations; whilst lower-income countries grapple with the lack of basic capabilities such as food, drinking water, medicines, infrastructure, etc.

We assert that upper-middle-income countries experience a larger variety of poverty witnessing both deprivation regarding basic needs on one hand, and regarding material deprivation and more elaborated well-being issues on the other hand. Accordingly, there are at the same time the following main profiles: (i) who do not experience neither accessibility problems to basic needs/subsistence nor material deprivation and live in very good dwelling, environment conditions benefiting advantaged positions in labour market, education and health services; (ii) who do not experience accessibility problems to basic needs/subsistence, live in good dwelling, environment conditions benefiting some advantaged positions but slightly suffer a certain number of material deprivation or are in difficulty to pay private education/private school and/or private health system depending on context; (iii) who slightly suffer deprivation regarding a certain number of basic needs/subsistence in terms of quality rather than quantity depending on context and live in mediocre dwelling, environment conditions with less access to private education, to private health system, and to more elaborated high-tech, electronical materials; (iv) who suffer deprivation regarding a certain number of basic needs/subsistence notably in terms of quantity and live in relatively not good dwelling and environment conditions with not at all access to private education, to private health system[8] and have difficulty to buy elaborated high-tech, electronical materials; (v) who completely or very severely suffer access to basic needs/subsistence. Furthermore, combination of profiles such as mix of some attributes of profile (ii) with some attributes of profile (iii) are also part of diversity of lifestyles, which makes even more complex the comprehension of poverty. For instance, people from lower-middle or lowest classes live in bidonvilles but at the same time obtain a certain variety of materials such as mobile phone, television, furniture; or, people from middle-class live dependent on social allowances or work with no social security and live at the same time in good dwelling conditions with access to high-tech materials, furniture; or, people from high or

[6] According to World Bank's income-oriented classification. Please refer to the link below for further information (accessed 6 August 2019): https://datatopics.worldbank.org/world-development-indicators/stories/the-classification-of-countries-by-income.html. Upper-middle-income countries have economies a GNI per capita (2017) between $3,896 and $12,055.

[7] "The G20 is the premier forum for international economic cooperation. G20 members represent around 80% of the world's economic output, two-thirds of global population and three-quarters of international trade" (Source, accessed 5 January 2020: https://g20.org/en/about/Pages/whatis.aspx). More information on the G20 (accessed 5 January 2020): https://dfat.gov.au/trade/organisations/g20/Pages/g20.aspx

[8] Except scholarship.

upper-middle classes live in very good monetarily and material conditions with a large variety of advantages but suffer at the same time social pressure, anxiety.

Second, in upper-middle-income countries, despite upper and upper-middle-income classes have a substantial level of material and monetary prosperity they also face security, democracy, human rights, corruption related problems affecting the whole country. From this point of view, it will be interesting to find out capabilities' patterns of both economically well-off and various vulnerable social groups to see which configurations of poverty, or in other words, which combinations of non-achievements, achievements, advantages, and disadvantages emerge as the most important ones in the society. The patterns information will strengthen the multidimensional perspective on poverty and will demonstrate the insufficiency of monetary aspect.

Third, compared to other developing countries, the access to macro level indicators and/or national micro data sets is easier for some upper-middle-income countries, but especially good quality of data and rich socio-economic conditions surveys provided by the national statistical offices have played a crucial role in determination of the countries we have studied in empirical part of this thesis.

In line with our previous explanations, this thesis aims one theoretical, one methodological and three empirical contributions. At *theoretical* level we will typologize functionings (see Sen 1981, 1984, 1987a, 1999, 2009) from macro and micro level prospects. In accordance with the core of capabilities, our study deals with how to better apprehend and unveil real liberty which in turn defines poverty situation of the person. This goal innately requires a comprehensive understanding of individual's living conditions. At this point, through our typology we will explain in more detail the components of capabilities of the person reflecting his or her living conditions including social context. We mainly distinguish (i) micro level functionings as the actions, activities done by the person for him or herself and the states as a result of the micro level functionings; (ii) macro level functionings as the actions, activities done by a group of people for society, community and the states of the society (i.e., societal issues such as inflation rate, safety level in the country); (iii) mixed macro-micro level functionings as the actions, activities done by a group of people or by one person for another person, or the actions, activities done by one person for a group of people or for another person. We categorize the outputs of micro level functionings as achievements or non-achievements and the outputs of macro level functionings, the outputs of environmental functionings (i.e., nature conditions), the outputs of innate functionings (i.e., talent) as advantages or disadvantages. The main reason behind this typology is to emphasize the role of social context in defining the capabilities of people. The term 'achievement' sounds more a personal success and 'non-achievement' as a personal failure; however, people don't live alone, and complex interactions define the social environment and capabilities of the person. In the original capabilities framework, social context is considered as the indirect factor of capabilities; however, we will explain that different types of social bonds that the person has with his or her entourage in daily life (via different types of contacts such as close direct relationship, or indirect relationship through media) as his or her social attachment profile shapes his or her capabilities. This link between social bonds and capabilities will be based on Paugam's (1991a, 2005) framework on social disqualification, social attachment and different forms of poverty. And, in practice the macro level and the mixed macro-micro level functionings we will mention in our typology of functionings (chapter 2) will show the place of social bonds in the capabilities approach.

To clarify the relationship between social bonds and capabilities, we give two short examples: first, the person might find herself receiving social help and become dependent on social assistance over time due to labour market conditions. This dependency puts the person in 'assisted status', which decreases real liberty/capabilities of the person. In such case, information on the nature of relationship of the person with labour market, social assistance, and his or her social environment (e.g., family, community) will help to better understand how real liberty of the person evolves over time or what is the real liberty/capability profile of the person at a given time. We can say that detecting the nature of relationships of the person with his or her environment (i.e., where the person lives, realizes activities in his or her daily life, social interactions the person encounters, situation of country, community, family, household, workplace, public places) will intrinsically inform about the capabilities set of the person. This process will help to define accurate and efficient social protection and other social policies aiming at providing various opportunities, sources, means, possibilities to people, with various social assistance when it is necessary by at the same time recognizing human, social, economic rights of the person without putting him or her in a vulnerable assisted status (see Paugam 1991a, 1991b, 2005, 2016 for the notions of protection and recognition). Second example: Suppose that a Formula 1 (F1) driver called Eddie would like to win a race. In such case, driver's personal abilities, skills, past achievements, intellectual and analytical capacity would not be the sole functionings to make him win the race. There is a technical team behind him, which works for defining strategy before and during the race and for preparing the car in terms of engineering work, and the driver has another colleague who attends the race as well in a separate car. The F1 driver's capability set to win the race is composed of not only micro-level functionings but as we can see there are some macro and mixed macro-micro level functionings such as engineers' work/performance for Eddie's car as a mixed macro-micro level functioning having advantages or disadvantages on Eddie's performance, the collaboration of Eddie's colleague's with the team during the test drives for the whole team as a macro level functioning having advantages or disadvantages on Eddie's performance. Eddie needs a certain solidarity but also, he will face a certain level of competition within the team and with the other teams during the race. To ameliorate Eddie's performance on the race, his capability set should be understood carefully. To be fully free to drive or to increase his chance to be fully free to drive as he would like to in terms of speed, he has to analyze his relationship with the team and other relationships existing in his work and in other social environments. These two examples are at micro level focusing on two persons' real liberty and capability set situation via social bonds. Later, in our multidimensional poverty analysis, we will have a macro level eye for explaining real liberty or in other words poverty situation in the country (we will not have information on specific goals of individuals, but we will have information on different aspects of life for the representative data). However, these two examples show the key aspects of the place of social context in the capability framework.

At *methodological* level we will show how to discover capabilities patterns of different poverty groups, namely the relevant *multidimensional poverty configurations* per poverty group based on Breiman et al.'s classification trees method (1984) and we will extract various profiles per poverty group. This methodology will engender a new way of measuring multidimensional poverty. At *empirical* level we will first present in a descriptive way the scores of various macro level socio-economic indicators reflecting different aspects of living conditions in select upper-middle-income countries, then we will summarize poverty measurement techniques applied by the national statistical offices of select upper-middle-income countries. After that, we will more

specifically focus on four upper-middle-income G20 countries, namely the Federative Republic of Brazil (Brazil onwards), the United Mexican States (Mexico onwards), the Republic of South Africa (South Africa onwards), the Republic of Turkey (Turkey onwards), and measure poverty on the basis of our methodology by employing national micro data sets[9]. We will then carry out comparative analyses between these four countries in accordance with the empirical results.

To explain the goal of this thesis, we have mentioned a number of notions in previous paragraphs: deprivation, disadvantage, advantage, achievement, non-achievement, functionings, and capabilities. We use these notions in the following way: each *functioning* represents an aspect of life such as the highest grade, the current health status, and functionings are operationalized via the *indicators*. Given that the outputs of these functionings are achievement, non-achievement, disadvantage, or advantage (see chapter 2 for more information on typology of functionings), the situation of *deprivation* corresponds to *the set of disadvantages and non-achievements* and the situation of *non-deprivation* corresponds to *the set of advantages and achievements*. According to the types of outputs, we consider that at a given time *the set of deprivations* (i.e., being deprived regarding all indicators/functionings at the same time) designates *total poverty*, *the set of non-deprivations* (i.e., being non-deprived regarding all indicators/functionings at the same time) designates *total prosperity*, and *the set of combinations of deprivations and non-deprivations* (i.e., being deprived regarding some indicators/functionings and being non-deprived regarding other indicators/functionings at the same time) designates *in-between poverty* (i.e., being in between of total prosperity and total poverty). The *total poverty* and *in-between poverty* situations show different *configurations of poverty*, which reflect *actual capabilities* of people at a given time. As we include different aspects of life via several indicators and we extract different configurations of poverty, we can say that this thesis aims at revealing *multidimensional poverty configurations*. In other words, the *capabilities patterns* indicate the nature and extent of *real liberty* of people at a given time. Greater non-deprivations mean greater actual capabilities, less poverty, and greater real liberty.

Currently, the comparative analysis of multidimensional poverty between countries is on the basis of headcount ratio of multidimensional poverty, for instance, ranking countries as a function of the United Nations Global Multidimensional Poverty Index (MPI) score. This popular ranking is annually published in the United Nations Human Development Report. Such index allows compensation of deprivations by non-deprivations and reflects only the number of multidimensionally poor people in the countries. However, such comparison is not correct. Suppose that there are two countries *X* and *Y*, and none of them is strictly better than the other. But by aggregating deprivation scores (e.g., sum up the scores), the country *X* might appear better than the country *Y* in terms of aggregate deprivation/multidimensional poverty score. Such ranking system elevates the position of some countries and vice-versa for others. In this thesis, we will fill this gap in the comparative literature by focusing on the *comparison of* countries based on the *multidimensional poverty configurations/patterns* and not focusing on the comparison of countries based on the multidimensional poverty headcount ratio. By extracting each single country's profiles, which correspond to the combinations of deprivation and non-deprivation (and not the aggregation of deprivation scores!), we compare the countries as a function of their performances regarding each available combination of deprivation and

[9] Although Argentina, China, and Russian Federation also fall within this classification, they were not included in the analysis, as available data for these 3 countries are relatively restricted in comparison to Brazil, Mexico, South Africa and Turkey.

non-deprivation. This way of comparison shows both the nature of configurations of poverty each single country experiences and the frequency of each configuration of poverty among the countries. Our empirical results clearly show that the ranking of countries on the basis of the aggregate multidimensional poverty score (e.g., the global MPI) hides information and provides inaccurate comparison of countries.

We also note that there are few common variables between the four countries at the same time, which makes less complicate the comparison analysis. But, on the other hand, less common indicators diminish the extent of comparison between the countries. It is normal that national surveys include specific indicators according to country context, but these surveys should also include some fundamental questions related to access to nutrition, sanitation facilities in terms of both toilet and bathroom, access to drinking water, access to electricity, medical coverage, access to education, access to medical service. The fundamental elements of life inspired by the universal human rights chart should be included in national general household, living conditions, socio-economic surveys. More questions about the living conditions will not only enrich the quality of national surveys but also will elevate the quality of comparisons between the countries. In our comparison analysis, we address this issue, and we emphasize that case studies (of South Africa, Turkey, Mexico, Brazil) are particularly important to show the diversity of indicators; for this reason, we do both the four separate case studies and the comparison analysis between the four countries. Moreover, we also show that in case of not many common variables between the countries, the profile analysis can be easily done instead of classification trees analysis to reveal the frequency and the nature of available configurations of poverty. For instance, if there are two common indicators, namely X, and Y, then the possible configurations are: {deprivation regarding X and non-deprivation regarding Y}, {deprivation regarding Y and non-deprivation regarding X}, {deprivation regarding both X and Y} and {non-deprivation regarding both X and Y}. The countries can be ranked regarding each configuration meaning that one country may have better performance in one configuration but may have worse performance in another configuration compared to other country. Such way of looking at multidimensional poverty will increase the transparency and robustness of rankings between the countries.

We will carry out the quantitative analysis by using the following cross-sectional micro national data sets: for **Brazil**, *National Household Sample Survey*[10] *(PNAD) 2011*, with the dimensions of dwelling situation, material situation, educational attainment, and economic activity; for **Mexico**, *Socio-Economic Conditions Survey*[11] *(MCS) 2012* with the dimensions of dwelling situation, material situation, educational attainment, health, economic activity, and nutrition; for **South Africa**, *General Household Survey (GHS) 2011* with the dimensions of dwelling situation, material situation, educational attainment, social assistance, environment, health, economic activity, and nutrition; for **Turkey**, *Statistics* on *Living Conditions and Income Survey (SILC) 2011* with the dimensions of dwelling situation, material situation, educational attainment, environment, health, economic activity, and financial situation.

Before introducing the comparison analysis, we will start off by case studies. In some surveys, we will find the opportunity to focus with a more specific eye on certain dimensions of life, such as: the dimension 'economic activity' in Brazil's National Household Sample Survey

[10] Pesquisa Nacional por Amostra de Domicilios
[11] Módulo de Condiciones Socioeconómicas

(2011), the dimensions 'social grant and services (social development, social assistance)', 'environment' in South Africa's General Household Survey (2011).

We recall that in this thesis, the focus is on conceptualization and measurement of poverty from a multidimensional perspective. Related to this purpose, we focus on quantitative analysis on the basis of micro national data sets. We also note that searching for causalities or determinants of poverty in upper-middle-income countries, comparing and evaluating poverty reduction policies from the political aspect do not enter into the scope of this thesis.

1.2 World poverty

Poverty manifests in different ways in societies ever since people exist in this world. Scholars from different disciplines at universities, international organizations, non-governmental institutions, other research institutions, and governments have attempted to measure poverty and to implement poverty policies and specific programs at both national and international levels. Generally, people perceive poverty as a negative phenomenon, but an unanimously adopted definition of poverty has not been emerged yet. On the one hand, the lack of an international agreement on the concept of poverty makes difficult poverty comparisons between countries, regions, communities, but on the other hand it generates various views on conceptualization and measurement of poverty. In the 18th century the pioneers in poverty investigation have implicitly uncovered the multidimensional nature of poverty in their studies, however monetarism has dominated the definition and measurement of poverty. Notably since the early 1990s the conception of poverty as a social phenomenon including different aspects of life has reappeared in poverty debates thanks to the United Nations Development Program (UNDP)'s annual human development reports which have emphasized the vulnerable situation of the Global South and the significantly increasing social and economic inequalities in the world. In 1990 Mahbub ul Haq and Amartya Sen, in collaboration with the United Nations Development Program, have developed the first prominent poverty and development tool, namely the *Human Development Index* (see Haq 1995; UNDP 1990). To put in evidence the importance of human development, Mahbub ul Haq mentions that "the real wealth of a nation is its people. And the purpose of development is to create an enabling environment for people to enjoy long, healthy, and creative lives. This simple but powerful truth is too often forgotten in the pursuit of material and financial wealth" (see UNDP 1990, 9). As a continuation of UNDP's powerful initiative, in 1995 the United Nations, during the Copenhagen World Summit on Social Development, has emphasized the multidimensional aspect of poverty and its various components such as "lack of income and productive resources sufficient to ensure sustainable livelihoods; hunger and malnutrition; ill health; limited or lack of access to education and other basic services; increased morbidity and mortality from illness; being homeless and inadequate housing; unsafe environments; and social discrimination and exclusion. It [poverty] is also characterized by a lack of participation in decision-making and in civil, social and cultural life. It [poverty] occurs in all countries: as mass poverty in many developing countries, pockets of poverty amid wealth in developed countries, loss of livelihoods as result of economic recession, sudden poverty as a result of disaster or conflict, the poverty of low-wage workers, and the utter destitution of people who fall outside family support systems, social institutions and safety nets" (see United Nations 1996, 57).

After the World War II (1939-1945), international organizations such as United Nations (UN onwards) and the World Bank have implemented poverty reduction programs in many developing countries through the development projects financed by either UN member

countries or other aid donors. Especially, UN's efforts for promoting multidimensional perspective on poverty (in opposition to monetary perspective) have successfully manifested in UNDP's annual human development reports, in the agenda of Millennium Development Goals (MDGs), in the applications of Human Development Index (HDI) and Multidimensional Poverty Index (MPI), and most recently in the agenda of Sustainable Development Goals (SDGs)[12]. In UN's works, multidimensional approach does not only suggest including various spheres of life but also points out human rights along with civil, political, social and economic rights. For instance, in Universal Declaration of Human Rights (adopted on 10 December 1948 by the UN General Assembly)[13], the Article 1 is "all human beings are born free and equal in dignity and rights", the Article 21 mentions the right to equal access to public service, the Article 22 mentions the right to social security, the Article 23 focuses on employment and working conditions such as the right to equal pay for equal work, the right to work, protection against unemployment, the Article 26 focuses on education such as the right to education, the Article 25 focuses on standard of living such as everyone has the right to "…including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control". The Millennium Development Goals (MDGs) determined under the leadership of United Nations Development Program (UNDP onwards) during the Millennium Summit in 2000, include universal human rights in promoting poverty alleviation programs worldwide. The MDGs focus on eradicating hunger, achieving universal primary education, promoting gender equality, reducing child mortality, improving maternal health, combatting AIDS, malaria and other diseases, ensuring environmental sustainability and global partnership for development. Moreover, that Multidimensional Poverty Index (MPI) comprises of three dimensions, namely education, health, living standards, shows the application of human rights in measurement of global poverty. And in the most recently adopted Sustainable Development Goals (SDGs), the human rights framework is extended involving additional dimensions such as clean water and sanitation, decent work, peace, justice, strong institutions, affordable and clean energy for everyone.

According to the 2016 Human Development Report (see UNDP 2016), the Human Development Index score on average among developing countries, is 0.756 in Europe and Central Asia, 0.751 in Latin America and the Caribbean, 0.720 in East Asia and the Pacific, 0.687 in Arab States, 0.621 in South Asia, and 0.523 in Sub-Saharan Africa. On the other hand, the human development index score on average among the countries with very high level of human development is 0.892. These scores indicate that Sub-Saharan Africa is the most backward region in the world. The 2015 Millennium Development Goals Report (see United Nations 2015) points several statistics out on monitoring MDGs. Related to that, some of them are: (i) extreme poverty rate (living on less than $1.25 a day) in developing countries has

[12] Please see the links below for further information on MDGs, MPI, HDI, SDGs, respectively (accessed 22 August 2018):
http://www.un.org/millenniumgoals/, http://hdr.undp.org/en/content/multidimensional-poverty-index-mpi, http://hdr.undp.org/en/content/human-development-index-hdi, https://www.un.org/sustainabledevelopment/sustainable-development-goals/

[13] Please see the link below for further information (accessed 22 August 2018):
https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf.
Please see also: (i) the International Covenant in Economic, Social and Cultural Rights entered in force in 1976 (source: https://www.ohchr.org/Documents/ProfessionalInterest/cescr.pdf. Accessed 22 August 2018); and (ii) Fukuda-Parr et al. (2015) for social and economic rights.

reduced from 47 per cent to 14 per cent between 1990 and 2015; (ii) the primary school net enrolment rate in developing regions has reached 91 per cent in 2015. Notably, Sub-Saharan Africa has achieved the best performance with 20 percentage points increase between 2000 and 2015; (iii) the share of women performing vulnerable employment amongst all employed women has declined 13 per centage points; (iv) the target of eliminating gender disparity in primary, secondary and tertiary education has successfully achieved by 2015 in developing region; (v) the number of deaths amongst children under five years has reduced from 12.7 million to 6 million between 1990 and 2015; (vi) the maternal mortality ratio has declined by 45 per cent worldwide since 1990. Notably, in Southern Asia, the ratio has declined by 64 per cent and in Sub-Saharan Africa it has declined by 49 per cent between 1990 and 2013; (vii) new HIV infections have reduced from 3.5 million cases to 2.1 million cases between 2000 and 2013. The global malaria incidence rate has diminished by around 37 per cent; (viii) the number of people having access to piped drinking water has increased from 2.3 billion to 4.2 billion between 1990 and 2015; (xi) the official development assistance (ODA) has increased from $81billion to $135 billion between 2000 and 2014[14]. Given the results published by the UN, the progress made over fifteen years back (between 1990 and 2015) can be considered important for the Global South but there is still much to do in terms of eliminating poverty in its all different forms especially when it comes to gender inequality, urban vs. rural disparities, gap between the poorest and the richest households, hunger, starvation, access to some basic needs, war and conflicts.

In the literature, researchers have criticized poverty and human development indices regarding some technical issues. To mention some of them, for instance, Mack and Lansley (1985, 39) highlight that "the concept of a minimum separates the poor from the rest of society and labels them second-class citizens. For the poor, a 'minimum' living standard is good enough but the rest of us expect far more and may indeed feel we are entitled to far more." In that vein, Valentine (1968, 12-13) criticizes binary separation of society, so that, "the primary meaning of poverty is a condition of being in want of something that is needed, desired, or generally recognized as having value. A moment's reflection on such related concepts as indigence[15], penury, and destitution makes it clear that there are variable degrees of poverty. As the primary definition implies, poverty varies in terms of recognized values. Starvation, death from exposure, and loss of life due to some other total lack of resources are the only absolute forms of poverty. Even here we are dealing with only one end of a scale that extends toward a quite indefinite opposite extreme. Everyone knows that being poor in the United States means something different from being poor in India; the same is true of America in the 1960's compared with the 1930's or another decade." Notably Cimadamore et al. (2016) emphasize the negative effect of aggregation process (e.g., indices such as Multidimensional Poverty Index) in the sense that overall poverty rates, due to compensation effects, hide various types of deprivation individuals experience[16]. For instance, aggregative indices do not reflect neither intra-household distribution (i.e., intra-household inequalities), nor multiple needs of people

[14] See full report for all results at the following link below: http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20%28July%201%29.pdf. Accessed 22 August 2018.

[15] "A person who is indigent is in need and, in most uses of the term, lacks the mean for subsistence" (see Gordon and Spicker (eds.) 1999, 81).

[16] There are also non-compensatory aggregation methodologies; for instance, see Mazziotta and Pareto (2016), Maggino (2017b, 87-114).

above the poverty line. The authors suggest focalizing on more specific studies including the needs of women, children, elderly, and disabled. Similarly, Mack and Lansley (1985, 41) express that "there are many aspects to our 'way of life' and some people may fall below an acceptable minimum in some respects but not in others. Deprivation among the poor surfaces in different ways according to particular circumstances. In addition, there are different degrees of deprivation: for some people, the deprivations they face will be relatively marginal; for others, it will affect their whole way of life". Moreover, Sen (1981, 157) indicates that "the lack of discrimination between different circumstances leading to poverty gives rise to a lack of focus in policy choice. Evaluative grossness can also distort. With the use of the headcount measure of poverty, the best rewards to poverty-removal policies are almost always obtained by concentrating on the people who are just below the poverty line rather than on those suffering from deep poverty. There is indeed a certain amount of empirical evidence that gross characterizations of poverty do lead to distortions of public policy" (see also Sen 1976).

If we look at recent monetary (absolute) poverty trends in the world, regarding the 2013 UN report for instance, the proportion of people living on less than $1.25 a day between 1990 and 2010 has diminished by more than 50 per cent in South-East Asia (from 45% to 14%), by 50 per cent in Latin America and the Caribbean (from 12% to 65) and in Northern Africa (from 5% to approximately 1%), and by almost 50 per cent in Southern Asia (from 51% to 30%). But, in Sub-Saharan Africa, the situation remains still problematic because it has only diminished by around 14% (from 56% to 48%). The World Bank has estimated that around 700 million people will be living in acute poverty[17] in 2020, and over 400 million in 2030. Notably, the region of Sub-Saharan African seems to be remaining the least developed region in the world in the future. Furthermore, the World Bank predictions have emphasized that in 2030, 81 per cent of the world population living in acute poverty will be living in Sub-Saharan Africa (see Cimadamore et al. 2016). We recall that there is still much to do for reducing poverty in the world despite some progress made in the past. Concerning the World Bank's calculation of (monetary) poverty, there are ongoing methodological discussions especially concerning the poverty line (see Pogge and Reddy 2010[18]; Reddy and Minoiu 2007, 53). Gordon et al. (2000b, 84) criticize World Bank's monetary poverty lines (e.g., $1 or $1.25 or $2 per day); so that, these lines "led to a form of apartheid between industrialized and developing countries. The problem was defined differently in these groups of countries and, as a consequence, comparison relevant to the construction of priorities of policy where either avoided or were thoroughly confused." The World Bank has determined initially the (monetary) international poverty line[19] as $1 per day in terms of 1985 US currency (i.e., including purchasing power of 1985 US currency). In 2008, the line has been revised and set at $1.25 per day in terms of 2005 US currency. As mentioned by Sundaram (2016), if the $1.25 per day line is used, the number of poor people would decrease from 1,923 million to 1,128 million between 1990 and 2010-which

[17] Acute poverty is also known as absolute poverty or destitution in the literature.

[18] Pogge and Reddy (2010, 42-54) mention that the number of poor calculated based on the World Bank's poverty lines varies significantly as a function of the threshold used. For instance, it is found that if the poverty line is set at $1.25 per day, 1.38 billion people would be considered poor, if the poverty line is set at $2 per day then 2.56 billion people would be considered poor, and if the poverty line is set at $2.50 per day then 3.08 billion people would be considered poor.

[19] Notably, monetary indicators are generally chosen for facilitating international comparisons of poverty because regarding living standards international comparisons seem becoming more difficult as national survey questionnaires do not necessarily include same items. We recall that there is no consensus on the definition of poverty items and each country may have its own preference.

corresponds to 41 per cent reduction. Another critic is that the 'dollar a day' line have yet adjusted to the US inflation rate. According to (Pogge 2016, 31-32)[20], if the $1 per day line was adjusted for the U.S inflation for between 1985 and 2005, the poverty line would reach to $1.815 per day (in 2005 US current). In this latter case, the number of poor people (living on less than $1-85 per day) would decrease from 2,698.42 million to 2,146.68 million between 1990 and 2010. This finding indicates a reduction around 20.5 per cent over twenty years, which is far from the target of 50 per cent reduction of acute poverty. Another point is if the monetary poverty line is halved, the proportion of people *not* hungry should be significantly decreased as well. Nevertheless, hunger still remains a global problem. For instance, the prevalence of chronically undernourishment[21] (in percentage) in Southern countries has reported around 24 per cent in 1990 against around 18 per cent in 2012, by the Food and Agriculture Organization (FAO onwards). The reduction has yet reached 50 per cent regarding this aspect. More specifically, between 1990 and 2012 in China, the number of chronically malnourished population has decreased from 21 million to 12 million (43 per cent reduction); in Latin America and the Caribbean, the number has decreased from 15 million to 8 million (47 per cent reduction); in India, the number has decreased from 26 million to 18 million (31 per cent reduction); in Sub-Sharan Africa, the number has decreased from 23 million to 15 million (35 per cent reduction); but in East Asia, when we exclude China, the number has increased by 20 per cent from 10 to 12 million (see Sundaram 2016, 30). The FAO, International Fund for Agricultural Development (IFAD), and World Food Program (WFP)'s 2015 numbers have showed that the three regions, namely Southern Asia (281 million people in hunger between 2014 and2016), Sub-Saharan Africa (220 million people in hunger between 2014 and2016), and Eastern Asia (145 million people in hunger between 2014 and2016) were the most vulnerable in terms of hunger.

Besides the United Nations' global efforts for a world free of poverty, many scholars have used the Washington Consensus to explain poverty and development process in Southern countries. The notion of 'Washington Consensus' has developed by J. Williamson in 1990 and supported by Washington D.C. based institutions[22] and the most powerful nation-states. The Washington Consensus has embodied the neoliberal economic policies such as structural adjustment programs and austerity policies (e.g., fiscal discipline and tax reform, balancing the national budget (i.e., balance-of-payments) for diminishing budget deficit and for providing economic stabilization, as recipes to solve the economic and financial problems in Latin American countries (see Dezalay and Garth 1998, Suter 1995). The neo-liberalism approach has focused on how to reduce the role of state (i.e., diminishing public expenses in education, health, sanitary facilitates), how to increase the role of private sector (e.g., privatization), the liberalization of trade and financial markets (e.g., 'laissez-faire' policies promoted by 'liberal' school of thought[23], suggest the free play of agents in the market-a self-regulating market), the

[20] According to Cimadamore's personal communication with Pogge.

[21] The prevalence of undernourishment (denoted as 'PoU') is "an estimate of the proportion of the population whose habitual food consumption is insufficient to provide the dietary energy levels that are required to maintain a normal active and healthy life. It is expressed as a per centage". This indicator aims at measuring the progress in SDG (Sustainable Development Goals) Target 2.1. Source: FAO (see http://www.fao.org/sustainable-development-goals/indicators/211/en/), accessed 15 June 2018.

[22] The International Monetary Fund (IMF), the International Bank for Reconstruction and Development, known as the World Bank, U.S. Treasury Department.

[23] The counter idea points out that the intervention of the state is important for protecting local agents in the country, for regulating cycles of recession and boosting economic growth.

protection of property rights (e.g., deregulation of public goods and services), to boost economic growth. In Northern countries, from the 1970s, neo-liberal policies have seen as good solutions against economic recessions and inflation. Especially, public expenses have been recognized as the main driver of inflation and unemployment, which has led to diminish the role of state in economy. Northern countries, therefore, have developed new non-traditional adjustment programs (against the central role of state in the economy) under the umbrella of free-market conception. Afterwards, due to new structure of Northern economies and notably due to increase of oil prices in 1973 and in 1979, the Global South economies (except oil producing countries) have encountered recessions and financial difficulties- which have led them to borrow massively. In the early 1980s, debt crisis has emerged due to several reasons (see Suter 1992, Suter 2012), but mainly due to deterioration in commodity prices (i.e., decrease of primary products' prices has affected massively many developing countries whose economies were significantly depending on the exportation of primary goods), due to increase of interest rate, to improvidence of banks and financial institutions in developing countries. The debt crisis of the 1980s has gave rise to specific 'structural adjustment' programs for Southern Countries, prepared by the Bretton Woods Institutions and Northern countries for helping the Global South economies (i.e., Third World countries)[24].

As mentioned by the United Nations Research Institute for Social Development (UNRISD 1994, 9), the structural adjustment programs have included the following conditions:

"

- drastic reduction of trade barriers protecting the local economy from foreign competition
- deep reduction or elimination of subsidies and price controls, which distorted internal prices for a number of goods and services
- restructuring of the financial system and weakening or removal of controls on the movement of capital
- privatization of state-owned firms
- elimination of control on private foreign investment
- reduction of the role of the state, not only in the economy but also in the provision of social services".

Neo-liberal economists have convinced that structural adjustment programs will be contributing to development and to elimination of poverty in developing countries, especially helping African and Latin American countries to recover them from the economic and financial crises occurred in the 1980s. However, the results have shown that these policies have had catastrophic impacts on social development; even the policies have stimulated poverty in the Global South such as diminution of income per capita, decrease of quality of public services,

[24] This help was for rebalancing economies which were significantly affected by high debt, and for stimulating development in developing countries. But it is worth to note that not all developing counties had employed neoliberal policies of Washington Consensus. For instance, Brazil, an upper-middle-income country regarding the World Bank classification, preferred to give the state an interventionist role in managing investments in the country, such as monitoring social protection, providing income support (e.g., cash transfers), playing important role in setting active labour market policies. This type of governance is known as 'developmental state', in the literature. Having practiced the 'developmental state' approach, China has also recognized as one of the most successful countries in eliminating absolute poverty (e.g., notably, high increase of employment opportunities accompanied by increase of inequalities).

worsening social conditions, decrease of formal employment, increase of depression and economic inequality in many countries. The SAPRI (Structural Adjustment Participatory Review Initiative) and CASA- (Citizen's Assessment of Structural Adjustment)'s investigations have determined four main results (SAPRIN 2002, 173-174): (i) reduction of domestic production leading to unemployment, and demise of small producers; (ii) decline of small farms leading to food insecurity, income reduction for small farms affecting massively rural areas; (iii) increase of precarious work, lower wages, and diminution of workers' rights and protection status; (iv) increase of poverty, notably diminution of access to public services.

Structural adjustment programs have often expressed by development sociologists to explain global economy and poverty situation in developing countries. Especially, with the emergence of globalization, scholars have also focused on modern world-system such as core, semi-periphery, periphery countries (see Wallerstein 1974), dependency theory (i.e., Prebish-Singer thesis), exploitation of developing countries, colonialism, post-war era, governance (e.g., democracy, reliability of institutions, political instability, corruption), international trade (e.g., power relationship in terms of exportation, importation, agriculture, industry, services sector), implementation of foreign multinational companies in the Global South, foreign aid, debt and financial crisis to examine causalities of poverty (at national, regional, international levels in comparative perspective) and have proposed remedies for poverty reduction in Southern countries. In this perspective, American sociologist Harold R. Kerbo (2006) has provided, in his book titled *World Poverty: Global Inequality and the Modern World System*, an excellent view on actual poverty in the world at the beginning of the twenty-first century by regional comparisons and has attempted to explain why some countries have achieved more to reduce poverty and others have not. Previously, the Nobel Prize Winner Swedish economist, Gunnar Myrdal (1970) has also focused on world poverty in a political economy perspective and has proposed some reforms for underdeveloped countries[25], notably in terms of equality, agriculture, demography, education and corruption. In another angle, Pogge (2002b) has referred to global poverty in human rights and democracy framework where he has used moral arguments in describing many dimensions of global poverty and global 'rooted' inequality, and he has pointed out injustice that developing countries have been facing in global market (i.e., the impact of common institutions dominated by western societies, the share of developing countries in the use of natural resources, the impact of historical background). Pogge has clearly showed that the current global economic system is *not* acceptable in ethical regards. Most recently, Hulme (2015) has published on global poverty in post-MDGs era dealing with governance, Sumner (2016) has published on global poverty in post-cold war era with a special focus on capitalism.

We have seen two different approaches to poverty studies: first is quantitative sociology based on the use of statistical numbers for identifying the extent and nature of poverty in the world or in some specific regions or per country se, second is sociology of development based on political, mercantile, and historical aspects in explaining poverty causes and poverty reforms. In latter approach, the analysis of impact of structural adjustment programs, the poverty explanations based on historical events, modern world-system, dependency theory, modernization theory, foreign aid, debt crisis, policy reforms and other elements of

[25] In the literature, the terms backward, underdeveloped, developing, southern are used to mention countries in the Third World, or also known as the Global South.

development theory take central stage[26]. Whilst, in former approach, elaborated statistical data analysis and methodological discussions on poverty measurement (e.g., which concept is better than others or whether the concept is adequately operationalized or not for measuring poverty, discussion on indicators, technical issues in construction of poverty index or other aggregate poverty measures) become the fundamental process for investigating poverty[27]. Quantitative analysis requires two steps: conceptualization and measurement of poverty. In the literature, several concepts such as subsistence, absolute poverty, basic needs, monetary poverty, culture of poverty, underclass, capabilities, subjective poverty are proposed by scholars. For instance, as we have seen the World Bank's poverty estimations are built on monetary perspective in opposition to the United Nation's human development and multidimensional poverty analysis. In this thesis, we will employ the quantitative approach and we will measure multidimensional poverty based on Amartya Sen's capabilities framework with the aim of discovering different configurations of poverty in South Africa, Turkey, Mexico, and Brazil.

1.3 Structure of thesis

According to the five goals mentioned in the introduction, the remainder of this thesis is composed of seven chapters (except bibliography and appendix) and twenty sections structured as follows. In chapter 2 '*Theoretical framework'*, we will explain the two approaches, 'Capabilities' and 'Multidimensional poverty', we use in this thesis to define and measure poverty. Moreover, in chapter 2: to clarify whether the capabilities framework takes absolute or relative approach on poverty, we will also give insight into the concepts of absolute poverty, basic needs, relative deprivation; and to show the link between capabilities and social bonds, we will describe the approaches of social exclusion, social disqualification, and marginalidad from the poverty literature. However, to portray in detail the evolution of conceptualization and measurement of poverty under a socio-economic perspective, we present the other approaches existing in the literature, namely 'subsistence', 'assistance', 'monetary poverty', 'culture of poverty', 'cycle of deprivation', 'underclass', 'working poor', 'vulnerability', 'poverty dynamics', 'vulnerability to poverty', and 'subjective poverty' as the extensive literature review of poverty concept in Appendix 9.1. The comparison of these approaches in terms of advantages and disadvantages is given in Appendix 9.2. In chapter 3 '*Poverty in upper-middle-income countries'*, we will show the socio-economic background of select upper-middle-income countries and the poverty measurement practices applied by their national statistical offices, and more specifically the national poor relief and poverty eradication programs carried out by the governments of the 4 countries of our study, namely South Africa, Turkey, Mexico, and Brazil. In chapter 4 '*Methodology*', we will inform about the methodology we applied in this thesis to measure multidimensional poverty. The chapter 4 includes three sections: first, we will introduce the four main elements of a multidimensional poverty measure, which are indicators, dimensions, aggregation technique, and decomposition analysis; second, we will inform about the data and indicators we employed in the analysis; third, we will show the methodological steps used in the analysis. More information on data mining, knowledge discovery, probability theory, and classification trees method are given in Appendix 9.3. In chapter 5 '*South Africa: an exemplary case study*', we will see in detail the results of our multidimensional poverty analysis in South Africa. According to the content of the methodology, the chapter 5 includes

[26] But of course, it is important to note that some developmentalist works include also descriptive statistics, or other statistical analysis as an accessory but not as a core of study.

[27] In some of this type of studies, poverty reduction/elimination programs and reforms are also discussed in a developmentalist view.

the following five sections: the quality report of variables as the introduction to the analysis, detecting multidimensional poverty patterns, detecting dimensional poverty profiles of poverty groups, detecting other various profiles of poverty groups, and detecting multidimensional poverty profiles of specific groups. The case studies of Turkey, Mexico, and Brazil are shown in Appendix 9.4, 9.5, 9.6, respectively. Due to the length of thesis, technical details including the codes used in the statistical software R for the classification analysis and robustness tests are given in detail only for South Africa. Appendix 9.4, 9.5, and 9.6 include the results of multidimensional poverty analysis in Turkey, Mexico, and Brazil. We note that the same methodology is applied (except not all indicators are the same) in four case studies. The chapter 6 includes the results of the main comparative analysis, which is the comparison between South Africa, Turkey, Mexico, and Brazil at the same time. As we have mentioned, the indicators vary between the four countries because of different survey questionnaires. It is not possible to do the comparison between the four countries by using all indicators from the case studies. There are only few indicators common for making an adequate comparison between the four countries at the same time. We address to this issue in chapter 6. However, the other (two or three country) comparisons between South Africa, Turkey, Mexico, and Brazil depending on the available common variables are shown in Appendix 9.7. In chapter 7, we conclude the thesis. After the conclusion, the full bibliography including the references used in Appendix is presented. And the list of national statistical agencies of select upper-middle-income countries used as the sources of information in chapter 3, and the quality report of some variables are given in Appendix 9.8, and 9.9, respectively.

2. Theoretical framework

Social scientists have conceptualized and measured poverty from different aspects for hundreds of years focusing on monetary and non-monetary items, which has attracted both academic and political interests at local, regional, international (global) levels (see Suter et al. 2017; Shildrick and Rucell 2015 for sociological perspectives on poverty). Historically, to our knowledge, the pioneer poverty studies have realized in terms of nutrition and subsistence in Victorian England[28]. Following this perception of poverty, from the late nineteenth and early twentieth centuries, several sociological schools of thought on poverty have emerged notably in Europe and in the U.S. These schools have considered poverty as a *social problem* affecting society as a whole and not as an individual problem[29]. In this regard, comprehensive social surveys[30] and various empirical investigations have qualitatively portrayed or statistically examined the dependence on charities and social welfare, different forms of deviance as causes of poverty (e.g., being criminal, being addicted to alcohol, drugs, adopting violence, spending time as dilly-dallier, being unemployed, having scattered families). The issues of social allowance and causes of poverty have also occupied policy makers in the sense that which type of social policies would be efficient or whether it is necessary or not to design social policies to combat poverty (e.g., social welfare programs). However, there is always one central question appears often on academic and political debates, which is 'how to define poverty?' Up to today there is no agreement on how to conceptualize and measure poverty; this topic continues to attract attention of scholars from different disciplines (e.g., anthropology, economy, geography, sociology, psychology, political science) and remains as one of the important problematics on the agenda of international organizations such as the World Bank and United Nations Development Program. There are more and more new methodologies and applications from various perspectives, and the latest tendency is to analyze poverty from a multidimensional perspective (i.e., considering multiple dimensions and indicators to conceptualize and measure poverty).

Especially, from the 1960s, the prominent international organizations have stimulated the emergence of new poverty reduction/eradication programs and agendas. Within this initiative, as one of the most active international institutions, the World Bank defines poverty as a "pronounced deprivation in well-being" (Haughton and Khandker 2009, 1). Through this definition, the World Bank focuses on two main aspects: first, person's "the command over commodities" to meet needs, second, person's ability to obtain "a specific type of consumption good" (Haughton and Khandker 2009, 2). The former aspect is generally known as a monetary poverty that is identified in terms of income level or consumption patterns; for instance, a person having an income or a consumption level below the defined threshold is considered poor. This aspect is also known as the welfarist approach (i.e., focusing on 'household utility'

[28] Mestrum (2002) mentions that in fact poverty has become a substantial problem in Europe from the sixteenth century and the seventeenth century was the confinement time for marginals (for instance, see Mollat 1978, Goglin 1976, Geremek 1976). But, from the nineteenth century, social studies have revealed miserable living conditions of urban working-class population, especially in England. These urban poverty studies have significantly stimulated new poverty studies focusing on magnitude, extent and causes of poverty. In addition to British works that we will see in more detail in next paragraphs, Villermé's work on moral and physical conditions of labourers (1840) can be considered as one of the most prominent poverty studies of the nineteenth century in France.

[29] But, later, we will see that in underclass approach, poverty has seen as an individual problem.

[30] See Bulmer et al. (1991).

through household income or household consumption expenditure) where income or consumption patterns are considered as "inputs into generating utility" (Haughton and Khandker 2009, 20). But the latter aspect 'the non-welfarist approach' is a more specific matter reflecting violation of human rights, deprivation in various aspects of life such as accessibility problems to different items considered necessary in life or whether the household/individual has achieved or not a minimum standard level of consumption in various goods and services such as access to food, shelter, health facilities, educational programs. The World Bank qualifies this aspect as "measures of output, reflections of utility rather than inputs into the generation of utility" (Haughton and Khandker 2009, 20). This broader perspective of poverty is also known in the literature as multidimensional poverty, which has introduced by the 1998 Economics Nobel Prize Laureate Amartya Sen in his capabilities approach (see Sen 1987a). Sen emphasizes that in general poor people lack key capabilities to function in a way that they wish; for instance, not being able to have sufficient income, good educational attainment, good health, access to information, to public services, to political freedom.

Regarding the measurement of poverty, the World Bank points out four main targets: "(i) to keep poor people on the agenda; (ii) to be able to identify poor people and so to be able to target appropriate interventions; (iii) to monitor and evaluate projects and policy interventions geared to poor people; and (iv) to evaluate the effectiveness of institutions whose goal is to help poor people" (Haughton and Khandker 2009, 1)[31]. A good measure of poverty allowing us to create poverty profiles (of countries, regions, communities, etc.) can be a very useful tool to explore living conditions of the poor, notably for public debates but also for policy makers to determine specific poverty tackling policies (see Ravallion 1992, 1998). Poverty statistics aim to find out the extent and intensity of poverty at different levels such as in terms of gender, household composition, setting (urban vs. rural), region, geographical characteristics. The statistical figures obtained, therefore, help to pronounce the voices and needs of the poor, to detect the disadvantaged regions, communities, specific groups. In this perspective, for policy makers, statistical methods for poverty measurement become extremely important in terms of identifying the priorities for reducing and/or eliminating poverty. For instance, Haughton and Khandker (2009, 3) argue that "it is easy to ignore the poor if they are statistically invisible. The measurement of poverty is necessary if it is to appear on the political and economic agenda." Moreover, the author mentions that "institutions, including the World Bank and aid agencies, have limited resources, and would like to know how best to deploy those resources to combat poverty" (Haughton and Khandker 2009, 4-5). Researchers have been computing poverty statistics by employing general household, socio-economic national surveys at cross-sectional or dynamic levels through various poverty measurement techniques. Recently, in 2010, the United Nations Development Program has proposed the Multidimensional Poverty Index (MPI) to operationalize poverty in the context of annual human development reports for promoting multidimensional aspect to measurement of poverty, and for facilitating international comparisons of poverty across different countries in the world. However, in academic world, notably in the field of socioeconomics, a voluminous number of poverty studies reflecting different sides of poverty exist based on various quantitative and qualitative tools. For long years, poverty has seen as a lack of sufficient amount of money; but recent contemporary trends have showed that poverty has been increasingly defined and measured from a multidimensional

[31] In this regard, the World Bank has encouraged the Poverty Reduction Strategy Paper (PRSP) process, has implemented and carried out the Living Standard Measurement Survey (LSMS) to measure poverty in different parts of the world (see Haughton and Khandker 2009).

perspective. Initial works on poverty have rested on subsistence, basic needs, social welfare and assistance.

In this thesis, we use capabilities approach to measure poverty from a multidimensional perspective in select four upper-middle-income countries, namely South Africa, Turkey, Mexico and Brazil. Related to that this chapter aims at explaining in detail the capabilities and multidimensional poverty approaches, which constitute the theoretical framework of this thesis. We will start off by introducing the capabilities approach in section 2.1, and then we will continue with multidimensional approach in section 2.4. In between of sections 2.1. and 2.4, in order to discuss whether the capabilities framework takes an absolute or a relative approach on poverty, we will see in detail the approaches of 'absolute poverty', 'basic needs', 'relative deprivation'; and to show the link between capabilities and social bonds, we will describe the approaches of 'social exclusion', 'social disqualification', and 'marginalidad' from the literature. Moreover, the other poverty approaches, namely 'monetary poverty', 'culture of poverty', 'cycle of deprivation', 'underclass', 'working poor', 'vulnerability', 'poverty dynamics', 'vulnerability to poverty', 'subjective poverty' are explained in detail in Appendix 9.1. A comparison of the poverty approaches in terms of advantages and disadvantages is given in Appendix 9.2.

2.1 Capabilities

The literature shows that poverty in southern countries is generally defined by advanced industrialized countries' suggestions, opinions, or poverty is ignored or accepted as a normal phenomenon or as the lowest social class. It is easy to say that poverty is part of life and it is normal that some people are poor, and they will be always poor. This primitive approach has widely accepted by the elitist groups in both Northern and Southern countries. Notably elitist groups, for protecting their place and power in society, have attempted to spread this approach in measurement practices of poverty. Typical examples consist of defining one particular list of items saying that if a person can consume some basic items, so that this person is not poor; or of defining a certain monetary threshold of poverty such as 1USD per day saying that if a person has at least 1USD per day, so that this person is not poor. If 1USD per day is enough to not be poor, then earning more or wanting to consume more or even consuming luxury goods would not be necessary. In the literature, Townsend (1993, 28) has criticized the dominance of monetary thresholds to define and measure poverty as these thresholds are mainly arbitrary, by saying that "…for many decades poverty in advanced industrialized countries has been interpreted differently from poverty in poor countries, because established elites have sought to marginalize the concept and deny, or minimize and trivialize, the phenomenon, and because poverty is not, or not only, an arbitrary idea of its beholders".

In the 1980s, Amartya Sen (1998 Economics Nobel Prize Laureate) has introduced the approach of capabilities, to explain well-being, poverty, quality of life, social justice, gender justice issues from philosophical (i.e., moral approach) and socio-economic perspectives. During the Dewey lectures in 1984 at Columbia University (i.e., Wellbeing, Agency and Freedom), he talked about his initial works on capability (see Sen 1985), and he outlined for the first time his capability approach in 1979 at Stanford University during the Tanner lectures on human values, namely 'Equality of What?' (see Sen 1979, 1980). His main motivation was to give new direction (i) to conceptualization and assessment of human wellbeing, moving beyond Rawls' primary

goods, neo-classics utilitarianism[32] (i.e., welfare economics and social choice), economic growth[33], income, gross national (GNP) or gross domestic (GDP) product approaches; and (ii) to public debates on poverty policies[34]. He has proposed the concept of '*capability to function'*, which means "the freedom to live the kind of life that we have reason to value" (Sen 2006, 35). Capabilities include various items such as having access to sanitation facilities, having access to education, being able to avoid hunger and starvation, having access to political opportunities, personal skills, environmental factors, but also social items such as "being able to appear in public without shame and being able to take part in the life of the community"[35] (2006, 35). Sen emphasizes that the capabilities is fundamental in any social order, which should be totally fulfilled for maximum prosperity. If someone is deprived of capabilities, this means that the person does not have real liberty/ freedom. Sen considers liberty/freedom as absolute necessities that everyone needs for leading a decent life. But he adds that type of capabilities (i.e., items) may vary among societies and within society depending on life context of individuals[36].

---

[32] Nussbaum (2006, 48) remarks that the utilitarian framework (i.e., measuring well-being based on utilities) does not keep in view adaptations of the poor and disadvantaged people (i.e., adaptive preferences). For instance, in order to continue to live even in difficult conditions, people may be contented with their resources, abilities, environment in general (see also Sen 1990, 1995a, 1995b).

[33] Regarding economic growth, Nussbaum (2003) qualifies it as a bad indicator for assessing the quality of life of countries because it does not yield any information on how the poor experience their deprivation status.

[34] Sen (2003, 5) has mentioned that the capability approach opposes to the "commodity-based systems of evaluation" (e.g., Gross Domestic Product (GDP) approach, basic needs approach) and to the "utility-based assessment" (see welfare economics, which defines individual utility based on pleasure, happiness, life satisfaction. As mentioned by Robeyns (2005c, 97), "economists use utility as the focal variable in theoretical work but translate this into a focus on income in their applied work. While income generally is an important means to well-being and freedom, it can only serve as a rough proxy for what intrinsically matters, namely people's capabilities"). Sen notably has argued that in long standing deprivation process, the person may adapt herself to be happy despite severe hardships (e.g., being happy with small mercies, living in a modest way, cutting down many desires). In such case, utility-based measures may not show all deprivations. For instance, he has mentioned that most of time women in disadvantaged position get used to live with their inferior status (i.e., they adapt themselves and become satisfied) compared to men's power status in society (Sen 1990, 1995b). Based on this example, Sen has pointed out that especially the utilitarian approach fails to capture the *real* well-being of women. A woman may be happy or be satisfied with her lower status and with less opportunities; however, from a moral perspective considering her as not poor is problematic. because in such case her human rights are fundamentally violated. Unfortunately, welfarist approaches do not take into account moral based information for evaluating and assessing well-being of people. Moreover, Fukuda-Parr (2003, 311) has compared human development (capabilities based), neoliberalism (welfarist approach) and basic needs approaches from four aspects namely, 'philosophical underpinnings', 'evaluative aspect', 'agency aspect', 'development strategy'. Human development approach includes social justice as the evaluative aspect whilst the neoliberalism considers monetary situation and economic growth and the basic needs approach considers income and some basic social services (e.g., education, health) as the foundation for evaluating quality-of-life of people.

[35] Concerning shame and social participation, Sen suggests reading Adam Smith's Volume 2, Book V, Chapter II (1776, section on "Taxes upon Consumable Commodities" on pages 469-471).

[36] Capabilities is the set of different types of capabilities (i.e., real liberty, freedom). Everyone needs real liberty, freedom for living in a way he or she values by respecting each other in society. However, real liberty and freedom of the person includes many subsets of real liberty, freedom such as real liberty to go to school, to have an adequate nourishment. If all real liberties of the person are satisfied, then he or she is totally well-off.

Capabilities is fundamentally based on two notions: wellbeing and agency (see Sen 1993; 1985, 169). On the one hand, Sen considers *wellbeing* for exploring how a person functions (see 1985, 197-198, see also Sen 1993). On the other hand, the term *agency* reflects the responsibility of the person in terms of behaviours and attitudes (Sen 1985, 204), in the sense that "the people have to be seen…as being actively involved-given the opportunity-in shaping their own destiny, and not just as passive recipients of the fruits of cunning development programs" (Sen 1999, 53; see also Alkire 2008b). The conception of agency leads to pay attention to the role of freedom in people's life in making decisions. In that perspective, normally, people should be able to take their own decisions by themselves freely and should be able to choose and apply means and resources for achieving the goals they value, and they should have reason to value their acts in their life. Alkire and Deneulin (2009, 37) have mentionned that "agency is related to other approaches that stress self-determination, authentic self-direction, autonomy, self-reliance, empowerment, voice and so on".

In his framework, Sen focuses on individuals (i.e., evaluation of individual's wellbeing, and not of groups). However, he recognizes that the individual acts including valuing something depend on the interaction[37] of the person with other people in society[38]. Giddens and Sutton (2014, 144) define *interaction* as "any form of social encounter, in formal or informal situations, between two or more individuals", which involves both verbal and non-verbal communication and everyone acts as a function of other(s). Or, according to Goffman (1990 [1956], 26), "an interaction may be defined as all the interaction which occurs throughout any one occasion when a given set of individuals are in one another's continuous presence; the term 'an encounter' would do as well".

Park (1921) designates interaction in a dynamic, evolutive aspect as a most fundamental social process, which can be happen between people or between groups. Park specifies social contact as a first phase of an interaction, but also as a preparatory, conditioner element for other future phases. Blumer (1986 [1969], 7-8) criticized the way that social interaction is described in classical sociological and psychological studies at that time. He mentioned that traditional studies in these two fields have treated social interaction as "merely a medium through which the determinants of behavior pass to produce the behavior". More specifically, cultural factors such as the impact of traditions, customs, social norms and rules, values, status, prestige or other social structure related terms are generally considered as the components of social interactions. But Blumer argued that social interactions cannot be seen as "an interaction of social roles or an interaction between the components of a social system" neither as "an

[37] The *interactionism*, based on the qualitative works, was flourished in the 1950s at the University of Chicago in the United States. As Durand and Weil (1997, 243) has mentioned this field of research includes various approaches such as the comprehensive sociology of Alfred Schütz (1970), the phenomenology of Berger and Luckmann (1966), the dramaturgical analysis of Erving Goffman (1956, 1961a, 1961b, 1963a, 1963b, 1967, 1969, 1974, 1981), the ethnography of Harold Garfinkel (1967), the conversational analysis of Harvey Sacks (1992 [1964-1968, 1968-1972]), and the cognitive sociology of Aaron Cicourel (1974). In addition, the works of Edwin Lemert (1951), William L. Warner (1952, 1960 [1949]), Anselm L. Strauss (1992 [1959]), Robert Redfield (1955), George Herbert Mead (1934), Herbert Blumer (1986 [1969]), Everett Hughes (1984 [1971]), Howard Becker (1963) contributed significantly to the related literature.

[38] Related to that, Sen (2009, 246) has mentioned that "in valuing a person's ability to take part in the life of society, there is an implicit valuation of the life of the society itself, and that is an important enough aspect of the capability perspective."

interaction between the attitudes held by different people"; instead, the social interaction is in fact about the relationships between people. In this line, he explains society as a group of individuals interacting with one another. In another similar approach, Schütz (1970) qualifies daily life as not something personal, but as an intersubjective world experienced by the person, interpreted by others, or in other words as a common world shared and formed by other human beings. In this sense, he defines social interactions as interrelation that the person has with other human beings, influenced by others but also influencing others through these relationships. In a broad sense, he designates other human beings as being part of an individual's personal situation as he or she has been part of their lives too.

A person does not possess full control or full power in deciding on something in life; there is also interdependence of relations, and external events or shocks might affect the person's decision process. Through social interactions people tie some social bonds with their entourage, they become interconnected and interdependent in one way or another during their life (Paugam 2018 [2008]; 2015 [2007]; Paugam et al. 2020). Such process leads to power-based relations which may cut down real liberty of some people compared to others but also some people may become subject to obligations, strains, responsibilities implied by others[39]. According to Nobert Elias's understanding (1970, 1987), people can never be seen on their own (i.e., individuals do not live isolated), and there is no make sense to study people as an 'individual'. Similarly, Charles H. Cooley (1902, 183-184) has emphasized that individual and society cannot be separated, they are both part of the social reality in the sense that individual exists only by society and society exists only by individuals: he defined the *Self* (personality) as a function of the relationships of the person with others, and the *Self-image* ("*the looking glass self*") as a function of the views of others toward the person[40].

Some scholars as Deneulin (2008), Evans (2002), Ibrahim (2006) have emphasized the importance of relations as determinants of human life. For instance, Deneulin (2014, 36) has

[39] Elias (1987) mentions that villages, cities, universities, factories, casts, classes, families, socio-professional groups, feudal vs. industrial societies, communist vs. capitalist states are the social networks formed by individuals.

[40] The social psychologist George Herbert Mead (1934), influenced by the works of Cooley, has also worked on the notion of *Self* that is, for him, not a something innate but something in progress/evolution appearing as a result of the person's societal experiences[40] including her/his relationship with other participants to this process. He expresses that the relationship of individual organisms to the whole of which they are members is similar to the relationship of individual cells of a multicellular organism to the organism as a whole. In the perspective of Mead, society is a system of communication between individuals. Mead distinguishes the social self "Me" and individual self "I" in the sense that the former is constructed based on the conducts of others (e.g., including geographical, political, family factors) whilst the latter emerges as the reaction to the "Me". In other words, he supposes that human being has a self that allows the person to be in connection with herself or in the sense that the person is able to take herself as an object that he or she can communicate with and act toward. Thanks to this self-interaction, the person would (i) examine the things existing in her/his mind; (ii) change her/his behaviors after interpreted in her/his mind the things (meaning(s) of objects, past experiences, others' behaviors), and perception on herself. This dynamic aspect qualifies the self as a process. When a person strikes an attitude to others as "Me", he or she will then react to the attitudes of others to her as "I". The individual self "I" evinces and reacts to the social self "Me". The person adapts herself to the conditions of environment where he or she lives and gives reaction to this context. So that, the two components "I" and "Me" form together the *Self* as a social process.

evoked that "collecting information about individual wellbeing only, including the ability of individuals to participate in society, omits a very important aspect of human life, namely that human life is embedded in a complex web of structural relationships which do not belong to any individual as such." Related to that, as mentioned by Alkire and Deneulin (2009, 28), the implementation of agency becomes quite complex in practice because "many decisions are made by groups rather than individuals, and a person's ability to affect collective choices varies greatly in different contexts." However, despite this gap, capabilities approach has brought a new perspective to development and poverty studies by highlighting the question of freedom. For assessing social phenomenon such as poverty, wellbeing, and human development, Sen suggests that "the appropriate 'space' is neither that of utilities (as claimed by welfarists), nor that of primary goods (as demanded by Rawls), but that of the substantive freedoms-the capabilities- to choose a life one has reason to value" (1999, 74).

Concerning poverty, in contrast to previous studies in the literature, Sen did not conceive poverty neither in terms of money nor in terms of absolute or relative deprivation; but he explained it as a specific situation in which poor people do not have the possibilities, opportunities to have/make/being what they have reason to value. Particularly he has emphasized the inadequacy of monetary perspective, and the need of new instruments to "go deeper to understand what makes human life rich and human freedoms effective" (2006, 35), by reminding the sentence of Aristotle: "wealth is evidently not the good we are seeking; for it is merely useful and for the sake of something else" (1980, book I, section 5, p.7, cited in Sen 1999, 14). Before Sen, non-monetary poverty studies were focusing only on resources (i.e., money) and results (i.e., meet or not meet needs, achieve or not achieve something). However, Sen's theoretical framework (see Sen 1992, 63) has taken attention to the *real liberty* and has made a clear distinction between two notions: *achievemen*t (or called *accomplishment*) and *liberty of achievement*. An achievement corresponds to the result of having realized something successfully. On the other hand, the liberty of achievement is the real possibilities and opportunities enabling the person to achieve something (e.g., educational attainment, participation in social life) that he or she has reason to value. In some cases, the person can have all necessary possibilities and opportunities for obtaining what he or she values and the person can achieve the goal(s) by using possibilities and opportunities; in that case, the person is considered having the real liberty of accomplishing what he or she valued.

In the language of capabilities, the term 'value' has an important place because when we would like to say something that the person wants or desires to achieve, Sen considers it as something that the person values. This is why, the term value is usually used. It is important that when the person values something, this should not damage or violate human rights of others (e.g., crime, harassment). Only this type of values is taken in account in capabilities framework. Sen qualifies it as the acts and states the person '*values*' and the person has '*reason to value*'. Specifically, on the issue of values and freedom in capabilities approach, Alkire (2002) has provided a comprehensive review with some examples from southern countries. In addition, Sen uses both the terms freedom and liberty; generally, he uses the freedom for expressing liberty in general sense including everyone in society and the liberty for mentioning personal liberty (see 2009, 362).

For achieving something, people need some resources and means. For social evaluations, it is important to know which type of resources and means the person has effectively. Besides, not everyone has the same conditions for converting resources and means into achievements. The

conditions will be shaped based on opportunities, possibilities, personal abilities, socio-demographic features (e.g., age, sex), nature (i.e., climatic conditions), physiological state (e.g., pregnant women), psychological state, living standards, social interactions, external shocks (i.e., the person's environment as a whole). In this sense, all the resources, means, and other life conditions constitute the real liberty of the person for accomplishing something. Traditional wellbeing and poverty studies focus on income and resources instead of real liberty. Sen has pointed out that these studies ignore five factors (1999, 70-71): "personal heterogeneities, environmental diversities, institutional variations, differences in relational perspective, distribution within family." Thus, we can say that in Sen's framework, there are three main elements for evaluating quality of life of the person: (i) what achieved by the person; (ii) resources and means available to the person for achievements; and (iii) other conditions the person encounters in daily life. Regarding this formulation, the *(real) achievements* correspond to the first item, and the *liberty of achievements* refers to the second and third items. Meanwhile, Sen (1992, 72) touches two additional subjects concerning liberty. He has mentioned that on the one hand liberty can be the result of anterior accomplishments, and on the other hand some future resources and means can help the person to have the liberty to achieve other things in the future. In this context, he suggests inter-temporal analysis.

Sen recognizes monetary perspective and agrees that economic prosperity and real liberty are linked in some aspect. But he sees poverty as a multidimensional phenomenon and argues that economic performance cannot explain alone the multi-facets of poverty[41]. He recalls the situation of underclass people; for instance, the Afro-American have a low life expectancy in the U.S. With such example, he mentions that even the situation of the most disadvantaged people in developed countries can be similar to those in developing countries (Sen 2009, 279). Another example is that he considers that a person can be prosperous enough for paying health services fees; however, the availability of medical services will depend on many other factors such as public service, educational, and social security systems. Herewith, Sen points out the importance of observing the life situation (i.e., living conditions) of people, and the importance of how people manage their life, for understanding poverty and its non-material components. With this motivation, Sen and his colleague Mahbub ul Haq, has introduced the human development index for measuring human development and poverty. The index, which is first presented in the United Nations Development Programme's human development report in 1990 includes three dimensions, namely education, health, and living standards. This initiative was the first international attempt to measure human development and poverty by using a non-economic tool.

In Sen's capabilities framework (2009, 279), the concept of liberty has a fundamental role. He mentions that the freedom of living the life the person values is the most precious thing in life. Sen often referred to the works of John Rawls -the former political philosophy professor at University of Harvard- as an important source of inspiration. John Rawls' well-known contribution *A theory of justice* is published in 1971 (see Rawls 1958). The theory of justice is known as one of the most influential theories in modern moral philosophy. Regarding Rawls, justice should be related to the idea of equity. His theory starts at an *imaginary initial equal*

[41] However, he argues in favour of considering income inequality in poverty studies; he mentions that "clearly the assessment of poverty in the space of capabilities cannot be divorce from the extent of income inequality. In this issue, he even notes that "the absolute deprivation of social capabilities depends on relative deprivation of incomes" (see Sen 2006, 36).

point, where no one knows the identity and interests of others. Everyone is equal at this imaginary initial equal point. The principles of justice should be *unanimity* determined at this initial point, which will lead to construct fair social institutions in society. In other words, the *initial status quo* defines the political system.

Rawls mainly formulates the notion of justice through two ideas (see Rawls 1971, 44): 'the initial status quo', and 'the fairness of fundamental agreements' Rawls accepts that in society people can be different in terms of culture, belief, living style, conception of life, or ideas on life; however, the initial status quo will allow them to understand these differences and, therefore, to decide on the core principles of *fair justice for everyone*. Once the principles are decided with unanimity by taking in consideration differences, then later the principles will influence the behaviours of individuals in the spirit of *justice as equity*. But it is clear that there is one major question: What will be the social contract that everyone will accept with unanimity? (for more detailed discussion, see Sen 2009). Rawls' theory of justice can be summarized as a chain of five successive conditions: the initial status quo followed by fair fundamental agreements on justice, fair social institutions, fair functioning of society, social justice, and the creation of new fair social decisions such as the legislation system.

Regarding Rawls, a consensus at initial status quo would make emerge the following principles of justice (see Rawls 1993, 347): (i) each person has equal rights in the system of equal rights for everyone based on freedom of equal basis for everyone and freedom for everyone; (ii) social and economic inequalities should fulfil two conditions under the distributive equity; first, these inequalities must be emerged in such circumstances where the opportunities are fairly open to everyone, second, they should be a kind of positive inequalities in the sense that they should be to benefit of the most disadvantaged people in society. As we have seen from the first principle, in Rawls' justice framework, the notion of freedom has a particular importance. Freedom is considered as the first principle of fundamentals of justice. The notion of freedom has also appeared in Rawls' *primary goods* concept. As the essentials for a society, the primary goods include rights, freedom, possibilities, opportunities to everyone, income, wealth, and self-respect[42]. Sen (2009, 95) has mentioned that Rawls' primary goods significantly influenced the concept of essential needs and unmet basic needs implemented for eliminating poverty in developing countries. Likewise, Sen has used the notion of liberty in his concept of capabilities; however, he does not totally agree with Rawls on this issue. In next paragraph, we will see how Sen considers the notion of liberty.

On the one hand, Sen argues that the notion of liberty should have some limits and should define some priorities. For instance, hunger, malnutrition, insufficient medical services are not acceptable; so that, the system of liberty for everyone should first of all give guarantee every citizen equal access to some items such as food, education, sanitation facilities, medical services. On the other hand, Sen argues that Rawls' primary goods can be seen as the means for achieving some purposes that the person values. Rawls was seeing these goods as the

[42] Rawls (2001, 58-59; 1982) lists the primary goods as follows: i) the basic rights and liberties (freedom of thought and liberty of conscience); (ii) freedom of movement and free choice of occupation against a background of diverse opportunities; (iiii) power and prerogatives of offices and positions of authority and responsibility; (iv) income and wealth; and(v) the social bases of self-respect. See also (accessed 20 June 2019): https://plato.stanford.edu/entries/rawls/

principles of fairness in society, and the freedom[43] (i.e., liberty) as one of these primary goods. Thus, Rawls focuses on the evaluation of equity and justice by evaluating the distribution of primary goods including the distribution of liberty. But Sen argues that this method does not take in consideration the question of how people can convert primary goods to have a pleasant life. For that reason, Sen mainly does not agree with Rawls concerning the place of liberty in the primary goods concept. Sen makes clear that looking just what a person has or not in terms of primary goods is not a sufficient solution because each person has his or her own living conditions[44]. Congenital factors (e.g., abilities, genetic diseases), environmental factors (e.g., cholera, pollution, natural disasters), or many other factors can affect the potential use of primary goods and income. At that point, Sen emphasizes the necessity of focusing on the evaluation of real liberty and taking in consideration a larger range of factors such as the environment, social and economic conditions, family life, personal skills, abilities for evaluating the real liberty of people.

In opposition to Rawls whom main concern is the equity (the liberty is one of the essential goods for providing the equity in society), Sen defines liberty as the core element for ensuring human development and social justice in society. To explain it, Sen gives some examples (2009, 97): for instance, disabled people will not have the same possibility of enjoying income and primary goods compared to non-disabled people; or, a pregnant woman will need a different diet compared to a non-pregnant woman. However, Sen highlights that Rawls' ideas were significantly important because Rawls' theory of justice called attention to human liberty, which is very important to understand human development and poverty issues. Keeping in mind Sen's ideas on liberty and his critics, we will explain Sen's framework of capabilities in next paragraphs.

As we have seen Sen focuses on real liberty of people. Normally, real liberty should give the person the style of living that he or she values. The set of possibilities and opportunities that a person possesses in her/his environment to realize effectively what he or she would have valued in life shows the nature and the degree of her/his real freedom. In this perspective, for evaluating the real liberty, Sen has proposed the capabilities approach. In this approach, the main idea is that the evaluation should not only be based on what the person has obtained but also on the process of choice, her/his possibilities of choices as well as the limits of her/his actual ability to come up with what he or she would care to get, do, or become. In other words, Sen does not focus directly on the status of being deprived or not regarding an item (e.g., having a television), but instead he suggests first determining the living activities, actions, states that the person has actually valued and achieved in her/his environment and second, he suggests assessing whether the person has or not the capability to achieve other living activities, actions, states that he or she values to accomplish in her/his future life. Sen considers (2009, 284) that an advantage of a person is lower in terms of possibilities if that person has lower capabilities to achieve what he or she considers valuable to achieve. In this sense, Sen (2009, 285) uses the term *capabilities* to reflect the real liberty of the person, as he mentions that capabilities are

[43] Sen (2009, 362) uses the notion of freedom for expressing liberty in general sense including everyone in society. And he uses the notion liberty for mentioning personal liberty.

[44] See Sen (1990, 2009) for discussion on the place of freedom and capabilities in theory of justice, see the papers of Pogge (2002a, 2010) for a critic of Sen's capability approach, see the paper of Robeyns and Brighouse (2010) for more information on both Rawls' primary goods and Sen's capabilities regarding the theory of justice. Moreover, see Anderson's article "What is the point of Equality?" (1999) for more information on the question of equality.

"the substantive freedom [a person] enjoys leading the kind of life he or she has reason to value" (1999, 87). He points out that the aim is not to propose an optimal way of organizing a society or suggest a typical good governance; the capabilities should be rather considered as an 'informational base' to compare the overall individual benefits in a society. It is important to note that the evaluation of societies, social institutions, performances of governments, or functioning of state or of any institution will be depending on the construction of capabilities. Sen qualifies this feature of his theory as its principal contribution to the literature. Regarding this issue, there is always an ongoing debate because some academicians find this part of theory quite large and borderless; however, other academicians find it very suitable for especially social analysis given there is no any specific suggestion or implication of a fixed list of capabilities. Besides, Sen argues that the list of capabilities should be determined as a function of the context of study, or of the purpose of study, or of the purpose of development agenda, programme, so on (see Sen 2004 for discussion regarding the issue of capabilities list).

As one of the fundamental elements of the approach of capabilities, Sen introduced another concept called *functioning*[45] which he defines as "the various things a person may value doing or being" (1999, 75). Functionings, as a set of potential or achieved 'doings and beings'[46] valued by the person, are the fundamental conditions for having the capability to realize an activity, an action, a state that the person has right to value and would like to achieve (e.g., living in an adequate shelter, voting in an election, being literate). Sen mentions that "the functionings are constitutive of a person's being, and an evaluation of a person's well-being has to take the form of an assessment of these constituent elements" (2003, 5). Specifically, *being* means the state that the person would like to be in (e.g., being self-confident, being well-nourished, being healthy), *doing* means the actions and activities that the person would like to do (e.g., visiting a museum, eating, voting in an election)[47]. Herewith, *doings* and *beings* represent the activities, actions, and states that the person values and most importantly has a reason to value them. Sen (1992, 76) qualifies them as the fundamentals for the existence of the person. The original idea behind the functionings comes from his perception on wellbeing. As we mentioned in the first paragraph of this section, Sen was interested in wellbeing of the person in terms of how he or she 'functions' and how he or she can 'function' regarding opportunities, and possibilities. These two events are expressed by the term 'functionings'. This is why, functionings correspond to the doings and beings representing all of them different activities and states of the person[48]. He used functionings for assessing 'what kind of a life is she [a person] leading? What does she succeed in doing and in being?' (1985, 195). Sen does not provide any specific list of functionings, instead he expresses them as 'to do or be what they reason to value' (1985, 200). The importance that he gives to personal value judgment emerges particularly here. He

[45] *Functioning* is of Aristotelian origin.

[46] What the person effectively able/achieved to do (can do) or to be (could be). So that, functionings correspond to the activities and actions valued by the person (e.g., performing an economic activity, voting in an election, travelling) as well as the state he or she valued (e.g., being self-confident, being well-nourished, being educated).

[47] Blumer (1986 [1969], 5) defines an *action* as something "consists of the multitudinous *activities* that the individuals perform in their life as they encounter one another and as they deal with the succession o situations confronting them".

[48] Sen (1993, 31) mentions that functionings "represent parts of the state of a person – in particular the various things that he or she manages to do or be in leading a life.... Some functionings are very elementary, such as being adequately nourished, being in good health, etc. ... others may be more complex, but still widely valued, such as achieving self-respect or being socially integrated".

perceives functionings as pluralist concept, he accepts that each person may function in different way meaning that there will be different types of activities to do and states to be.

Sen distinguishes 'refined' functionings as the "relevant alternatives that were available, but not chosen" (2003, 8) by the person. To explore the functionings that the person selected vs. the functionings that the person did not select, or in a broad sense different lifestyle available to the person, we need in fact a panel data in order to follow the person through some years. Or, by using a cross-sectional data, if we have information on whether the person did not use an available functioning because he or she did not want, we can at least distinguish 'refined' functionings at one point of time. However, it will be important to know why the person did not use an available functioning. This latter information can be obtained by a direct question during the interview or can be indirectly captured via the interaction between different functionings and deficits he or she has met. Moreover, Sen includes 'refined' functionings as a part of the 'intrinsic' view of freedom versus 'instrumental' view of freedom. In 'intrinsic' view, alternative functionings available to the person becomes crucial; however, in 'instrumental' view, the information on the actual functioning chosen by the person remains sufficient (the latter view is also called 'elementary evaluation', see Sen 1987). Sen recognizes that in practice the instrumental perspective is easier to be captured than the intrinsic view ("in terms of practical application, the intrinsic view is much harder to reflect than the instrumental view, since our direct observations relate to what was chosen and achieved. The estimation what was could have been chosen is, by its very nature, more problematic (involving, in particular, assumptions about the constraints actually faced by the person). The limits of practical calculations are set by data restrictions, and this can be particularly hard on the representation of capability sets in full, as opposed to judging the capability sets by the observed functioning achievements" (2003, 8). Nevertheless, he also emphasizes that 'intrinsic' approach is better to explicit the type of freedom behind an achieved functioning. For instance, he says that the person who starves because he or she did not want to eat by her/his own choice and the person who starves because of penury will have both the same deficit, being starved. But, clearly, the first one is opted for being starved and the second one is forced to being starved. The freedom behind these two events is completely different. When looking at the achievements or deficits, it becomes therefore important to distinguish if someone did an activity with her/his wish or rather he or she did not have any choice than that.

The combinations of functionings can also be seen as the indicator of which style of life that the person is free to choose from possible livings and to live actually or in her/his future. On the one hand, a person may achieve all the functionings that he or she values (e.g., total prosperity). But, on the other hand, it is possible to be in a position of not being able to achieve all or some functionings (e.g., destitution, partial poverty). The difference between these two situations is explained by real liberty. In latter case, the real liberty of the person is limited. For evaluating her/his limited position in her/his society (e.g., its disadvantaged situation), it is important to focus on the combinations of her/his functionings, including both her/his achievements and deficits, which will indicate the capabilities of that person. The combinations of functionings indicate not only to what the person ends up doing, having, or becoming, but also to what he or she is really able to do if she would like to do so. With this idea, Sen passes from the person's actual functionings to the capability of the person to function. In terms of functionings, Sen formulates capability as "the various combinations of functionings (beings

and doings) that the person can achieve" (1992, 40)[49]. If the person can successfully achieve what he or she values, then the person is considered having capability or in other words the person has got real freedom to achieve what he or she values. In a broad sense, Sen mentions that "capability reflects a person's freedom to choose between different ways of life" (2003, 5)[50]. In normal living conditions, the person should be able to live the life he or she would like to live in the sense that to be able to obtain or realize what he or she values without damaging other human beings and by respecting natural environment. Notably, Sen emphasized that the more alternative combinations of functionings the person has, the more the person will be free to choose the style life he or she values. So that, more alternatives show the greater volume of capability that the person possess, which also means that "the capability approach thus covers all dimensions of human well-being" (Robeyns 2005c, 96).

In this perspective, for instance, suppose that a person liked a watch during shopping trip in a commercial centre. However, he or she does not have a willing to buy it immediately. But he or she is in a position that if he or she wanted to buy it, she could do so. This means that she has a real liberty to buy this watch, or in other words, his or her capabilities allow him or her to buy it if he or she would like to do. This aspect of the capabilities helps also to observe the life situation of an advantaged person regarding an action, an activity, a state (i.e., doings and beings). For ameliorating the poor condition of disadvantaged people regarding an aspect of life, for policy makers, it is also important to find out the prosperous condition of advantaged people live with. This will also guide for observing the repartition of functionings in a society. For instance, to evaluate the accessibility conditions to legal services in a country, it is both important to know the capabilities of the person who would like to reach it because he or she needs it to do, and the capabilities of the person who does not need to go to legal services immediately but he or she can do easily if he or she needs it in future. Sen (2009, 292) provides another example: guaranteeing basic health insurance increases capabilities of people for taking care of their health. It is not important that all citizens benefit them immediately, but it is important to have this possibility for people need it immediately and also for people who will need it in future given the life circumstances can be quite uncertain as well. To emphasize the importance of this feature of capabilities approach, we focus on an example concerning schooling of young girls. A young girl who finished her obligatory school may want to go to university but unfortunately the rules in her society do not allow her to do. In another society, a young girl of same age does not want to pursue her superior degree but in all circumstances, she is in a position that she can do it freely if she values it. Thus, in former case, clearly, the girl does not have the liberty of choosing the style of life she wishes and values to live. As Sen mentioned (2009, 290), the idea of capabilities has significantly grasped this importance distinction since she is oriented towards freedom and possibilities. Usually, when we are not in the circumstances of another person, we don't really get into her/his life conditions; however, it is important to understand that the disadvantaged circumstances of other person may happen to everyone in an unexpected moment in life. This is why, it is important to focus on capabilities

[49] More specifically, a combination of doings and beings among many others is considered reflecting a part of the living conditions of the person, named as *functioning n-tuple*. And the set of functioning n-tuples is named as *capability set* (Sen 1985, 1992, 1999).

[50] Robeyns (2005c, 100) mentions that in early works of Sen, the capability set corresponds to the person's opportunity set in terms of combination of her/his functionings. However, there is also another point of view that defines the capability set as an ensemble of capabilities. This latter is notably employed by Martha Nussbaum in her works.

for maintaining solidarity within society and between societies in favour of universal human well-being.

As we can see, capabilites inform us on life situation of people rather than focusing specifically on consumption elements (e.g., goods, services, income). Many real examples from different parts in the world show that the money has not a unique power. People generally have tendency to think that it does because it is reflected in that way by various communications tools. However, if we look at it more carefully and in detail, we can see that real liberty or freedom in general sense cannot be bought by money. For instance, a young girl living in a monetary prosperous family would like to go to school to follow the studies she would like to do, but unfortunately other factors such as her family, traditions, war, conflict situation in her country can make it impossible at the time she wants to go to school. Another example is a young girl living in monetary poor family would like to go to school to follow the studies she would like to do, but unfortunately other factors such as lack of sufficient household income, lack of school in her living area, lack of transports, lack of daily food can make it impossible at the time she wants to go to school. In these two examples, the final situation (i.e., output, result, outcome) is the same; however, the picture behind the output differs because they have different sets of opportunities/possibilities. If we look at in terms of achievement, there is a deficit, but this angle does not give us the detailed information on their own singular context. Sen particularly mentions that for comparisons it should be focused on 'the actual opportunities that different persons have' (1985, 201) rather than focusing on the actual activities or states of persons. At this point, as Drèze and Sen mention (1989, 42), the perspective of *freedom to achieve* will inform us on the distinction between these two cases. In the example of two girls, clearly, their capabilities for both of them are limited. To understand and ameliorate their situation, we need to look at their life situation in global (i.e., her positive and negative capabilities) by taking in consideration different aspects of their life simultaneously. Achieving successfully a functioning requires an adequate extent of freedom. For making possible the enrolment to school for girls, policy makers first of all need to ameliorate girls' freedom, in other words their capabilities. Herewith, Sen points out that it is crucial to distinguish the *capability to function* and the *achieved functionings*.

In the light of the explanations in previous paragraphs on functionings, Drèze and Sen (1989, 12) define the capability of the person as "the set of functioning bundles, representing the various alternative 'beings and doings' that a person can achieve with his or her economic, social, and personal characteristics." In this regard, Tsui (2002, 72) notes that "the capability of a person is an opportunity set of bundles of functionings and not the functionings." Besides, for emphasizing the place of freedom, Sen gives another definition of it (1990, 118) as "capability reflect a person's freedom to choose between alternative lives (functioning combinations), and its value need not be derived from one particular 'comprehensive doctrine' demanding one specific way." Based on these two definitions, we can see that the capability of the person reflects her/his ability to achieve valuable doings and beings. Her/his ability is mainly determined by her/his real freedom. He does not specify the valuable capabilities, instead he lets to decide them by individuals or societies or depending on the context[51]. He (see Sen 1992, 79) differentiates capabilities such as elementary capabilities (i.e., the set of functionings

[51] Sen accepts that the capability is a broad concept that, regarding him, includes two main concerns of life: living standards and beyond living standards. The distinction between these two elements is that the latter focuses on "being socially useful and influential" (Drèze and Sen 1989, 12).

allowing to the person to avoid undernourishment, morbidity, mortality), social capabilities (i.e., achieving self-respect, participating to economic and social life in society, being active in community, social inclusion), capabilities contributing directly to well-being (i.e., bringing various possibilities into the life of the person to make thoughtful choices).

Herewith, based on capabilities, Drèze and Sen (1989, 15) defines poverty as "a severe failure of basic capabilities." Some *basic* capabilities[52] such as "the ability to meet one's nutritional requirements, the wherewithal to be clothed and sheltered, the power to participate in the social life of the community" (Sen 1982, 367-368) are considered important to reach a certain minimum acceptable level of living a decent life (see Sen 1992, 182). He argues that *basic* capabilities help significantly the person to develop herself, achieve various valuable activities and actions, increase her/his living conditions, and open new ways of life with various lifestyle options. He particularly suggests employing this category of capabilities in assessment and evaluation of poverty (Sen 1987a). For instance, Robeyns (2005c, 101) emphasizes that basic capabilities fundamentally address surviving and poverty traps in developing countries whilst a more elaborated list of capabilities can be also drawn for evaluating and assessing well-being or inequalities in developed countries. In same line with Sen, Nussbaum also emphasizes the eventual importance of basic capabilities (1999, 43) as "it seems likely, at any rate, that moving all citizens above a basic threshold of capability should be taken as a central social goal. When citizens are across the threshold, societies are to a great extent free to choose the other goals they wish to pursue." Sen mentions that combinations of functionings of basic capabilities may vary from one society to another. For instance, well-nourished, conveniently dressed and lodged, having a good health, avoiding morbidity, being socially included, not having shame to show up in society are given as examples of basic capabilities by Sen (see Sen 1992, 183). Thus, in a broad sense, Sen conceptualizes *poverty* as a *deficit of capabilities* or in other words as a lack of real liberty to accomplish the actions, activities, states that the person values. In other words, being not poor corresponds to being able to obtain having and beings that the person values in her/his life. For instance, if a person has not access to sanitary and educational services, on the one hand he or she is clearly in deficit regarding these items but on the other hand these deficits will also likely reduce her/his potential capabilities and make the person more vulnerable against external shocks and some other negative facts in her near/far future (e.g., as many studies have already shown in the literature, a high educational level is one of the capabilities for finding an accurate and well-paid job which may help the person to build a new way of living and escape poverty). Due to diverse living conditions, personal expectations, satisfactions, hopes, desires, there will emerge various forms of lack of real liberty and therefore, various forms of poverty. His contribution to the poverty literature, compared to other works, does particularly rely on the notion of 'real liberty' he has used to describe poverty in a multidimensional perspective. In the literature, often, poverty has seen as being deprived or not regarding something, or whether the person has meets or not her/his basic needs or some other life needs, or simply as not having a sufficient amount of income for leading a pleasant life at a specific time. But as we can see Sen has brought a larger perspective on poverty evaluation

[52] In fact, Sen has had intention of developing a *potential* list (but not a fixed final list!) of basic capabilities. He suggests adapting that list according to social context, he says "some of the basic capabilities (with which my 1979 Tanner Lecture was particularly concerned) will no doubt figure in every list of relevant capabilities in every society. But the exact list to be used will have to take note of the purpose of the exercise" (Sen 2004, 79).

and assessment with a special focus on democracy and fulfillment of fundamental human rights (against monetary and material view) in a dynamic process.

He also links deprivation and poverty. When he mentions 'failure of human capabilities', he describes this situation as the person's well-being is deprived. For instance, regarding him, "if a person does not have the capability of avoiding preventable mortality, unnecessary morbidity, or escapable undernourishment, then it would almost certainly be agreed that the person is deprived in a significant way", or another example is "the inability to appear in public without shame because of one's evident penury[53]."

In this paragraph we will focus on the roots of *real poverty* in terms of deprivation of capabilities. In his argumentation, Sen focuses mainly on two issues: his opposition to purely economic definition of poverty, and his disagreement on Rawls' primary goods in particular regarding the freedom. We already seen his position vis à vis Rawls' theory of justice. Now, we will more see why Sen emphasizes particularly the inadequacy of economic measures such as income as the solely indicator of poverty[54]. We know that in his perspective, observing the environment where the person lives give information on the structure of real liberty of the person got which is in turn shows the degree and nature of poverty manifesting in in the life of that person. He specifies that income itself cannot explain the entire picture of real liberty of a person. Income or consumption expenditure as proxy measures of poverty, in fact risks misidentifying the poor and multiple deprivation profiles. Sen sees income as a mean for obtaining or achieving something. At that point, it is very important to look at how income can be used and effectively converted into beings and doings that the person values and would like to get them in her/his life. And not everyone can convert income to achievements in same manner. For instance, we can consider that if a person gets at least 2 dollars per day he or she can be considered non- poor. However, the main question is not every person has the same real opportunities, environment for converting that money into her/his desires lifestyle. So, we have to consider these conversion differences for defining poverty. Monetary sources are considered as a kind of mean for obtaining desired lifestyle, by pointing out that they are one of the various possible means that a person can get them. So, why for instance income cannot be considered as the unique mean? Sen (2009, 311-313) gives four main reasons to explain it: personal heterogeneity, diversity of physical environments, variety of social climates, the differences in terms of relational perspectives. We explain these four factors in next two paragraphs.

First of all, everyone does not have the same physical characteristics. As a function of age, sex, health status the needs change, and depending on needs, the necessary amount of money for living a pleasant life will change as well. For instance, a disabled person to meet her/his needs will need probably a higher amount of income compared to a non-disabled. Or, some chronic illness requires a high cost treatment on daily basis. But also, money cannot always treat all type of illness. Even a millionaire does not find always solution to her/his illness. Second, the environmental factors such natural disasters, pollution, climate play important roles for converting income into satisfied living conditions, and functionings. For not being affected by natural disasters, or other climate factors, generally, collective effort is needed. For that reason, Sen advices to take environmental issues as a non-modifiable reality. Herewith, conversion

[53] See also Sen (1984), Drèze and Sen (1989).

[54] For instance, Ruggeri-Laderchi et al. (2003) found that in India, 43 per cent of children and more than half of adults were deprived regarding education and health although they were not monetarily poor.

process and its monetary necessities will change depending on environmental conditions where the person lives. Third, social conditions such as social backgrounds, distribution of services among districts, regions affect the life of people in many aspects. For instance, public services such as health facilities, schooling, transports, legal services are not always the same everywhere. Even within a country, depending on regions or districts in a city, the quality and quantity of services may vary significantly. In some cases, even the income won't be a solution immediately. For instance, a child living in a small village far from everything and where there are no school and no necessary transports for reaching out other school in other very far village, will be severely deprived regarding schooling even if her/his parents have enough income. Fourth, the patterns of behaviours such as social norms, social bonds, traditions, habits, conventions may vary among societies and within society. For instance, depending on milieu, the cost of social activities may differ. The parents of a student enrolled in a private school will need to spend more money for their child's social activities at school or outside of school.

Sen (2009, 312) yields another interesting example based on Adam Smith's *The wealth of nations*: "to behave in public without shame" may involve a highly cost outfit in rich society compared to less prosperous one. Smith, in his book *Theory of Moral Sentiments* (1812 [1759], 51), also pointed out gravity of being poor in moral, socio-psychologic perspective emphasizing the risk of social isolation, emotional obstacles that the poor faces: "The poor main …is ashamed of his poverty. He feels that it either places him out of the sight of mankind, or that if they take any notice of him, they have, however, scarce any fellow-feeling with the misery and distress which he suffers…To feel that we are taken no notice of, necessarily damps the most agreeable hope, and disappoints the most ardent desire, of human nature. The poor man goes out and comes in unheeded, and when in the midst of a crowd in the same obscurity as if shut up in his own hovel". On the other side, for assessing social inclusion, in some cases such as underclass or disadvantaged people in U.S, the money will not be sufficient because of ethnic, racial discrimination even a monetarily rich person may find herself in a socially excluded position. Sen also pays attention to the following two cases regarding the issue of conversion of resources such as income to real liberty: the case of invalid persons, and unequal distribution of household income among household members such as gender inequality (i.e., boys are consistently preferred to girls in resource allocation in some societies). He emphasizes that the cumulation of disadvantages makes complicated the conversion of resource; particularly for that reason, social welfare and assistance programmes had to consider carefully the question of conversion. Herewith, if we summarize, on the one hand, the necessary amount of income for converting personal sources to functionings, and to desired living style will vary significantly. On the other hand, income is not always the unique solution for compensating deprivation and obtain real liberty. Therefore, logically, determining a unique monetary poverty line does not make sense neither in scientific platform nor in practice.

In the framework of capabilities, Sen also focused on famine[55] and hunger[56] in developing countries. Sen (1981, 39) mentions that famine imply starvation, but it is not vice versa. On the

[55] It is worth to note that Ó Gráda (2010) provides an excellent review of the history of famine, in his book.

[56] Influenced by Sen's capabilities approach, the Millennium Development Goals (MDGs)'s second target aims to halve, between 1990 and 2015, the proportion of people who suffer from hunger in the world. For monitoring this aim, the three selected indicators are: (1) the proportion of the poorest quintile in national consumption, (2) the prevalence of underweight children under five years of age, and (3) the proportion of the population below minimum dietary consumption levels.

one hand, being hunger/starvation corresponds to nutritional deprivation as being in a situation of not being able to eat. If being hunger repeats persistently, it calls chronic hunger. On the other hand, famine is the acute starvation (i.e., known also as destitution or absolute poverty) causes particularly widespread death, for instance, in 1845-51 in Ireland, the potato famine killed about one-fifth of the total Irish population. Generally, starvation is perceived as due to lower consumption of food compared to usual level, or tendency of consuming less food, known as the *food availability decline* (FAD onwards). However, famine is characterized by an unexpected dramatically diminish of the level of food consumption in society or a dramatic increase of regular starvation. There are many possibilities for being hunger such as to not have enough money to buy food, or not enough food in the country (i.e., failure of production system, limited trade opportunities), not enough food in region or in local market, or not enough transportation for getting reach to food market, or no refrigerator or storage in kitchen, or not having a kitchen in the house, or insects in the house preventing to enough store the food in the house, and so on. Money can be directly linked, but in many other cases money is just indirectly related or even it is far being linked; for instance, if there is a war in the country and there is no food anywhere in the country so no matter the question of money. To understand what the origin of starvation or hunger is and to find out recipes for eliminating this human shame, it is very important to focus on capabilities of people in the situation of starvation and hunger.

In his approach on starvation and famine (1981, 1989), Sen and Drèze and Sen develop arguments to explain these two-phenomenon based on entitlements in an economic perspective. Sen's departure point in his framework relies on the argument that having food is one of the oldest property rights. In this light, the entitlement approach to starvation and famine aims to point out the "legal rights of ownership" of people (i.e., person's entitlements, see Sen 1989, 10) to get food among the commodity bundles that he or she can afford given the legal, political, economic situations in society. The presence of food in the market does not mean that all persons can consume it, even it does not mean that all persons can acquire all types of food available in the economy. For that reason, it is important to highlight the entitlements of the person, referred as "the set of alternative bundles of commodities over which a person can establish such command" (Drèze and Sen 1989, 9). The person's entitlements will depend on her/his endowment (i.e., ownership rights) which he or she will use for exchanging her/his endowment for commodities in the market. For instance, the endowment of a labourer is her/his labour power in a form of wage which he or she will exchange for commodities in the market. Drèze and Sen (1989, 10) define "the alternative bundles of commodities a person can acquire through exchange for each particular endowment are the person's 'exchange entitlement' for that level of endowment.[57]" Moreover, Sen (1981, 45) mentions that "a person starves either because he or she does not have the ability to command enough food, or because he or she does not use this ability to avoid starvation. The entitlement approach concentrates on the former, ignoring the latter possibility." Of course, a person can prefer also to be in starvation and avoid the necessary food; however, in that case, this is her/his own wish. In this perspective, starvation

[57] Drèze and Sen (1989, 10) also define 'extended entitlements' to take in consideration "the broader form of accepted legitimacy rather than legal rights enforceable in a court", such as whether the male head of household benefits of more commodities consumption in the household. Sen, particularly, suggests extended entitlements for examining intra-household distribution, inequality, or social analysis in a broad sense.

and famine correspond to be not entitled to a basket of enough quantity of food. Herewith, Sen does not extend his capabilities approach to this latter case.

In next paragraphs, we will shortly focus on four famine situations happened in the past, the Great Bengal, the Ethiopian, Sahel, and Bangladesh famines respectively, as mentioned by Sen in his book (1981). These examples will be important to understand the reasons behind the famines and to concretely get the fact that poverty is beyond a monetary issue. We can consider famines as a form of destitution. We will see that the cause of this destitution (i.e., food entitlement deprivation) is not always due to not having a sufficient income for buying food. There are multiple capabilities failure provoking famines, as delicately described by Sen. So that, we start by the 1943 Great Bengal famine (see also Ó Gráda 2010, chapter VI for entitlements discussion and for the brief history of Bengal famine). It is reported by Famine Inquiry Commission, India that around 1.5 million people were dead in India in 1943 because of Great Bengal famine disaster (however, this estimation is considered under-estimated by a member of the commission, see Sen 1981, 52; Famine Inquiry Commission India 1945, 109-110; Sen 1977). Sen, in particular, was concerned by the role of food supply and exchange entitlements as the causes of famine. The investigations widely used the FAD approach. Specifically, the substantially dramatic rice shortage is considered as the primary cause of the Bengal famine. However, Sen (1981, 58), based on the computations from the data yielded in the report of the famine inquiry commission (1945), puts in evidence that in 1943 the rice supply "was only about 5 per cent lower than the average of the preceding five years. From that point, Sen becomes suspicious regarding the role of FAD in that famine. It was, in fact, 13 per cent higher than in 1941, and there was, of course, no famine in 1941." Even, (i) the growing population taken in consideration, still the per capita numbers for rice and wheat (as one of the considerable consumed food); (ii) the available food grains and per capita available food grains; and (iii) two-year average of carry-over of old rice ending in 1943, were higher in 1943 compared to 1941 (see chapter 6 in Sen 1981 for more detail). In the end of his analysis, Sen was sure to conclude that the Bengal famine was not due to over-all food grains shortage. This is why, it was essential to look at carefully to exchange entitlements.

Regarding exchange entitlements, Sen first of all emphasizes the rural dimension of the famine. In fact, urban settings, notably Calcutta (i.e., capital of the West Bengal in India, one of the most important cities in East India), was offered by state-subsidized distribution of food process preventing them from the high food prices. Second, Sen gets the result that the famine had lower effect on peasants and sharecroppers compared to agricultural labourers and some sellers such as fishermen, craftsmen, barbers, due severest deteriorations of exchange entitlements (see Sen 1981, 69-70). When Sen (1981, 75) analysed the fluctuations of rice prices over different phases, he concluded that the 1943 famine was mainly related to "powerful inflationary pressures initiated by public expenditure expansion." Millions of people died, notably a large proportion of agricultural labourers, whereas the most important quantity of rice crop was produced in the history of Bengal in 1943. The prices were started to rise in 1942, and government and other institutions were not able to adjust people's food needs given the situation. As we have seen from this example, state's inability to arrange public distributional system of food because it could not predict correctly sharply increasing prices of food in the market caused millions of people's lives. Obviously, the capabilities of these vulnerable persons were very limited and even not sufficient for coping with inflationary situation. Government did not consider how to increase purchasing power of the disadvantaged people. The main lesson can be learned from this disaster is even for estimating correctly the increase of wages

or the subsidies to the poorest and the most disadvantaged people in such situation requires well-understanding of capabilities of people. If real liberties of people from different classes and occupations were regularly well documented, the prediction could be done in better conditions.

We saw in previous two paragraphs the entitlement approach in 1943 Bengal famine. Now, we will see the nature of capabilities deficit of who suffered of famine in Ethiopia which is occurred between 1972 and 1974 in two different regions: north-east, especially the Wollo province in 1972-3, and in southern provinces especially in Harerghe in 1973-4. According to estimates, around between 50,000 and 200,000 people were died because of the famine. The origin of famine was drought and as a result crop failure. Typically, in similar cases, the first impression is that there was not enough production of food meaning a significant decline of FAD. However, calorie consumption estimates indicated no significant decrease of crop production in 1972-3 (see Sen 1981, 92). In the famine year of 1973, there was a remarkably shortage of food in the province of Wollo. To explain it, Sen focuses on entitlement constraints (direct entitlement failure due to Wollo farmers and trade entitlement failure for labourers and services providers), and transport constraints (difficulties of transporting food to Wollo from other parts in Ethiopia). Sen argues that even the roads were very bad in Wollo especially in rural settings, the explanation of the famine is more likely due to entitlement approach. Based on the figures of National Bank of Ethiopia, it was clear that the average food prices were kind of similar as in pre-famine periods. So, there was no question of substantially high prices of food in Wollo markets.

Sen concludes that clearly there was a lack of enough available food in Wollo but it was not due to the roads, in opposite it was due to dysfunctional market command. He explains the later as follows: In fact, Ethiopian farmers get used to eat the food gown by family without involving in food exchange process in the market. So, when the drought happened massively in that region, the Wollo agriculturist was not able to do sufficient demand for food in the market. In 1973, the food output was substantially reduced, and people were unable to command the food outside of Wollo as they were lacking enough purchasing power to buy food. As a result, food prices were not high and there was no dramatical shortage of food supply in the region as a whole. The famine in this north-eastern part of Ethiopia affected massively agricultural community, especially the pastoral population. The pastoral population were mainly nomadic and semi-nomadic groups. Drought push them to move from their traditional dry-weather grazing land disrupting their animal stock. This population was mainly obtaining grain and food via the livestock. As their livestock was diminished dramatically due to drought and the market mechanism was not able to regulate the market, the pastoral population were led to starvation. Through this case, again we can see that income is not a sufficient mean for escaping poverty. Unfortunately, the real liberty of Wollo and pastoral population was limited due to inelasticity of market to unexpected situation, and the government was not in the measure of understanding the living style of agriculturist and pastoral population.

In this paragraph, we will see another type of capabilities deficit through the Sahel case. The Sahel is the transition band between the Sahara (i.e., desert) and the Sudanian Savanna (i.e., tropical rain forests) in Africa-six countries in question: Mauritania, Senegal, Mali, Upper Volta (i.e., Burkina Faso), Niger, and Chad. The drought happened in this zone between 1968 and 1973 due to low rainfall. The pastoral and agriculturist population and the related economy were affected significantly by the drought (e.g., destruction of crops, death of animals). In

spring 1972, the situation became emergency. In 1973, it was estimated around 100,000 people died because of starvation. The food and agriculture organization (FAO) figures on calorie and protein consumption per head showed that apart from Senegal, in this region the decline in food consumption was observed and the per capita food availability was also substantially decreased (see FAO production yearbook 1976; Sen 1981, 117). Similar to Ethiopian case, the nomadic pastoral population was severely affected by rapid loss of livestock because they were exchanging the animals for cereals and affected also by the emaciation of animals causing the sharp decline in prices. Sen (1981, 123) emphasizes an important point: "droughts may not be avoidable, but their effects can be." Herewith, it is very important to analyse capabilities and entitlements to diminish the vulnerability of people living in this region. He mainly suggests that instead of worrying about food availability, it is more important to focus primarily on providing adequately food entitlements via regulation of markets, organization of institutions, well examination of production and exchange systems in the market. The issue of food distribution appears significantly in most of famine cases as we have seen through previous examples. For guaranteeing in a substantially way the food entitlements, Sen (1981, 129) emphasizes the importance of focusing on how the capabilities of people regarding social security and employment protection can be increased in this region.

As a final example regarding famine, we will focus on the 1974 famine in Bangladesh which is due to floods and the excess of the river Brahmaputra. In middle of August 1974, floods became at disastrous level having substantially affected transplanted rice crops, and at the early September 1974, Brahmaputra crossed the danger line. Between July and October 1974, the rice price doubled. Afterwards the intensive floods, the famine has started and in late September the government of Bangladesh officially announced the famine (see Sen 1981, 131; Etienne 1977, 113-114). Langar khanas where the free cooked food is delivered to the people in need, are opened by the support of the government in early October. This food relief provided help to more than 6 per cent of the total population of Bangladesh, and it is officially reported that around 26,000 people were died because of famine. In 1974, Bangladesh was already severely dependant on food importation. With massive floods and the start of starvation, the country was clearly in a very difficult situation. In 1974, compared to 1973, the government had imported less food/food grains even during the famine. Especially, it is known that Bangladesh was receiving an important quantity of food aid from U.S. However, at the same time of the catastrophe, the relationship between Bangladesh and U.S. was constrained because of the commercial relationship between Cuba and Bangladesh, and also Bangladesh was severely deprived of U.S dollars and had to cancel some grain purchase order for autumn 1974.

Herewith, the food stock shortage, and the inability of Bangladesh government for food grains importation from abroad seem made vulnerable the country against unexpected natural disaster and therefore played substantial role in Bangladesh famine. At this point, Sen particularly investigated whether the famine was due to the food shortage resulting from the floods, or it was due to the general food shortage in that period in the country. Sen shows that in terms of production of rice, per capita rice output, and food grains availability, the 1974 was the peak year of production of rice in the country. Based on these figures, the famine would not be expected; however, Sen recommends looking at these numbers at regional level for checking whether a substantial decline of food availability happened or not in the districts most affected by the famine. But the numbers showed that (i) only two districts were remarkably different whereas the famine was more widespread; and (ii) the most affected districts had a positive growth of food availability. In this light, Sen concludes that the Bangladesh famine was not a

FAD famine. So, he interested in the target population of famine and exchange entitlements for explaining famine. People in the langark hanas were mainly labourers and farmers. Regarding exchange entitlements, firstly, Sen mentions that generally the massive employment happens in June-August. Unfortunately, this period in 1974 was remarked by intensive floods which in turn reduced the employment opportunities. Secondly, the rise of rice price was not only due to floods because before floods around early 1973, but rice prices were also already dramatically rising which made decrease the rice-entitlement of wage. Traditionally, people in rural areas were getting the food by exchanging or selling their own products. Shifting from payments in kind to wage system made more vulnerable some groups. As a result, famine was mainly due to less employment opportunities because of floods, and the decline in the rice-entitlement wage. Based on this case, Sen suggests not relying solely on the aggregative numbers (e.g., aggregate food availability statistics) and the FAD, but more examining in detail the entitlement system which will inform us more on capabilities of people for commanding the food and coping with the famine.

Through all examples mentioned above, Sen wanted to explain that a person's situation (i.e., he called also person's ability) for commanding food depends on various factors that we can include them under her/his capabilities such as what the person possess, exchange possibilities/opportunities surrounding the person, quantity and quality of food relief or other help given to the person, and external shocks or other factors affecting negatively the person's conditions such that he or she loses job, food, livestock, house, and so on. He emphasized that income is just one of the possible entitlements to food. For instance, if a person losses her/his job then that person can lose the opportunities and sources for reaching to food. Possible entitlements to food are part of the capabilities of the person. Food shortage, income shortage or insufficient purchasing power are the first factors coming to mind as main drivers of starvation and famine. However, income can be only useful if two conditions are fulfilled: it should be available food in the market and this market should be available to people. If someone does not have enough money to buy food, we have to focus on her/his capabilities deficits for earning money and to ask this question: what happened that this person does not have any more money? Earnings depend on what that person sells in the market, in which conditions he or she sells and at what price. We saw that the unexpected modifications on market along with no governmental support, peasants and agriculturists found themselves in unable to sell or exchange her/his production in the market. Otherwise, we saw that being part of a specific group, being vulnerable to ecological disasters, market dysfunctional, policy failure, led as well to catastrophe situations.

The famines can happen both in good or bad economic conditions. As Sen mentions (1981, 165), "in the fight for market command over food, one group can suffer precisely from another group's prosperity, with the Devil taking the hindmost." (see the case of Bengal famine). One of the heartbroken pictures in these realities is that people who contributed to food production have been the people who have suffered most from famine. They were rural labourers, farmers, pastoralists. In somewhere, as we have seen in first paragraphs of this section, a real liberty of the person as the combinations of functionings (i.e., food availability, loss of endowment, exchange entitlements, direct entitlement failure, trade entitlement failure- see Sen 1981, 163) should be carefully analysed for finding the solutions to complex social problems such as

unexpected or poorly predicted famine[58]. Sen (1981, 156) summarizes that "the main advantage of the entitlement approach rests not in simplicity, but in providing a more comprehensive account of a person's ability to command commodities in general and food in particular." Drèze and Sen (1989, 13) make a clear distinction between the entitlement of the person and the capability of the person. The latter focuses on "the set of alternative *functionings* bundles", however, the former focuses on "the set of alternative *commodity* bundles." Based on these arguments, he highlights that for public action (e.g., for diminishing hunger, famine), it is important to examine the relationship between entitlements and capabilities. However, the ultimate focus should be on basic human capabilities and in this perspective, the entitlements are for instrumental purpose for guaranteeing basic human capabilities.

Besides famine and hunger, Drèze and Sen (1989, 14) also express the questions of undernourishment and of undernutrition in the framework of capabilities. More specifically, if a person eats better or feeds more, this means that her/his 'nutrition-related capabilities' will increase. This type of capabilities covers different aspects such surviving, having a better physical condition, being able to do sport, feeling energetic for life, being more motivated, etc. But, of course, as Sen has mentioned, these capabilities depend also other numerous factors such as educational attainment level, social security, access to health care services, access to drinking water and sanitation, epidemics and disease, age, sex, weather conditions etc. In his analysis, he prioritizes the nutrition-related capabilities which are crucial to human well-being such as ability to survive. Drèze and Sen (1989, 14) define on the one side the undernutrition as "shortage of food intake" as related to having less quantity of food, or less variety of food, on the other side he defines the undernourishment as "unsatisfactory state of being" like having less energy or feeling less strong.

After reviewing the core notions of the capabilities approach including functionings and capabilities, we will continue with the typology of functionings we propose in this thesis. This typology aims at explaining functionings in terms of *time* and *macro vs. micro* aspects. Based on this typology, we will also show the mechanism between functionings and capabilities and we will define poverty.

As we have seen functionings represent various activities, actions, states (i.e., *doings* and *beings* such as going to school, voting in elections, being well-nourished) the person values and has right to value[59]. In his approach, Sen specifically focuses on the functionings that the person valued and obtained or accomplished in her/his life, which he calls *achievements*, for measuring quality of life and poverty (e.g., United Nations' Human Development Index, Multidimensional Poverty Index). However, at a specific t period that can be the present for instance, or the interview moment, the person will have the two types of functionings: the already achieved ones at t (i.e., those the person had valued before t)- which Sen employs in practice, vs. the not-yet achieved but wished for the future $t+1$ (i.e., those the person values at t with a wish of achievement after t). In this regard, we propose to call these two types of functionings,

[58] However, we don't forget to keep in mind that as Sen mentioned (1981, 164), starvation and famine can be emerged because of personal choices rather than entitlement failures. For instance, people may refuse to eat unfamiliar or some type of food that they are capable to buy.

[59] Meaning that the doings and beings that the person values should not damage the life of others, should respect the universal human rights and the nature.

respectively: **'*actually achieved functionings*'**[60] (AAF onwards) and '***valued functionings to be achieved in near or far future***' (VFTA onwards). The former ones are the outcomes of a process started before *t*. Whilst, the latter ones are the valued specific personal goals to be achieved or/and other functionings that the person values to achieve in a period after *t*; so that they are not yet the outcomes. The person will pass by a process between *t* and *t+1* to obtain or accomplish the VFTA at *t+1* (see Figure 1). In that process, the person may use all or some of her/his AAF. Therefore, the AAF become the *means* for accomplishing the VFTA. For instance, at a present time (*t*) a person defines a specific goal such as obtaining a scholarship for studying in Australia next year (*t+1*). For that he or she fundamentally needs to achieve her/his final exams. In order to be successful in final exams, he or she will use his or her AAF. So here, obtaining the scholarship is a specific personal goal and passing exams is other functioning to be achieved at *t+1*[61]. Or another example we suppose that at present (*t*) a person values to be able to work at *t+1*. For that, he or she needs to search for a job. She evaluates her environment as pretty adequate for proceeding to seeking a job. In that case, wishing of working is her/his personal specific goal (VFTA), searching for a job is her other functioning to be achieved (VFTA), and his or her environment represents his or her AAF which will be used by the person as the means for achieving his or her two VFTA.

The AAF reflect the person's actual situation in terms of valued achievements; in other words, this is her/his actual *basket* of valued and achieved functionings. We also note that some modifications can be occurred in her basket during the process time of achieving her/his VFTA. However, we can't know in advance how exactly the AAF may modify or the new functionings can be occurred over time. But based on the nature of the AAF as a reference tool, we can predict the likelihood of achievement of the VFTA. When the process is over, for instance at *t+1*, the occurred modifications in her/his basket of functionings between *t* and *t+1* can be captured. Moreover, by only focusing on the AAF at one time in a static aspect, we can get the actual living conditions (i.e., actual life situation) of the person at that time as a proxy indicator of her quality of life[62].

[60] When we say 'actually achieved functionings' at *t*, this does not mean that these functionings are happened just before *t*. Some of them may had happened for instance ten years before *t* and they are still in the life of the person.

[61] It does not mean that the date of final exams and the publication of results are realized at *t+1*. We use time periods in large sense. The exams can be for instance one day or ten days before *t+1*. But, in any case, the person has to get her/his results latest at *t+1*.

[62] We say 'proxy indicator' because we only focus on valued achievements. However, there are also non-achievements (i.e., deficits) in life.

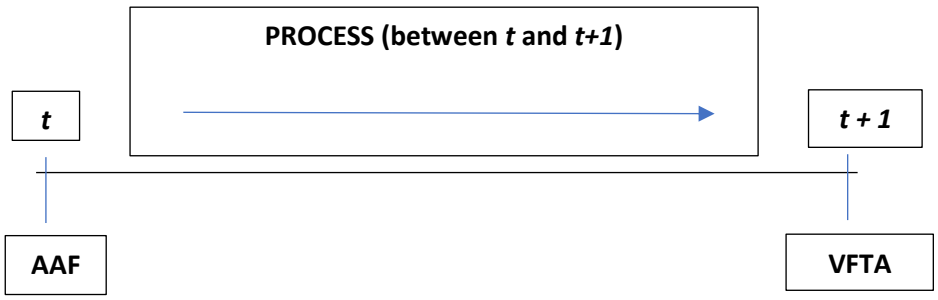


Figure 1. Actually achieved functionings (AAF) and valued functionings to be achieved (VFTA) between the two periods t and t+1.

Regarding *macro vs. micro* aspect, we distinguish the two types of functionings: the first one reflects societal functionings whilst the second one reflects individual functionings. The ***micro level functionings*** (**MIF** onwards) are effectuated and achieved directly by the person (P) only for him or herself. On the other hand, the ***macro level functionings*** (**MAF** onwards) are the societal achievements, states accomplished/created by the social environment of the person such as her/his household, family, friends, neighbor, workplace, community, region, country, in which we cannot directly observe the participation of P (e.g., a country with a low level of inflation or a community where there is an available public school, or a community where there is a low rate of crime). The outputs of the MAF are also part of the person's capabilities as individuals do not live isolated in society, they do face social facts and they are affected somewhere by societal events. People live in perpetual interaction with her/his social environment including behaviors of her/his entourage, government decisions, social regulations, social norms, traditions, globalization, and all different types of actions and activities going on in her/his actual society and in the world as a whole. The aim here is not to search for the impact of the MAF on the MIF or to predict how specifically a MAF may determine the living conditions of the person. In our framework, we integrate *social context* into the individual's life situation analysis by assuming that a human being interacts with her/his entourage/environment from social, political, cultural, economic, ecological perspectives, and we take the MAF in consideration for understanding the society in which the person lives, in other words for depicting in a comprehensive way the ensemble of living conditions/life situation of the person.

There are mainly two approaches in considering the concept of interaction in sociology: face-to-face encounters (such as analyzing body language, oral expression, gesture, facial expression) vs. intersection of different social relationships. The former one is more observable whilst the latter is generally captured via its effect on an observable phenomenon. Park (1921, see chapter one in social contact part of his book) emphasizes that touching is the most intimate contact sort whilst the visual and auditory senses (i.e., speaking and hearing/listening) during the face-to-face encounters are also part of the social contact. He includes not only these senses but also other kinds of communication tools such as language, literature, works of art, phone, radio, television, theatre, cinema in explaining different means of social contact. He argues that a person is in a social contact with other people even if he or she does not know them in person given that such contact is indirect and complex; for instance, global economy makes many people in connection, or the work of someone else in another continent may be useful for another person living in another continent, or interdependence between two people due to global

market mechanisms/division of labour across the world, or a group of people manifesting for a protection of a natural environment in another country will be indirectly in social contact with other environmental activists living in other countries[63]. In a society, people living in different regions/districts/cities/places can be in contact due to their emotional or intellectual exchanges; such societal contact may lead to a collective behavior such as organizing protests or strikes in different parts of the country for defending the same thing/for the common purpose.

Blumer (1986 [1969], 87) indicates that society is composed of "acting people" and based on that any social analysis focusing on the nature and extent of society should include the actions of individuals. He adds that "the acting unit necessarily has to identify the things which it has to take into account tasks, opportunities, obstacles, means, demands, discomforts, dangers, and the like; it has to assess them in some fashion and it has to make decision on the basis of the assessment". Blumer includes also household, family, institutions (e.g., school, company, labour union), group of people, community, organization as other types of acting units in this context. He summarizes that "any particular action is formed in the light of the situation in which it takes place". At this point, Blumer's ideas can be linked to Sen's capabilities in the sense that to evaluate better the capabilities of the person to achieve what he or she values, first we have to understand the living conditions of the person (the person will act to achieve what he or she values).

We assume that in any case environmental, personal, societal characteristics as well as individual actions and activities simultaneously play role in decision process of which functionings the person values to achieve (i.e., personal goals) and of how he or she will employ her actual functionings to achieve what he or she values or to achieve what he or she has to effectuate. So that, we propose to obtain the person's profile of her/his actual functionings at a given time as a picture of her/his life situation. For instance, the fact that a child cannot go to

[63] Park (1921) defines five main topics in discussing social contact in a sociological perspective: (i) land as a basis ground of social contacts: location of people and communities in the world gives us a spatial picture of social contact. As examples, he mentions comparisons of rural areas across different countries to learn about more different rural life styles, or different living conditions/behaviors of families from different areas such as country sides, villages, little towns, cities in the same country, or templating people's behaviors by day, week, month, or year. He argues that such examination will show the similarities and differences in terms of affinity and intensity of social contacts; (ii) tangency as a physiological ground of social contacts: touching someone is the as close as possible way of contact in spatial terms. The senses of affection, hate, desire, repulsion, sympathy, hostility emerges or becomes more or less intense in advanced phases of the contact; (iii) contacts with 'in-group' and 'out-group': good faith and solidarity between the members of in-group vs. hostility and distrust against the members of out-group; (iv) historical continuity and mobility: traditions transmitted from one generation to another in preserving the historical past and cultural life of a community. As an example, he defines the spring of cultural degeneration when the contact is split between young and old generations. He indicates discoveries, popular books, newspapers, trends, influences as some of the various tools of change and social movements; (v) primary and secondary contacts: the first one represents sincere, face-to-face encounters and the second one corresponds to external, distant relationships (e.g., primary contacts including greatest contacts (family, parents-children, wife-husband relationships) vs. acquaintance contacts (outside family, lovers, friendship). In other words, these relationships mark the development of human nature and personality. Parks especially highlights that people's reactions to other people, whether voluntary or non-voluntary, correspond to social phenomenon. As Park, Schütz (1970) has also topologized different types of communication tools between people. He has designated physical movements and activities (e.g., eating, speaking, walking), and gestures/physical expressions (e.g., smiling, crying, be shy) as the symbols comprehended by the interpreter during face-to-face encounters.

school gives us information only for her/his educational attainment status however her/his life is more than school so that we have to capture her/his actual living pattern/living profile in order to understand in what other aspects also he or she lacks her/his real liberty. Or another example, for a child to be successful in her/his school exam, a number of functionings in her/his life such as her/his psychological and physical conditions, school materials at the disposable of the person, quality of school he or she attends, safety, family relationship, household financial context, friendship will play simultaneously in child's performance before and during her/his exam process. It is clear that multiple types of social interactions emerge on our daily basis, which we cannot always detect all of them. These interactions happen very fast and as Strauss (1992 [1959]) mentioned they are too complex that we cannot draw in detail, or they are almost invisible. But at least it will be important to add social context (e.g., macro level functionings, environmental functionings) as an inclusive element into the agency-based capabilities framework to emphasize that people do not live isolated, and they are exposed directly or indirectly to the millions of social interactions during their life. Berger and Luckmann (1966) mention that the person should be continually socialized with actions and activities, due to anthropological and biological reasons.

Park (1921) argued that through different types of social contacts the person becomes connected to other people, and the person becomes what he or she is today due to existence of other people, and to her/his interaction with other people in the past and present. Park also designated social interactions between individuals and social interactions between groups as the indicators of society. Agency approach makes relevant individual's freedom to achieve what he or she values. But on the other part, social factors (i.e., person's social environment) model the extent of freedom in society as a whole, and through social interactions it will define the meaning of freedom the person has in her/his mind, and therefore how he or she perceives other things and her/his acts. For instance, Schütz (1970) considers social groups, institutions, legal, economic systems as the complementary tools to people's lives, which may modify as a function of time and space. The best way of understanding the real liberty of the person will be understanding of the person and her/his life. Therefore, in order to capture at best level, the real liberty of a person, we need to get knowledge of both the person's individual actions (i.e., micro level functionings) and societal facts surrounding the person as the reflection of his or her social contacts, and the groups to understand the nature of society where he or she lives.

In original capabilities framework, social context is used as (i) "individual conversion factors"[64] in transforming goods and services[65] into the capabilities (i.e., people get first their goods and services and later social context may influence or not the capability of using goods and services or the capability of transforming the characteristics of goods and services to achieve what he or she values); (ii) "social influences on decision making" when the person selections which combinations of her/his functionings to be used for achieving (see Robeyns 2005c, 98). These two points of views are the heritages of moral individualism and economic perspective which does not put the concept of society and its components at the heart of poverty analysis. Robeyns

[64] Sen mentions three types of conversion factors: social, environmental, and personal. Robeyns (2005a, 36) mentions that "social conversion factors are determined by a number of societal aspects", "environmental conversion factors are determined by the environment in which a person lives", and "the personal conversion factors are determined by one's mental and physical aspects".

[65] Robeyns includes "non-market production market production, net income, transfers-in-kind" as goods and services (2005c, 98).

(2005c, 99) mentions that "social conversion factors, …, and environmental conversion factors play a role in the conversion from characteristics of the good to the individual functioning". She gives an example: "if there are no paved roads or if a government or the dominant societal culture imposes a social or legal norm that women are not allowed to cycle without being accompanied by a male family member, then it becomes much more difficult or even impossible to use the good to enable the functioning". In this example, she uses bicycle as the good and cycling with the bicycle as the functioning. This viewpoint prioritizes and uses as a base the individual moralism in getting the meanings for the objects. Of course, it is of paramount importance individuals could be able to value and perform the way they value objects. But, as Blumer has mentioned the meanings towards objects does not appear in the mind of the person as an only function of her/his personal characteristics and/or actions and activities he or she achieved. Or, similarly, Mead (1934) highlighted the existence of both characteristics of objects and personal experiences in construction of an action. In fact, individual moralism can be flourished and shaped by the inclusive perpetual interconnexion with social context, so that the social life elements should be part of the set of functionings. For that reason, we propose integrating macro, environmental, mixed macro micro level functionings in the capability-based poverty analysis. In practice, some indicators of social relationship such as activities with friends, family, neighbor are included in multidimensional poverty analysis but societal indicators such as democracy level, insecurity, economic instability, corruption are rather considered as the indirect factors possibly affecting later the capability of the person to realize what he or she values. As a result, this type of social indicators is generally taken out of individual level-based poverty analysis.

Robeyns (2005c, 99) also adds that "knowing the goods a person owns or can use is not sufficient to know which functionings he/she can achieve; therefore, we need to know much more about the person and the circumstances in which he/she is living". We can see that Robeyns' argument is coherent with our way to look at capabilities as a life situation at a given time. However, there is a major concern: measuring real liberty of a person grounded fundamentally on goods and services as a consumer puts materialistic aspect at the core on the one hand and societal features at the periphery as a touch-up factor on the other hand. Such primarily individualistic approach elevates risk to show poverty as an individual problem such as not being able to access to market for buying goods and services, or not being able to employ in good way social context, or not being able to overcome disadvantaged influence of social context for achieving an action or an activity. We can say that capabilities approach is not completely against including social context factors in social analysis such as poverty but the way that they are included are either indirect or not clearly explained in practice. In parallel, Robeyns (2005a, 36) has also admitted that "not all capabilities require some good or service as an input; for example, being respected by your peers often only requires respectful behavior from other people". Sen also mentioned in his framework that social, environmental, and personal conversion factors may highlight the specific circumstances of the person such as if he or she needs additional medical treatment, or higher income due to her/his medical treatment, or additional sources given her/his social and environmental conditions. But we argue that by just focusing on the profile of the person it is also possible to detect her/his needs for achieving a specific functioning. Thus, it is not a reason to put social and environmental contexts outside of the mainstream of individual's capability set.

Suppose that a person wants to play a piano in public. For that, he or she needs to be able simultaneously to acquire a piano, have a self-esteem, have skills and necessarily education,

and a society allowing him or her to do that. Theoretically none of these criteria may have an absolute power. But in a subjective perspective, the person may feel more important to have a self-esteem as the basis and then acquire a piano as a second most important item. Or another example, elderly people may need to get a walker for achieving the functioning of walking in the streets; of course, in such case the primary necessary is to have a brain system capable of giving, perceiving, answering the command, and then having a walker can be considered a second need once the first is fulfilled. Moreover, for some activities such as looking and observing people and nature, or thinking something, we don't specially need goods and services. Depending on people's choices, goods and services may be the most substantial step to achieve a functioning but at least if we don't have information on how the person values her/his functionings for her/his specific purposes we cannot then suppose that goods and services should be the main reference for assessing capabilities of the person. It will be fair to keep equal with a same importance all different contexts of life. Why discriminate social context from the heart of capabilities and put material context as the fundamental element of real liberty? We, therefore, in our framework include directly the social context as well as ecological and person's psychological and physical conditions under the umbrella of macro, environmental, mixed macro-micro, innate level functionings besides the micro level ones in poverty analysis.

Sen clearly mentions that the nature of real liberty of the person, namely his or her capabilities, correspond to the combinations of his or her functionings. The person has a set of functionings at a given time and he or she can use all of them or some of them for effectuating her/his other functionings. Unless we don't follow the life of the person over periods or unless the person does not yield us the detailed information on which activity/action he or she effectuated by employing which of her/his functionings we cannot know the combinations of functionings that he or she used for specific purposes. For that reason, we generally obtain the information on the person's actual set of functionings at a given time, which corresponds to the combination of her/his all available functonings (i.e., the main initial combination). The actual set of functionings provide us the picture of life situation, in other words the nature and the extent of real liberty of the person at a given time. From the initial set, different sub-sets can be obtained (i.e., derived combinations); however, as we mentioned already, in practice we cannot know which derived combinations that the person exactly used in effectuating her/his actions and activities.

In Figure 2, we mention that: regarding the MIF the person (P) him for herself exclusively takes and achieves the activity or the action (i.e., doings) exclusively for him or herself (P); on the other hand, regarding the MAF, more than one people such as a group of people, a community effectuate and achieve the activity or the action (i.e., doings) for a group of people, or a social group, or a community, or a region, or the society in large sense (societal doings emerge due to complex social interactions that we may not know their departure point). When an action or an activity is undertaken, the *beings* (i.e., states) as a result of the doings are emerged. For instance, if a child gets daily adequate nutrition diet (i.e., doing), then we can say that the child is well-nourished on her/his daily basis (i.e., being). However, it is important to note that in some cases (as a result of interview, or survey) we will have a full information on both the nature and results of actions and activities, or in other cases we will have only information on the beings (and not on the nature of the doings, without knowing which action or activity is exactly taken for and by who for achieving the being).

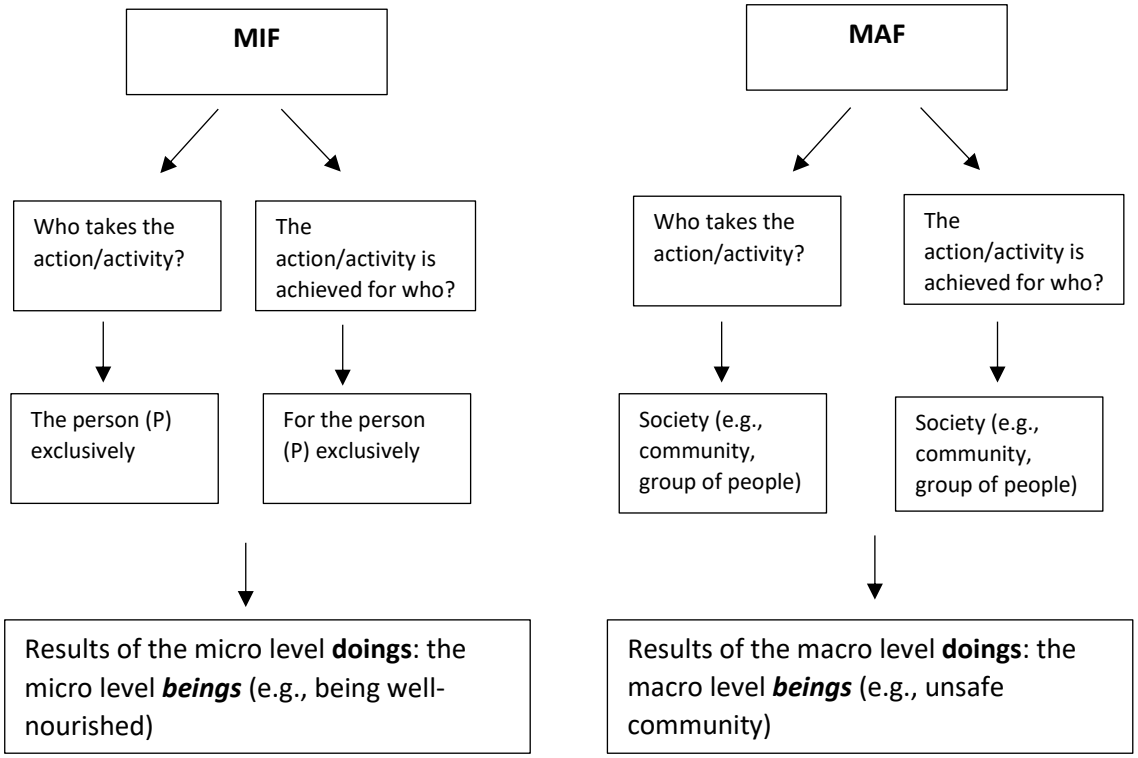


Figure 2. The micro level functionings (MIF) vs. the macro level functionings (MAF).

Besides there are also functionings corresponding to a mix of macro and micro aspects, called ***mixed macro-micro level functionings*** (MACMICF onwards) as follows:

(**MACMICF1**): an activity or an action (i.e., doings) is performed by both the person (P) him or herself and other person(s) from P's social environment[66] for other(s) (e.g., P effectuates some tasks as a voluntary in collaboration with other(s) for someone else's farm such as mutual help in farming community).

(**MACMICF2**): an activity or an action (i.e., doings) is performed by both the person (P) him or herself and other person(s) from P's social environment for both him or herself and other(s) (e.g., P participates in organizing a manifestation as a social movement with many other people).

(**MACMICF3**): an activity or an action (i.e., doings) is performed by both the person (P) him or herself and other person(s) from P's social environment, specifically for P (e.g., P's elderly sibling helps P to do her/his homework).

[66] Here, the term *social environment* is employed here in large sense. It means the society where the person P lives.

(**MACMICF4**): an activity or an action (i.e., doings) is performed by the person (P) him or herself for other(s) (e.g., P takes care of elderly in the household, or a child labour working for her/his parents).

(MACMICF5): an activity or an action (i.e., doings) is performed by the person (P) for both him or herself and other(s) (e.g., P cleans alone the entire house for the benefit[67] of all people including P living in the household).

(**MACMICF6**): an activity or an action (i.e., doings) is performed by other(s) specifically for P (e.g., P's elderly sister prepares the breakfast of P in the morning before P wakes up and goes to school).

The results of the mixed macro-micro level doings correspond to the mixed macro-micro level beings. For instance, from the example of the MACMICF6, we can say that P finds him or herself (i.e., P's state) prepared her/his breakfast by her elderly sister in the morning before going to school. From P's point of view, it is a MACMICF6. However, from P's elderly sister point of view, as she prepares the breakfast for her little sister/brother it is a MACMICF4.

After typologizing functionings, we will show the relationship between functionings and capabilities. In his approach, Sen focuses solely on achievements to explain capabilities of the person in a sense that more variety of achievements means more capabilities for achieving more functionings and for having a more variety of lifestyle. A large set of capabilities reflects a higher level of real liberty for the person to lead the life he or she values[68]. In a dynamic perspective, various combinations of the AAF will be indicating the *actual capability* of the person to accomplish or obtain her/his VFTA. However, in reality, the output of a functioning does not necessarily correspond to an achievement. For instance, a person may be unsuccessful as well to obtain or accomplish what he or she values (i.e., non-achievement, or deficit). Besides, the person does not always choose the doings and beings that he or she values to obtain or accomplish. In other words, the person may simply find him or herself doing an action or an activity or being in a state that he or she does not value, due to random circumstances, obligations, social rules, traditions, adaptations to the conditions and requirements of society. But in any case, achievements should not violate her/his human rights, and should not put in danger the life of the person[69]. For instance, suppose that a person values and performs successfully an economic activity but in hazardous working conditions. In that case, the person will be considered achieved regarding his job. However, the fact that he or she is in dangerous working conditions is a harmful achievement for the person. In capabilities, it is important take in consideration all types of achievements including the harmful ones as well in order to understand the living conditions of the person, but the harmful achievements won't increase the real liberty of the person instead it will put her/his real liberty in danger[70]. In the literature, Wolff and Shalit (2007) has proposed "genuine opportunities for (secure) functionings" saying that the realization of a functioning should not put in danger the life of the person. In that sense,

[67] When we say, 'for the benefit of other(s)', it does not mean that the action or the activity is not harmful for other(s).

[68] In practice, many capabilities studies such as those based on the Alkire and Foster methodology (see the Multidimensional Poverty Index) consider that a larger set of achievements corresponds to a higher level of real liberty for the person.

[69] Of course, it should also respect the life of others and the environment.

[70] For understanding the extent of the harmful MIF in society, researcher can look at national/regional/community level MAF statistics.

we follow Wolff and Shalit's idea in expressing the non-harmful achievements in our framework.

Based on our explanations above we can say that there are four types of functionings reflecting personal achievements and two types of functionings reflecting personal non-achievements[71]: (i) the personally valued and achieved non harmful functionings; (ii) the personally valued and achieved harmful[72] functionings; (iii) the personally not valued but achieved non-harmful functionings; (iv) the personally not valued but achieved harmful functionings; (v) the personally valued but not achieved functionings; (vi) the personally not valued and not achieved functionings[73]. But there are also innate functionings (**IF** onwards) that the person got from her/his birth (e.g., being born with a high level of intelligence quotient) which cannot be considered as a personal achievement/not achievement, it is rather her/his chance. Moreover, there are also environmental functionings (**EF** onwards) that determine the natural (ecological) conditions of societies. The IF, MAF, and EF are neither a personal achievement nor a personal non-achievement. We propose to employ the terms '*achievement vs. not-achievement (deficit)'* as the outputs of the MIF whilst the terms '*advantage, disadvantage or neutral factor'* as the outputs of the IF, MAF, and EF[74]. The role of these latter in realizing the personal goals, or the MIF may be positive, negative, neutral, or both positive and negative. In other words, the person may see the IF, MAF, EF as an advantage or a disadvantage or a neutral factor[75] for effectuating her/his personal goals and the MIF. We note that of course a MIF can also be evaluated as an advantage, a disadvantage, or a neutral factor by the person him or herself (e.g., a university diploma can be seen as an advantage for obtaining a better paid job); however, we prefer to apply a direct approach regarding the personal goals and the MIF by focusing directly on the results of their realization (i.e., achieved or not achieved harmful or not harmful) proceeded by the person him or herself alone. On the other hand, as the IF, MAF, EF are not proceeded by the person him or herself alone, we then take a proxy approach (i.e., advantage vs. disadvantage) for evaluating their role in the realization of the personal goals, of the MIF, and in the life style of the person.

[71] The term *personal non achievement* is not used for blaming the person for not achieving a functioning, but it is used in a broad sense for mentioning a non-realization/a non-accomplishment.

[72] If the realization of a functioning puts in danger the life of the person and/or violates the person's human rights, then this functioning is considered harmful.

[73] For simplification purpose, we don't distinguish the non-achievements as a function of harmful vs. not harmful.

[74] We assume that an IF cannot put in danger the life of the person or violate her/his human rights, but an EF may put in danger the life of the person (e.g., natural disasters) and a MAF may both put in danger the life of the person and/or violate human rights (e.g., a social rule saying that girls cannot go to school). We keep apart harmful/not harmful with advantage/disadvantage/neutral factor because a harmful functioning may also bring an advantage to a person. For instance, a society allowing children working in hazardous conditions (as a MAF) is in fact a harmful functioning for the child but at the same time it can bring an advantage to the child for earning the money to buy food, clothes, school equipment. However, this type of advantage does not increase the real liberty of the person.

[75] As in principle, the person evaluates herself if the functioning brings an advantage/disadvantage/neutral effect, so we don't consider in additional another personal valuation on the functioning. We can say that if the person evaluates a functioning as an advantage, then it corresponds approximately to a valued functioning, or if the person evaluates a functioning as a disadvantage then it corresponds approximately to a non-valued functioning.

We also note that regarding the *not achieved MIF*, in order to understand whether such situation concerns also many other people at community, regional, or country level, the MAF statistics (e.g., national unemployment rate, community safety ratio, national drop out of school rate) can be used as complementary information to the MIF. For instance, the person can be interviewed about her/his educational attainment and her daily routine of eating. Suppose that the person did not achieve her/his secondary school degree (he or she dropped out of school) and he or she is not able to have a healthy meal on daily basis that he or she values. These two outputs correspond to the two non-achieved micro level functionings (MIF). However, for understanding the magnitude of the two phenomena in society, the interviewer/researcher can look at national/regional/community level statistics of school dropout and of access to subsistence.

Figure 3 below illustrates the outputs of the four types of functionings (MIF, IF, MAF, EF) that we described above in terms of macro vs. micro level, achievement vs. non-achievement, harmful vs. not harmful, advantage/disadvantage/neutral effect[76].

As we can see, in our framework, we propose to not only take in consideration achievements, but we extend the capabilities framework by categorizing the outputs of the functionings in different terms instead (see above). In this line, we define the capabilities as the *combination* of the outputs of the five distinct types of functionings (MIF, IF, MAF, EF, MACMICF). The possible outcomes of the MIF, IF, MAF, and EF are as follows[77]: (1) personally valued and achieved non-harmful micro level functionings; (2) personally valued and achieved harmful micro level functionings; (3) personally not valued but achieved non-harmful micro level functionings; (4) personally not valued but achieved harmful functionings; (5) personally valued but not achieved micro level functionings; (6) personally not valued and not achieved micro level functionings[78]; (7) advantages of innate functionings; (8) advantages of harmful macro level functionings; (9) advantages of non-harmful macro level functionings; (10) advantages of harmful environmental functionings; (11) advantages of non-harmful environmental functionings; (12) disadvantages of innate functionings; (13) disadvantages of harmful macro level functionings; (14) disadvantages of non-harmful macro level functionings; (15) disadvantages of harmful environmental functionings; (16) disadvantages of non-harmful environmental functionings; (17) neutral effect of innate functionings; (18) neutral effect of harmful macro level functionings; (19) neutral effect of non-harmful macro level functionings; (20) neutral effect of harmful environmental functionings; (21) neutral effect of non-harmful environmental functionings[79]. The terms advantages/disadvantages/neutral factor (effect) are normally obtained based on personal evaluation (i.e., subjective opinion of the person on the MAF, EF, IF); however, subjective evaluation should be also cross checked by objective criteria for enabling non-violation of human rights and for guaranteeing non-harmful living conditions. As we have already mentioned, this cross-check by employing objective evaluation criteria will be important to uncovering adaptation, contentment to dangerous living conditions and violation of human rights.

[76] Here, neutral factor and neutral effect are used interchangeably.

[77] We suppose that a functioning valued by the person does not violate human rights of other(s) and respect environment. However, in case of violation, it should be indicated in addition, for instance by the interviewer or researcher.

[78] For simplification purpose, we don't use the term *harmful* for non-achievements.

[79] The MAF, IF, EF may involve at the same time several advantages, disadvantages and neutral effect.

Besides, concerning the types of outputs of the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 we will treat them as in the case of the MIF because P is involved in the realization of the action, activity. On the other hand, the types of outputs of the MACMICF6 will be as in the case of the MAF because the MACMICF6 focus on the actions, activities performed by other persons in P's social environment (i.e., the person (P) does not effectuate any action or activity by him or herself). *But, if the person (P) has settled a personal goal (i.e., an action) to be performed by him or herself or by him or herself in collaboration with other(s) as a solution to one of her/his disadvantages emerged because of a MAF, IF, EF,* t*he result of her/his action will be treated as in the MIF*. For instance, suppose that Mary lives in a not safe neighborhood, and she would like to move out. If she achieves to find a new home in a safer place, then we can say that she achieved a valued non-harmful functioning. Another example suppose that Alice is a long-term unemployed living in a city with a few opportunities of job, so she wants to move in a bigger city with more job opportunities. She achieves her goal and she finds a job, but it puts her life in danger because of her new working conditions in the city. Then we can say that she achieved a valued but harmful functioning. So that, we employ a proxy approach for determining the possible outputs of the mixed macro-micro level functionings. However, we will later mention the nature of these functionings in practice.

A higher number of valued (non-harmful) achievements and advantages will increase the possibility to lead the life that the person values. A person may use simultaneously all of or some of her/his valued (non-harmful) achievements and advantages to achieve one or more functionings and/or to lead a lifestyle that he or she values[80]. Using the right (non-harmful) achievements and advantages, overcoming and being resilient against disadvantages and deficits is a something personal due to personal characteristics, social, economic and ecological conditions. But theoretically we can say that greater valued (non-harmful) achievements and advantages, and less deficits, harmful achievements and disadvantages will increase the capabilities of the person.

In a dynamic aspect, we consider the two periods: *t* for the present and *t+1* for the future. Here, periods are defined in large sense. In real life, the present (*t*) may be today and the period t+1 may be tomorrow, the period *t+2* may be the day after tomorrow, and so on. At the period *t*, the life situation of the person can be reflected by (i) her/his actual capability (AC onwards) set and (ii) the functionings to be realized (FTR onwards) by the person by *t+1*, namely specific personal goals[81] settled by the person (denoted $Ends_{t+1}$), other micro level and mixed macro-micro level functionings which will be occurred between *t* and *t+1* (denoted $MIF_{t+1,}$ $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$). The latter group includes: (i) the MIF and the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 that the person does not value him or herself, but he or she has to effectuate them; and (ii) the MIF, the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 that the person values him or herself but at *t* he or she has not yet listed them as personal goals. The outcomes of the FTR (i.e., achievement or deficit) will be emerged at *t+1* as a part of her/his AC_{t+1}. In order to realize the FTR, the person will use basically her/his actual capability set. Therefore, the AC_t becomes the means for achieving the personal ends ($Ends_{t+1}$), the MIF_{t+1}, and the $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$. The

[80] The personally not valued but achieved non-harmful functionings may increase the real liberty of the person depending on the conditions.

[81] Personal goals are considered as valued functionings to be realized.

$Ends_{t+1}$ will become 'personally valued and achieved harmful' or 'personally valued and achieved not harmful' or 'personally valued but not achieved' micro level functionings or the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 at *t+1*.

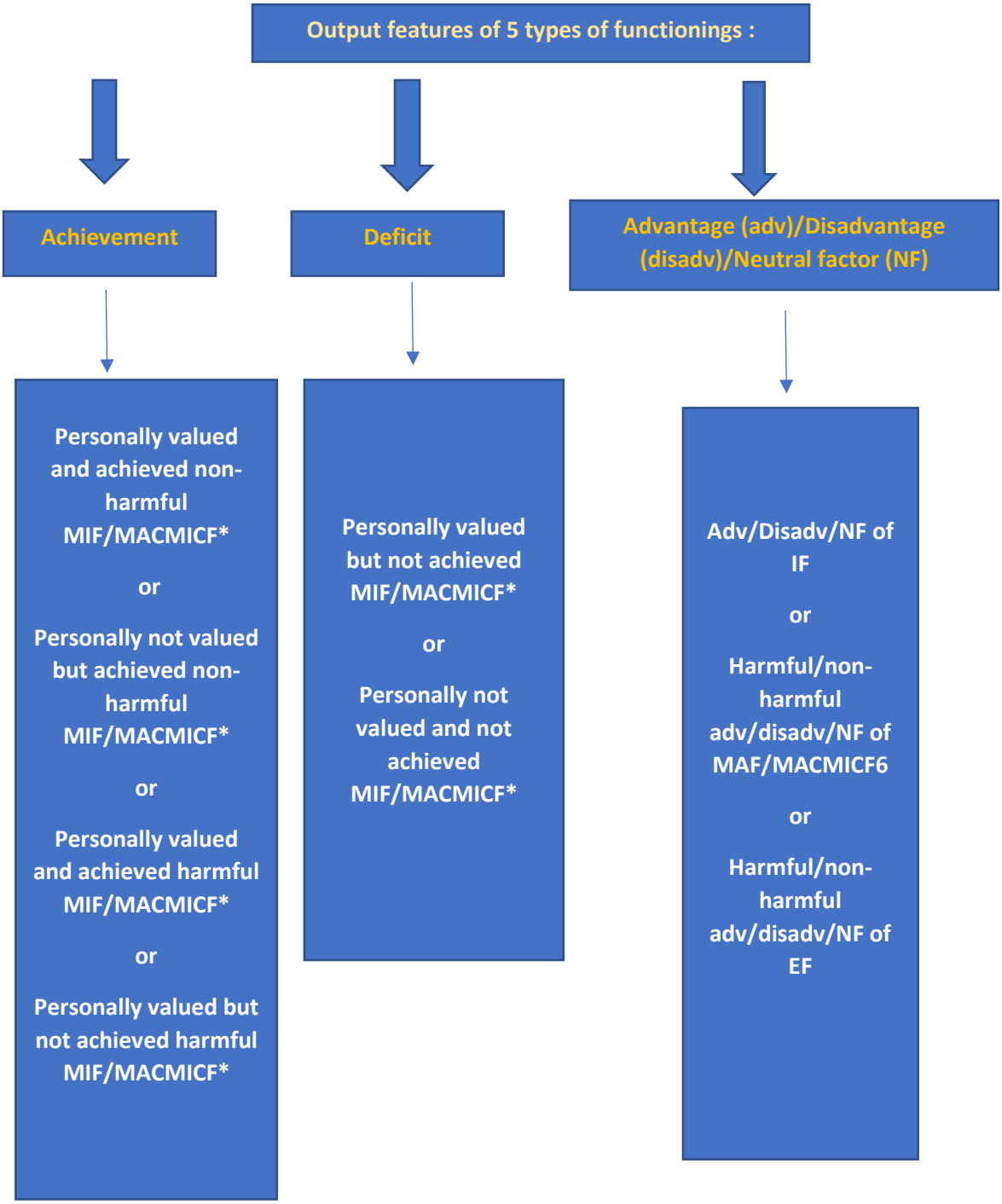


*Figure 3. The outputs of four types of functionings (MIF, IF, MAF, EF) in terms of macro vs. micro level, achievement vs. non-achievement, advantage vs. disadvantage factor. Note: * MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5. Source: Author.*

On the other hand, the results of the personal actions, activities realized between *t* and *t+1* (MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$) will become 'personally valued and achieved harmful' or 'personally valued and achieved not

harmful' or 'personally valued but not achieved' or 'personally not valued but achieved harmful' or 'personally not valued but achieved harmful' or 'personally not valued and not achieved' micro level or mixed macro-micro level[82] functionings. It is also important to note the person may also benefit of the micro level functionings, and the mixed macro-micro level functionings to be realized (MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$) between *t* and *t+1* for realizing her/his personal goals ($Ends_{t+1}$). The difference between the AC and the FTR is the former includes the actions, activities, states already happened at the present whilst the latter covers the personal actions, activities, states to be realized by the person between the present and the future. In other words, the AC of the person is composed of the results of the personal goals (Ends) and the outputs of the five types of functionings described in previous paragraph (see Figure 3), whilst the FTR do not include nor results of the ends neither outputs of the functionings.

Figure 4 below displays the actual capability sets (AC) and the personal functionings to be realized (FTR) at the three periods *t-1* (past), *t* (present), *t+1* (future). For instance, at the period *t*, the actual capability set of the person is composed of the five categories (denoted from a to e): (a) the results (achieved harmful/not harmful vs. not achieved) of the personal goals determined at *t-1* (denoted as $Ends_t$), (b) the valued/not valued achieved/not achieved harmful/not harmful micro level functionings (MIF_t), mixed macro-micro level functionings ($MACMICF1_t$, $MACMICF2_t$, $MACMICF3_t$, $MACMICF4_t$, $MACMICF5_t$), (c) the advantages/disadvantages/neutral effect of innate functionings (IF_t), (d) the advantages/disadvantages/neutral effect of harmful/non-harmful macro level functionings (MAF_t), mixed macro-micro level functioning ($MACMICF6_t$), (e) the advantages/disadvantages/neutral effect of harmful/non-harmful environmental functionings (EF_t). Between the two periods *t* and *t+1*, the capability set may have encountered modifications such as some functionings may have had new achievements, or new deficits, or may become more or less advantaged or disadvantaged, or the capability set may remain the same (status quo). For instance, the set of MAF of the person at *t* may have had modifications between *t* and *t+1* which in turn may have changed advantages and disadvantages of the person regarding her/his MAF at *t+1*. Concerning the FTR at the period *t*, it includes two categories: (f) listed specific personal goals to be realized by the person between *t* and *t+1* ($Ends_{t+1}$), and (g) other not yet listed micro, and mixed macro-micro levels functionings (MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$) to be realized by the person between *t* and *t+1*.

Regarding the arrows illustrated in Figure 4 below, the *orange* ones are used to indicate that the personal goals to be realized in the future will become the valued achieved harmful/not harmful or the valued but not achieved micro level, or mixed macro-micro level functionings in the future. For instance, the personal goals determined at the period *t* ($Ends_t$) will become the valued achieved harmful/not harmful or the valued but not achieved micro level, or mixed macro-micro level functionings at the period *t+1* (MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$).These are indicated under the name of 'Results of $Ends_{t+1}$' which is part of the actual capability set of the persion at *t+1*. The person can use the results of her/his $Ends_{t+1}$ in realizing her/his personal goals, micro level and mixed macro-micro level functionings for the next period (*t+2*). Concerning the *yellow* arrows, they indicate the micro level, mixed macro-micro level functionings the person will effectuate. For instance, at the

[82] Namely, MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5.

period *t*, the person has already obtained or accomplished micro level functionings; however, he or she will also proceed other micro level, mixed macro-micro level functionings between *t* and *t+1* in order to achieve them at *t+1*. The not yet achieved micro level, mixed macro-micro level functionings (FTR_t) will be denoted by MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$ which have six possible outputs at *t+1*: (i) valued and achieved harmful level functionings; (ii) valued and achieved not harmful functionings; (iii) valued but not achieved functionings; (iv) not valued but achieved harmful functionings; (v) not valued but achieved not harmful functionings; and (vi) not valued and not achieved functionings. The outputs of the MIF_{t+1}, $MACMICF1_{t+1}$, $MACMICF2_{t+1}$, $MACMICF3_{t+1}$, $MACMICF4_{t+1}$, $MACMICF5_{t+1}$ will become therefore part of the capability set of the person at *t+1*. It is also important to note that some achievements/non-achievements happened in the past may still exist in the life of the person. So that, the micro level, mixed macro-micro level functionings in the actual capability set of the person include both the results of the FTR and the existing results of other micro level, mixed macro-micro level functionings occurred in the past.

We note that in Figure 4, due to space limitations, we did not show the case of the MACMICF. But, as we have mentioned previously: the possible outputs of the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 will be the same as the MIF, and the possible outcomes of the MACMICF6 will be the same as the MAF in a dynamic perspective (between the two periods *t* and *t+1*).

Now, we will provide two small examples. We suppose that at the moment *t* a young woman called Jennifer has aimed at losing some kilos in one year (by *t+1*). To achieve her goal, she thinks that she needs to run every day 30 min. and she needs to go to gym saloon twice a week. And for that, she realizes that she should be able to have sufficient (individual) income level for paying the gym saloon fees and be able physically to run (have a good health conditions). Her actual capability set allows her to realize to go to gym and to run between *t* and *t+1*. Later, she informs that at *t+1* she achieved her personal goal of losing some kilos. In that case, given the terms defined in Figure 4 below, we can say that: losing some kilos is a personal goal to be realized at *t+1*, denoted as $Ends_{t+1}$; running every day 30 min. and going to gym saloon twice a week between *t* and *t+1* are the valued micro level functionings to be realized between *t* and *t+1*, denoted as MIF_{t+1} (they are not goals because they are counted as the intermediate activities for achieving her main goal). Moreover, having sufficient individual income level for paying the gym fees and having good health status for running are part of her actual capability set at *t* corresponding to the results of her previous ends or the outputs of her previous micro level functionings[83]. We can result that Jennifer has used her actual capability set at *t* for effectuating and achieving her MIF_{t+1} at *t*, but also, she has used some of her MIF_{t+1} to achieve her personal goal ($Ends_{t+1}$) at *t*. Referring to Figure 4 below, for instance at the period *t*, the categories a, b, c, d, e, and g are used for achieving the category f; on the other hand, the categories a, b, c, d, and e are used for achieving the category g.

Another example is we suppose that an adolescent called Max would like to go to public university for studying philosophy next year (at *t+1*). For that, he knows that he needs to pass

[83] In this example, we don't have comprehensive information on the nature of individual income level and having good health status functionings. In other words, we don't know on which activities and actions that these two functionings are based. So that, we consider them directly as the MIF supposing that they are got by the person for herself.

his final exams at high school at the end of *t* (*t* is present year). He looks at his capability set at *t* for being prepared to achieve his goal. He describes the main features of his actual capability set as follows : (1) he is working two hours per day in week for himself (light work allowed by the constitution of his country, for savings for his future); (2) he is doing some household tasks such as cleaning the house two hours per day (MACMICF5); (3) his parents are able to pay his university's first year tuition fee (MACMICF6); (4) he has a high level of intellectual ability for studying literature and philosophy; (5) he can regularly attend the school; (6) his parents are able to afford his high school books (MACMICF6; (7) he has to spend two hours in commuting between school and home; (8) and his parents are able to afford daily nutritional requirement for everyone in the household (MAF). In that case, we can say that: going to university is Max's personal goal to be realized at *t+1* ($Ends_{t+1}$); passing his exams is the micro level functioning to be realized between *t* and *t+1*. In addition, Max qualifies his intellectual ability, his parent's ability to pay his university fee and his high school books, and to provide sufficient financial situation for being daily well-nourished, as advantages helping him to increase his real liberty to lead the life he values. Concerning his MIF and MACMICF, he reports that he values working, doing some household tasks, attending the school, and commuting time between home and school. For realizing his $Ends_{t+1}$ and MIF_{t+1}, Max will use his actual capability set at *t* which is composed of the following items:

- (1) his employment status as a valued $MIF_{t.}$ But this performance should be checked regarding: (i) whether it is harmful or not harmful for him[84]; (ii) whether it is performed only for him or only for other(s) or for both him and other(s). For instance, Max may work in hazardous working conditions for his parents without benefiting of his own earnings (e.g., MACMICF4).
- (2) his role in effectuating some household tasks at home as a valued $MACMICF5_{t.}$ But this performance should be checked regarding: (i) whether it is harmful or not harmful for him[85].
- (3) his parents are able to pay his first-year university fee as an advantaged $MACMICF6_t$.
- (4) his intellectual ability for literature and philosophy as an advantaged IF_t.
- (5) he attends regularly the school as a valued MIF_t[86].
- (6) his parents' achievement for paying his high school books as an advantaged $MACMICF6_t$.
- (7) spending two hours per day in week between home and school for reaching his high school as a valued MIF_t. But this performance should be checked whether it is harmful or not harmful for him[87].

[84]If it is a harmful achievement, then it can be cross checked by macro level statistics whether it is a community or regional or country level practice.

[85]If it is a harmful achievement, then it can be cross checked by macro level statistics whether it is a community or regional or country level practice.

[86] Additional information (MAF) on whether society where he or she lives provides community the opportunity to go to school such as available school, available teacher, available and secure roads to school, available transports (i.e., societal achievements) can be checked.

[87] In case of available information on commuting type and time, some other MAF can be checked such as available roads, transports, school in the community (i.e., societal achievements).

- (8) His parents' financial situation allows everyone living in the household to have an adequate diet on daily basis as an advantaged MAF_t[88].

From Max's example, it is important to point out the two following situations: First, Max will be working, be performing household tasks, be spending two hours on road between school and home and be attending regularly his school between *t* and *t+1*. In other words, he already achieved these four functionings at the period *t* and he expects to perform them between *t* and *t+1* as well (like a continuum of the actual occurred functionings). Thus, some micro level, and mixed macro-micro level functionings to be realized can be the continuum of the actual occurred functionings. On the other hand, the fact that his parents are already paid his high school books is a finished occurred functioning at the period *t* (no continuum on this)[89]. Second, at the period *t*, his parents are actually able to pay his university tuition fee if he is accepted later at the university, and they are actually able also to provide an adequate nutrition to all members of the household. But as we have previously mentioned the situation of the functionings may meet modifications over time. For instance, at the end of *t*, Max's parents may also find themselves for not being able to pays his tuition fee due to financial crisis.

Thus, we can do three types of analysis based on the actual capability set of the person (P):

- analyzing directly the actual capability situation of the person based on only occurred functionings (outputs) without having information on her/his FTR. This type of examination puts in evidence the *actual living conditions* (or the *actual life situation*) of the person.
- predicting the outcomes of micro level, mixed macro-micro level functionings to be realized based on the actual capability set.
- predicting the future living conditions (or the future life situation) and/or the likelihood of the personal goals based on the actual information on the occurred functionings, on the predicted information on the micro level, mixed macro-micro level functionings to be realized, the MAF, IF, and EF.

As a summary, we mainly topologized functionings as a function of *nature* (micro/macro/macro-micro/ environmental/innate), *time* (already obtained or accomplished vs. to be obtained or accomplished) aspects. Based on that, different types of outputs are derived for each type of functionings: achieved/not achieved micro level functionings, achieved/not achieved personal goals, achieved/not achieved mixed macro-micro level functionings[90], and advantages/disadvantages/neutral effect of macro level, of mixed macro-micro level[91], of innate, and of environmental functionings. In addition, we focused on (i) if the person values or not the micro level, mixed macro-micro level functioning; (ii) whether the performance of the functioning is harmful; (iii) whether the result of an activity or an action that the person (P) performs (i.e., doings) is only for him or herself (P) or only for others or for the shared/societal

[88] We can also say that he is in a situation where his household is able to afford every member of the household an adequate daily nutrition. This achievement is his household's shared achievement (MAF). However, for his parents, this functioning will be MACMICF2.

[89] We assume that the IF do not change over time as they are the functionings that the person got from birth. But if a new situation emerges related to his or her IF, the other types of functionings (e.g., MAF, MIF, EF, MACMICF) behind this modification should be integrated in his or her capability set.

[90] The outputs of the MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5 are considered as in the case of the MIF.

[91] The outputs of the MACMICF6 are considered as in the case of the MAF.

purpose; (iv) whether the person (P) effectuates the action or the activity (i.e., doings) by him or herself alone (P) or he or she participates with other(s) in a joint action or activity; (v) whether another person(s) achieved an activity or an action for the person (P); (vi) whether the person (P) is in a situation (i.e., beings) as a result of a societal functioning; and (vii) whether the situation in which the person (P) faces (i.e., beings) is shared by more than one person at household, community, regional, country level.

We note that in Figure 4, for simplifying (due to space limitation) the illustration of the mechanism between functionings and capabilities over time, we did not integrate nor 'the not harmful vs. not harmful' neither 'the mixed macro-micro level functionings' (MACMICF). But the necessary information on them is already mentioned in the text above.

To put in context the functionings, we will see a fictional example. We suppose that a young very intelligent woman called Annie lives in a very clean and naturally rich island and she works in fishery sector (performing an economic activity she values on her daily basis in week). At the time *t* (e.g., see Figures 3,4) Annie has been interviewed by a scholar for a research project and she has reported information about her present/actual functionings (i.e., actual capability set_t and functionings to be realized at *t+1*[92]). At *t*, she has set two goals to be realized at *t+1*: participating in an international seminar on fishery (for her own interest) in another continent and finishing her study in anthropology that she has started two years ago. These two activities are valued by Annie and she would like to achieve them at *t+1*. At the same time, even if she does not particularly value, at *t+1* she has to get her driver license required by her workplace (for only the company benefit/for work purpose) and also, she has to participate regularly between *t* and *t+1* in pilates class with her mom (Annie will do it for her mom). In addition concerning her actual capability set at *t*: (i) she already achieved to keep good health; (ii) even if she does not particularly value, she achieved to become able to run every day for 30 min.; (iii) she bought the apartment she wished at *t-1*[93] (i.e., personally valued goal); (iv) she is very satisfied with the natural conditions of where she lives; (v) she thinks that her high level of intelligence quotient brings her self-confidence to accomplish many tasks; and (vi) the company where she works is very successful in fishery sector which brings her a high level of salary and many other travelling and seminar possibilities across the world. One year later, at *t+1*, Annie has reported about the results of her 'functionings to be realized at *t+1*': (i) she achieved to finish her study in anthropology; (ii) she obtained her driver license; (iii) she did not achieve to participate in the seminar because of her work intensity; (iv) she did not achieve to attend regularly her pilates class. Moreover, between *t* and *t*+1, no modifications had happened concerning her innate functionings (IF) and her social, ecological environment (MAF, EF), and we don't have any other additional information on her MIF or MACMICF. Given Annie's profile, we can qualify her functionings as follows, in Table 1 below.

[92] *t+1* is the period after *t*.
[93] *t-1* is the period before *t*.

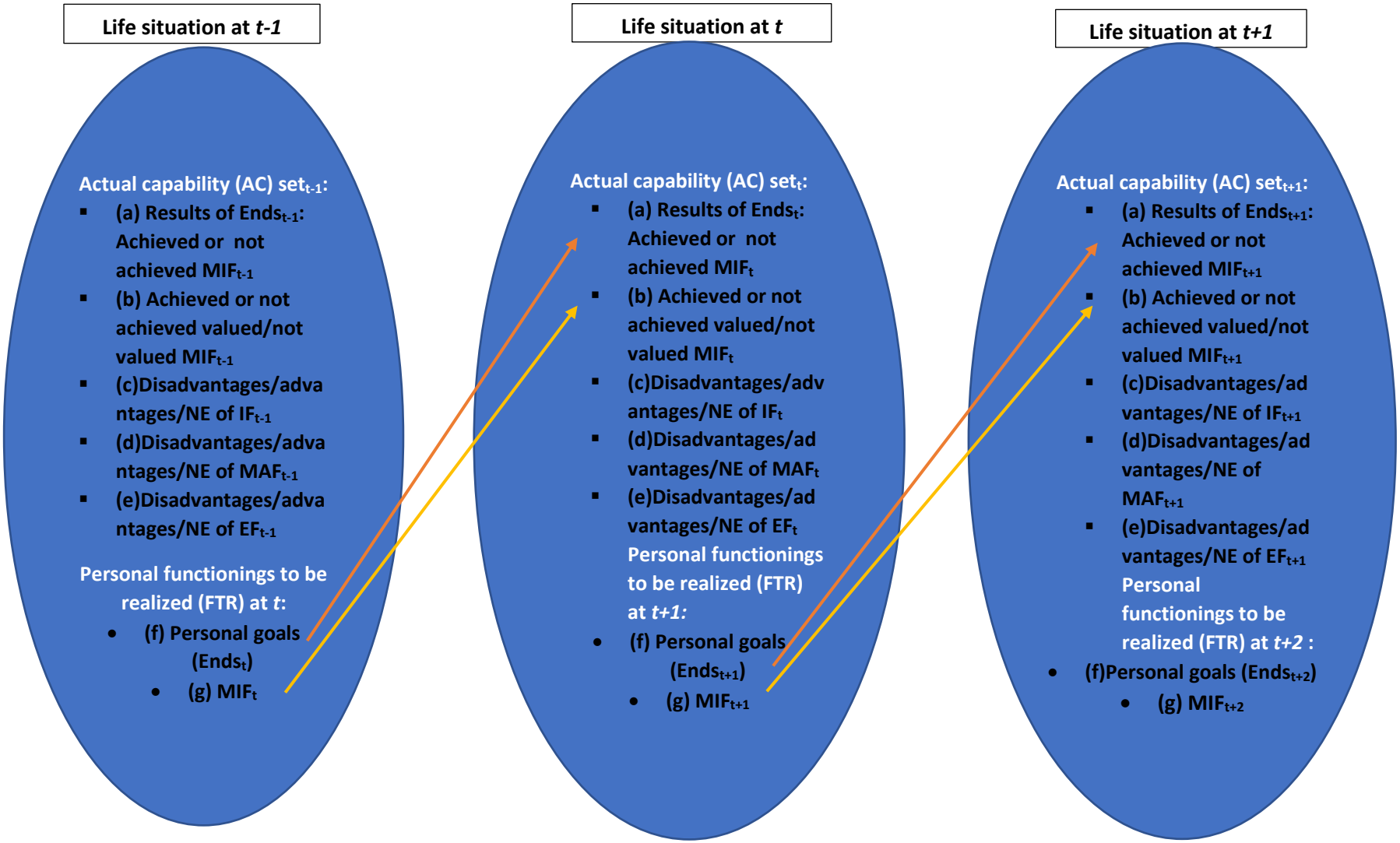


Figure 4. Functionings and capabilities in a dynamic aspect.
Note: (i) NE means 'neutral effect'; (ii) for space limitation in the figure, we did not add the MICMACF, valued vs. not valued, harmful vs. non-harmful aspects. Source: Author.

Table 1. Annie's functionings, an example for putting in context our typology of functionings.

Functioning name	**Nature/feature of functioning**	**Category (AC or FTR at t+1)[94]**	**Description of functioning**	**Valuation by Annie at *t***	**State of functioning at *t+1***
Living in a clean and naturally rich island[95]	Mix of non-harmful environmental and macro level functionings at *t*: EF_t and MAF_t	Category d and e (AC_t)	At *t:* the environmental (naturally rich) and the socio-environmental (kept clean by habitants in the island) conditions of the island where Annie lives.	Annie is satisfied with the environmental conditions of the island, so that she finds it as an advantage increasing her quality of life.	No modification between *t* and *t+1*.
Having a high level of intelligence quotient	Innate non-harmful functioning at *t*: IF_t	Category c (AC_t)	Annie has a high level of intelligence quotient from her birth.	Annie is satisfied with her intelligence and she finds it as an advantage for being a self-confident person.	No modification between *t* and *t+1*.
Having bought an apartment	Result of End_t (we assume that it is not a harmful action)	Category a (AC_t)	At *t* Annie achieved to buy the apartment she wished (as a goal) at *t-1*.	Annie values it personally.	No modification between *t* and *t+1*.
Having a good health[96]	Micro level functioning at *t*: MIF_t (we assume that it is not a harmful status)	Category b (AC_t)	Annie achieved to keep her good health status at *t*.	Annie values it personally.	No modification between *t* and *t+1*.

[94] See Figure 4.

[95] This is a specific functioning. The fact that Annie lives does correspond to her personal action for herself on her daily basis, so in this sense it is a MIF. However, we have additional information on her living conditions: island is naturally rich and habitants living in the island have achieved to keep clean the island. From these two latter points, it is a combination of EF and MAF respectively. If we don't have information on details, then we can keep it as the MIF. But, if we have additional information describing the natural and socio-environmental conditions where the person actually lives, then it is important to keep these two aspects in actual capability set of the person.

[96] We don't have additional information on the actions, activities, states allowed her to achieve a good health status.

Running every day for 30 min.[97]	Micro level functioning at *t*: MIF_t (we assume that it is not a harmful activity)	Category b (AC_t)	Annie achieved to run every day for 30 min. during the period between *t-1* and *t.*	Annie does not value it personally.	No modification between *t* and *t+1*.
Performing regularly an economic activity in week company	Micro level functioning at *t*: MIF_t (we assume that it is not a harmful activity)	Category b (AC_t)	Annie is working regularly in week in a fishery company.	Annie values this activity.	No modification between *t* and *t+1*.
The company where Annie works is very successful	Macro level functioning at *t*: MAF_t (we assume that it is not a harmful state)[98]	Category d (AC_t)	Working environment of Annie is quite successful in fishery sector. This is part of her social environment, reflecting not Annie's personal success but the company's. In other words, the company's success is the result of company's workers' activities for the company.	Annie is satisfied with her working environment. She finds it as an advantage for travelling and having a high-income level which increase her quality of life.	No modification between *t* and *t+1*.
Willing of participating in an international seminar on fishery at *t+1*	Personal goal to be realized at *t+1*: End_{t+1}	Category f (FTR at *t+1*)	Annie wants to attend an international seminar at *t+1*.	Annie values it personally.	Annie did not achieve.

[97] We don't have additional information on the actions, activities, states allowed her to achieve running 30 min. every day.

[98] We don't have detailed information whether Annie or her colleagues at the company work successful or not for the success of the company. In that case, we cannot consider MACMICF2. Only we know is the fact that the company is successful on the market, it is a result of complex non-directly observable interactions among the behaviors of all workers in the company, so we consider the company's success as the MAF. If the success of the company was not evoked and only we knew that Annie and her colleagues work in collaboration for a specific joint project for the benefit of the company, then in such case we can consider this activity as the MACMICF2 (similar to performing household tasks with all members of the household for the benefit of the entire household).

Willing of finishing her study in anthropology at *t+1*	Personal goal to be realized at *t+1*: End_{t+1}	Category f (FTR at *t+1*)	Annie wants to get her university diploma in anthropology at *t+1*.	Annie values it personally.	Annie achieved.
Getting her driving license at *t+1*	MACMIC4 to be realized at *t+1:* $MACMICF4_{t+1}$	Category g (FTR at *t+1*)	Annie has to obtain her driving license at *t+1*.	Annie does not value it personally, but it is required by her company.	Annie achieved (we assume that it is not harmful)
Participating regularly in pilates class between *t* and *t+1*	MACMICF4 to be realized at *t+1*: $MACMICF4_{t+1}$	Category g (FTR at *t+1*)	Annie has to attend regularly pilates class between *t* and *t+1*.	Annie does not value it personally, but it is required by her mom (so she does it for her mom).	Annie did not achieve.

Now, we will provide another example for showing both the functionings and the capability sets in a dynamic aspect. We suppose that a woman named Kate has been interviewed on June 2018 by a national statistical officer in the framework of measuring quality of life in the country where Kate lives actually. The interviewer has focused first of all on the *actual capability* set of Kate by asking her various questions on socio-economic aspect of life. Some of the questions are as follows: (i) what is her highest educational attainment; (ii) did she perform an economic activity over last week; (iii) did she go to vacation at least once in 2017; (iv) could she go to medical center if she needed it over last year; (v) does she have the electricity at home; (vi) did she afford without problem a daily adequate nutrition for the entire household over last week; (vii) is there any goal(s) she wished in 2017 and achieved by June 2018; (viii) is there any goal(s) she wished in 2017 but not achieved by June 2018; (ix) Does she think that she has some specific innate abilities, if yes which one(s)? Kate gave the following answers to these nine questions: (i) high school diploma; (ii) yes, as a sales representative at a clothes shop; (iii) no; (iv) yes; (v) yes; (vi) yes; (vii) yes, she got a certificate as a yoga instructor; (viii) no; (ix) no, not specifically. The interviewer has also collected some additional external information concerning the situation of her community in terms of social and ecological conditions, her household information, and the socio-economic context of her country. As a result, the interviewer has noted that there is a safety issue in Kate's community, notably after 8 pm with a quite high crime rate, and there are no sufficient green areas nearby Kate's home, and the country recent statistics have shown that due to economic crisis, the dismissal in clothing sector is predicted to increase next year. Moreover, in the city where Kate lives the snowing rate is quite high and often this causes transport concerns. Kate has also reported that she is living with her 5-year-old daughter. In order to understand better Kate's life situation, the interviewer decided to ask additional following questions: (i) does she value her educational attainment; (ii) does she value her job; (iii) does she value vacation at least once a year; (iv) does she value the quality of medical service; (v) does she have transportation problem when it snows to get her workplace? Kate answers yes to all of them and she adds that transportation problem in the city is an important issue. The interview finishes by a last question: does she has some specific goals

(ends) planned for the end of year (December 2018)? Kate has answered that she wishes her daughter can be successfully pass the entrance exam to the German private nearby primary school and she plans to buy a car for facilitating her mobility.

As a function of the collected information, the interviewer lists Kate's *actual capability* (AC) set as follows (we assume that there is no harmful functioning):

Table 2. Kate's capability set and functionings, an example for putting in context our typology of functionings and capability set in a dynamic aspect.

Functionings-Description of the Capability set of June 2018	**Type of functionings**	**Category (see Figure 2)**
Highest educational attainment: Valued high school diploma	Valued and achieved micro level functioning (MIF_{June})	Category b
Economic activity: valued job (sales representative)	Valued and achieved micro level functioning (MIF_{June})	Category b
Vacation at least once a year[99]	Valued but not achieved micro level functioning (MIF_{June})	Category b
Visiting medical center if needed[100]: yes	Valued and achieved micro level functioning (MIF_{June})	Category b
Household electricity: yes[101]	Advantage, macro level functioning (MAF_{June})	Category d
Daily adequate nutrition: yes	Valued and achieved mixed macro-micro level functioning ($MACMICF5_{June}$)	Category b
Obtain the yoga instructor certificate	Valued and wished goal in 2017, and achieved by June 2018 (End_{June})	Category a
Safety problem in community	Disadvantage, macro level functioning (MAF_{June})	Category d
Lack of sufficient green areas in community	Disadvantage, macro level functioning (MAF_{June})	Category d
Economic crisis in the country: likely increasing dismissal in clothing sector where Kate works actually	Disadvantage, macro level functioning (MAF_{June})	Category d

[99] We don't have information if she went to holiday with her daughter; if yes, the functioning will be MACMICF2.

[100] It can be important to collect information on MAF statistics related to the availability of medical center and health facilities in the community where Kate lives, as additional information reflecting societal achievements in this sector.

[101] We don't have additional information on who finance the electricity in her household. As the electricity concerns the entire household and we don't know who takes the action for paying electricity bills, we consider this functioning as the MAF. Otherwise, MACMICF2 or MACMICF5 may be possible.

Lack of adequate public transportation when it snows	Disadvantage, macro level functioning (MAF_{June})	Category d

In addition, as a function of the collected information, the interviewer lists Kate's personally valued goals to be realized *for December 2018*, denoted as $Ends_{2018}$ (i.e., FTR) see the table below.

Table 3. Kate's capability set and personal goals, an example for putting in context our typology of functionings and capability set in a dynamic aspect.

Functionings-Description of the Ends for December 2018	**Type of functionings**	**Category (see Figure 2)**
Educational attainment of her daughter (she can go to German school)	Valued goal to be realized at December 2018 (we suppose that she takes legally the responsibility of her daughter)	Category f
Buying a car	Valued goal to be realized at December 2018	Category f

On January 2019, the interviewer visits again Kate for actualizing her capability set. The interviewer got the following information: compared to her situation in December 2018, she reported that her situation remained the same except she bought a car, and her daughter was accepted for German school. Concerning her MAF, EF, and IF, there is also no modification. So that, her actualized capability set in January 2019 includes (i) the results of her two ends for December 2018 (two new achievements); (ii) the same functionings of June 2018.

As we have seen, regarding Sen's approach, the person's achievements correspond to her/his capability of achieving various actions, activities, states (functionings) and of leading the life he or she values. Greater achievements increase the probability of greater opportunities for achieving manifold functionings. More the person achieves to obtain what he or she values without damaging the life of others, more the person has a higher real liberty in her/his life. So that, more achievements lead to higher real liberty which in turn means less poverty. Thus, the capability set of the person reflects her/his real liberty profile. In practice, Sen focuses on the number of non-achievements (deficits) for measuring poverty in the sense that greater number of deficits leads to greater poverty (i.e., lower real liberty)[102]. For instance, United Nations' Multidimensional Poverty Index[103] methodology counts the number of non-achievements regarding the ten distinct functionings (nutrition, child mortality, years of schooling, school attendance, cooking fuel, sanitation, drinking water, electricity, housing, assets)[104] for each person in the household. Then, the total number of non-achievements adjusted by weights is computed for each household called the household total deprivation score which is later compared with the poverty line settled by the commission of the United Nations (if the household's deprivation score is higher than 33 per cent than the household is considered poor

[102] Sen suggests particularly employing the term 'basic capabilities' for measuring poverty in developing countries (see Capabilities). And he calls poverty as "capability deficit".
[103] The MPI is conceptually based on Sen's capabilities approach and technically based on Alkire and Foster's counting methodology (see Multidimensional Poverty).
[104] For the full description of the functionings of the MPI, see the following link (accessed 12 March 2019): http://hdr.undp.org/en/2018-MPI

and therefore, all people living in poor households are considered poor). Regarding this methodology, greater non-achievements[105] indicate poorer situation with less real liberty and less capabilities for being able to achieve the life the person values. Thus, the focus is solely on the number of non-achievements and on people below the poverty line as a target population. However, we argue that taking into account the combinations of achievements, deficits, advantages, and disadvantages based on the typology of the functionings we propose above will provide a larger picture of the actual living conditions of the person. Therefore, the nature of the actual capability set of the person reflecting her/his actual living conditions (see Figure 4) will indicate the nature of actual poverty situation of the person. This process we propose does not only include those below the poverty line but focuses on the whole population accompanied with different sub-group analysis.

In a dynamic prospect, based on the four main outputs of functionings namely achievements, deficits, advantages, and disadvantages, we propose that multidimensional poverty can be measured in three ways as follows[106]: (i) actual living conditions analysis of the person (*actual capability set analysis* leading to *actual poverty analysis*) at a specific time *t*; (ii) ex-ante poverty analysis at *t*: by taking into account the actual capability set_t, forecasting the likelihood of achievement of functionings to be realized at *t+1* ($Ends_{t+1}$, MIF_{t+1}, and $MACMICF_{t+1}$), (iii) ex-post poverty analysis at *t+1* (after *t*): comparison of the capability set_t and the capability set_{t+1} to see the evolution of poverty situation of the person between the two periods (*t* and *t+1*). We can illustrate this proposition by Figure 5. We also note that in this thesis, we focus only on the actual poverty analysis.

We propose that the actual living conditions analysis at *t* should be based on the outputs of the MIF, MAF, MACMICF, IF, and EF of the actual capability set described above (see Figures 3, 4). In opposition to traditional multidimensional poverty studies (e.g., MPI), we assume that poverty is not something experienced by only a certain share of population staying under an arbitrarily defined poverty line. But rather in a society everyone faces poverty in a different manner with different capability sets. So that our framework concerns whole population and don't use a poverty line technique to dichotomize the society as the poor vs. the not poor. But how to process from the individual actual capability set to the societal level actual poverty status? Firstly, the actual capability set of each individual in terms of the combination of the outputs of her/his functionings will be detected. Secondly, each person regarding her/his actual capability set will be assessed in one of the following three situations (i.e., poverty profile)[107]: (i) *'zero poverty'*: being totally prosperous (combinations of achievements and advantages); or (ii) *'maximum poverty'*: being totally poor (combinations of deficits and disadvantages); or (iii) *'in-between poverty'*: being between totally prosperous and totally poor. The latter option is more realistic including combinations of achievements, deficits, advantages and disadvantages. Moreover, being totally prosperous means that the person has only meet (i) achievements regarding all her/his MIF, MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5; and (ii) advantages regarding all her/his MACMICF6, MAF, IF and EF. This situation corresponds to the highest real liberty for leading the lifestyle that the person values to have. Whilst being totally poor means that the person has only meet (i) deficits regarding all

[105] In technical paper, non-achievements are called as *deprivations*.
[106] See Figure 5.
[107] By accepting that a better capability set with greater number of achievements and advantages will bring more opportunities for being able to lead a valued life with less poverty.

her/his MAF, MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5; and (ii) disadvantages regarding all her/his MACMICF6, MAF, IF and EF. In latter case, the situation corresponds to the lowest real liberty for leading the lifestyle that the person values to have. At the end, within each of the three poverty profiles, we will look for the most influential actual poverty configurations (combinations of achievement, deficit, disadvantage, advantage including information on valued vs. not valued, and harmful vs. not harmful).

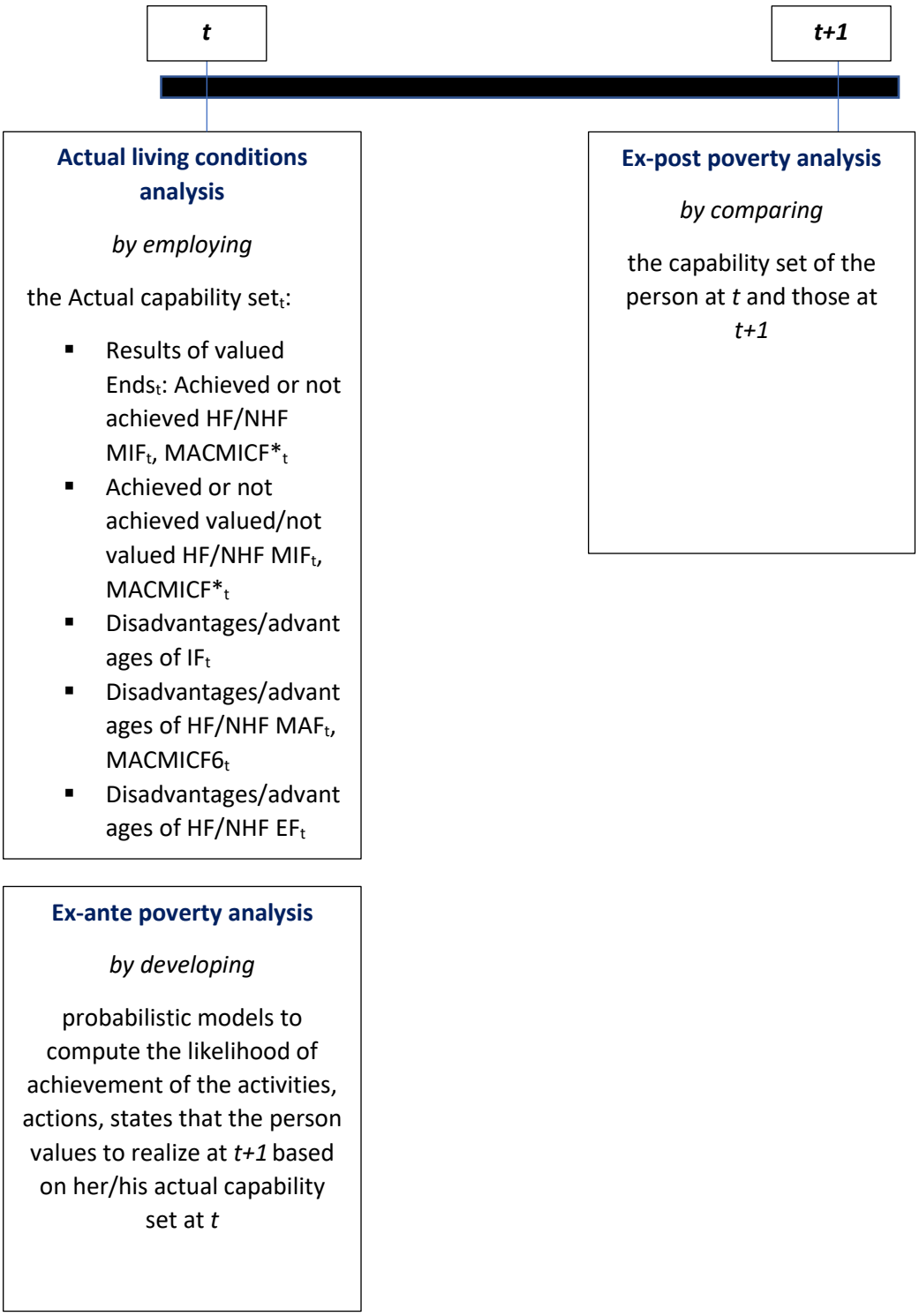


*Figure 5. Multidimensional Poverty analysis based on capabilities in a dynamic prospect, see Figure 3 for the terminology of functionings. Notes: (i) * MACMICF1, MACMICF2, MACMICF3, MACMICF4, MACMICF5; (ii) HF: harmful, NHF: not harmful. Source: Author.*

In practice, in addition to Sen, there are many other researchers who have worked on this topic, drawing directly on Sen's theoretical framework to develop their own theoretical frameworks or applying it for qualitative/quantitative analyzes on various quality of life topics such as poverty, well-being, inequality at global, national, regional, community level or even in more

specific sense by focusing on certain groups in society. In next paragraphs, we will explore an interesting qualitative based philosophical work on the question of being disadvantage in society, carried out by Wolff and de-Shalit (2007). In this work, authors investigate into different groups of disadvantages, and notably they point out the functonings and its measurement. They define disadvantage as "not being able to achieve the functioning" (2007, 38). Their analytical and empirical study is based on Nussbaum's (1988, 145-184; 2000; 2011; 2013) 10 "central capabilities" for being able to live a "life of fully human" (2007, 38)[108]. Additionally, authors added 4 other basic capabilities to Nussbaum's list (i.e., from 10th to 14th item in the list below). Their work has consisted of interviewing some very disadvantaged people in Britain and in Israel about their opinions on each of 14 functionings, and their perceptions about not being able to achieve them. Specifically, author's questionnaire to interviewees includes (2007, 190-191): "(1) *life*: being able to live to the end of a human life of normal length; (2) *bodily health*: being able to have good health, including reproductive health; to be adequately nourished, to have adequate shelter; (3) *bodily integrity*: being able to move freely from place to place; being able to be secure against assault, including sexual assault, child sexual abuse, and domestic violence; having opportunities for sexual satisfaction and for choice in matters of reproduction; (4) *sense, imagination, and thought*: being able to imagine, think, and reason-and to do these things in a way informed and cultivated by an adequate education. Freedom of expression, speech, and religion; (5) *emotions:* being able to have attachments to things and people outside ourselves; to love those who love and care for us; (6) *practical reason*: being able to engage in critical reflection about the planning of one's life; (7) *affiliation*: being able to live with and toward others, to recognize and show concern for other human beings, to

[108] The philosopher Martha Nussbaum used a teleological approach to link capabilities with human wellbeing. She proposed one particular list of valuable capabilities (composed of 10 items) that she calls as *central human capabilities* (see Nussbaum 1992, 1993, 2000, 2011, 2013) necessary to be fulfilled for being able to live well: : (1) life; (2) bodily health; (3) bodily integrity; (4) senses, imagination and thought; (5) emotions; (6) practical reason; (7) affiliation; (8) other species; (9) play; and (10) control over one's environment (see Nussbaum 2000, 2003). Regarding her, the set of these ten central capabilities may help to create a decent life for people in the sense that the fulfillment of this set will make enable various achievements that people value. This list is open-ended. Nussbaum remains open for any modifications in that list, and she has mentioned that these capabilities can be considered as general ones, whilst any local decision process may lead to create specific capabilities. For instance, Nussbaum has added the capability to hold property rights to the list, but also, she included the capability of living the life by taking care of animals, plant and the natural environment. Nussbaum proposed these ten capabilities as the must-have items in the constitutions provided by the governments to their all citizens. Due to this political aspect, Nussbaum aims at endorsing one particular list (see Robeyns 2005c, 103). However, Sen did not oppose to Nussbaum's list; he just argues that there should not be a single list of capabilities (see Sen 1993). In fact, the capabilities approach has known an on-going debate on whether a pre-determined list of relevant capabilities is needed or not. The lack of a rigorous framework to determine which capabilities should be retained, is often evoked in the literature by several academicians. For instance, Alkire (2007a, 6) mentioned that "Sen's stand offers no systematic guidance as to *how* to choose capabilities or domains in different contexts. Not all evaluative exercises can be open to public discussion in the same manner and it is still unclear what criteria *besides* public scrutiny there might be". Sen has argued that capabilities should be determined regarding the purpose of study and the context of society, and most importantly in a democratic environment such as public debate, or any social discussion.

engage in various forms of social interaction. Having the social bases of self-respect and non-humiliation. Not being discriminated against on the basis of gender, religion, race, ethnicity, and the like; (8) *other species*: being able to live with concern for and in relation to animals, plants and the world of nature; (9) *play*: being able to laugh, to play, to enjoy recreational activities; (10) *control over one's environment*: being able to participate effectively in political choices that govern one's life. Being able to have real opportunity to hold property. Having the right to seek employment on an equal basis with others; (11) *complete independence*: being able to do exactly as you wish without relying on the help of others; (12) *doing good to others*: being able to care for others as part of expressing your humanity. Being able to show gratitude, (13) *living in a law-abiding fashion*: the possibility of being able to live within the law; not being forced to break the law, cheat, or to deceive other people or institutions; (14) *understanding the law*: having a general comprehension of the law, its demands, and the opportunities it offers to individuals. Not standing perplexed facing the legal system."

Concerning Nussbaum's 10 capabilities, some of the results are following. First of all, several interviewees say that they feel continuously tired, and the life means something if and only if the life conditions are good (i.e., the life itself is not an important functioning). A disabled person affirmed that if the body is not perfect, it is not possible to achieve a good life. Drug addiction and sexual assault were mentioned in several cases as major problems concerning bodily integrity. Concerning sense, imagination, and thought, the interviewees agreed that the most important item is education. Emotions are also considered important by the majority of sample, however, based on some negative answers the authors recall that in hardship and misery, people can consider emotions as a burden. Concerning affiliation, for many interviewees, social relationships, being actively participate in social activities and community life are considered very important. For instance, during the interviews in Britain, one formerly homeless having had a difficult childhood in a drunken family living in slums and very poor conditions, expressed that the support from community helped morally his family a lot. Wolff and de-Shalit (2007) mention that the sense of belonging[109] somewhere (e.g., community) becomes crucial when a family is in a disadvantaged position, because many of them said that they lost social solidarity once they become disadvantaged. On the other side, there was not much agreement among interviewees on the importance of other species as functioning. But one formerly homeless mentioned that the relationship with other species especially with dogs is crucial because it helps them to retake care of themselves and learn to take responsibility. The issue of dependency appeared also very important. Many of them reported that they are completely dependent on others and they also affirmed that the dependency situation was even worse than being in extreme poverty. Authors mention that the fully dependent are mostly non-employed single mothers, disabled people, elderly people without assets. The question of having private property divided almost the sample into two opposite groups. And many of them claimed that they are not interested in politics.

Concerning the four additional capabilities proposed by Wolff and de-Shalit, first, for *complete independence,* there was a quite heterogeneity; for instance, some people mentioned that everyone is dependent on something in her/his life. Based on results, authors do not find any necessity of adding this option for further research because the 6th and 10th capabilities were

[109] Maslow (1962) emphasized that 'desire to belong' becomes important once physiological and safe needs are fulfilled. Authors also mention that for achieving a sense of self-worth, the sense of belonging should first of all be satisfied as priority human needs.

reliably answered. Concerning caring for others, humanity, and being able to show gratitude, most of answers showed that no one was against it, but they did not find it generally necessary for life. For living in a law-abiding environment, most people were rejected this capability. Many affirmed that even without law, people may be able to get a pleasant life. In opposite, concerning the 14th capability, understanding the law, general agreement on the importance of that capability was much more. For instance (2007, 59), "an Israeli lawyer who works with an Israeli NGO that offers legal services to Palestinians in the Occupied Territories, said that the least advantaged among those he worked with were those who either could not afford any legal service, or did not know the law, and did not realize that the law could often protect them. An interviewee who works with low-income families in the UK said that low-income individuals have much less access to the law. Rich people can afford having lawyers who will explain the laws to them". As we have seen, the importance of law knowledge and of accessibility to justice through improvement of legal service appeared as an important capability for leading a decent life. Nevertheless, the issue of humanity, helping others was not perceived as important despite generally people said that it is quite impossible to live without relying on someone for something in life. It is very interesting result regarding how people are in an ambiguous situation concerning social links and life independency. As we can see, the relative importance given to each capability varies as some capabilities were perceived more important than others.

Moreover, authors propose to use the terminology "(genuine) opportunities for (secure) functionings" (2007, 37), instead of the simple term functionings. First of all, we start by explaining why authors propose to use the term *secure* functionings. Their departure point is that the capabilities approach does not include "the freedom to sustain functionings" (2007, 65). However, authors insist that facing extreme risks comes first as the main disadvantage of vulnerable, poor people in their life. In other words, these people find themselves in a position of confronting by necessity high risks in their daily life in performing some functionings for fulfilling some other functionings. For instance, authors provide a case of Mr. Kedra Mia (see 2007, 65), "a Muslim daily labourer who worked in Hindu neighborhood in Dhaka, where Sen grew up as a child. While aware of, and deeply concerned about, the risk of going to look for a job in a Hindu neighborhood in troubled times, Mr. Mia had no other choice but to do so to save his family from starvation. Sometime later, very sadly, Mr. Mia was knifed on the street by Hindus, and later died." Mr. Mia took the risk, actually he was quite obligated to take his risk for being able to bring a bread to his family. In these cases, disadvantaged people put their life in danger for just being able to achieve some basic functionings. In Sen's framework, there is a question of valuing functioning; but authors insist that in such cases no one will normally have a joy of living by being exposed to danger every day. Authors distinguish three categories of being in insecure while performing functionings: (1) *risk to a specific functioning*: living in a permanent feeling of insecurity related to a functioning. For instance, authors give the example of threat of unemployment, employment insecurity[110]; (2) *cross-category risk*: a risk while performing a functioning can spread to other functionings. Authors give the example of a person who solely relies on her job for buying food may find him or herself in a situation that the risk of employment generates the risk of access to food; (3) *inverse cross-category risk*: securing one functioning leads to insecurity in another functionning. This latter case corresponds to the case of Mr. Mia mentioned above. Mr. Mia put himself in a danger for being

[110] Such issues were mentioned and investigated, particularly, in French literature notably by Robert Castel (2003, 2009) in his works on social insecurity.

able to provide food to his family. In other words, in such cases, the real liberty of a person is in danger. From this perspective, it can be seen that if their real liberty is limited, her/his capabilities of living a decent life face failure as well. This is why it is very important to not only look at the achievements, but also focus on the realization conditions of a functioning for understanding the real liberty of the person. Wolff and de-Shalit (2007, 72) define being disadvantaged as "one central way of being disadvantaged is when one's functionings are or become insecure involuntarily, or when, in order to secure certain functionings, one is forced to make other functionings insecure, in a way that other people do not have to do."

In addition to secure functionings, authors also propose to use the term "genuine opportunities" (see 2007, 74). The motivation behind this idea is that vulnerable, poor people may find themselves in a particular position of refusing some functionings due to their insecure character (see previous paragraph for *insecure functionings*). In other words, if they think that some functionings will bring them to insecure positions then they can refuse the opportunities related to these insecure functionings. In that latter case, "the individual *has the opportunity but fails to make use of it*", which is different than the case where an individual *lacks an opportunity*. This is why, authors would like to distinguish the opportunities which may bring the poor, vulnerable people to insecure functionings, compare to other opportunities. Concerning having opportunities but failing to make use of it, some discussions emerge. For instance, a person who has enough access to be well nourished, may refuse to eat sufficiently because of her/his ideological idea, her personal wish, belief; and as a consequence, the person may end up ill and is hospitalized (see 2007, 77). In that case, the person refuses by herself the opportunity of eating well.

On the other hand, authors yield another example (see 2007, 78) "where an unemployed single mother who has no savings and turns down a menial, low-paid, full-time job some distance from her home, in order to be able to see her young children to school and look after them in the holidays. She has chosen to reject full-time salaried work and let us add that she identifies with her role as an active mother who personally cares for her children[111]." In such case, the question of whether the mother has the right of refusing the opportunity of full-time job due to insecure functioning (i.e., the choice of mother is considered rational), or the mother do not take her responsibilities by refusing this opportunity, rises. Authors (2007, 79) explain that "this single mother is poor, and consequently has less opportunity for secure functioning than others. Allowing he states support even when she turns down a job will bring her closer to equal opportunity, and so a choice-based theory of responsibility, when combined with a theory of equality, would accept that this single mother, as a victim of inequality, should do not be held responsible for her choices made under conditions of inequality." If she would accept the full-time job, she could find her emotional well-being, health status decreasing, and children in risk. So, we should say that the poor has always right when they refuse the opportunities, and non-poor people are responsible of their decisions on the opportunities? Answering to that dilemma, authors clarify that if performing an opportunity will cause an undue cost (i.e., negative impact on other functionings) or a risk to be performing other functionings (i.e., *inverse cross-category risk*, see below), then the opportunity in question cannot be considered as a *genuine* opportunity. In other words, authors mention that a formal opportunity becomes genuine opportunity if and

[111] Among the interviewees, authors mention that a single mother in Israel has to refuse a better job opportunity for taking care of her children because she afraid that her children would be drop school, and become attracted to crime (2007, 206).

only if it is reasonable for the person to perform this opportunity (i.e., for authors, opportunity corresponds to the capability).

As we have seen from previous paragraphs, notably with her list of the central human capabilities, the philosopher Martha Nussbaum has worked very closely with Sen for developing and promoting capabilities approach. She contributed on numerous works on capabilities (1997, 2000, 2002a, 2002b, 2003, 2004). Nussbaum has employed qualitative research, notably the narratives for explaining capabilities. She has been working on the use of capabilities in political area with the aim of building one particular list of capabilities that each government should possess in her constitution and should provide it adequately to their citizens[112]. She is particularly interested in women's capabilities and poverty status and the link between poverty and inequality, based on various narratives and qualitative works[113]. She argues that the classical economic measure of well-being, the GNP is far from capturing adequately inequalities, equal distribution and diverse aspects of development such as education, health, and notably the living conditions of women. For having a comprehensive understanding on the quality of life of people, and for being able to compare the well-being of nations, she suggests using capabilities approach focusing on "what people are actually able to do and to be" (2006, 50), in other words focusing on the real liberty of people. Linked to women situation, she expresses that "agency and freedom are particularly important goals for women, who have so often been treated as passive dependents"[114] (Nussbaum 2006, 50). Regarding the inequalities (e.g., regarding opportunities, resources, rights) women face on their daily basis within family and society, from Nussbaum's perspective, it is very important to focus on the '*equality of capabilities*' as a "central goal for women who pursue social justice" (2006, 51)[115]. However, she finds Sen's conception of freedom too vague, and she mentions that a definite list of capabilities is necessary for making clear the notion of freedom especially regarding gender perspective because "some freedoms limit others; some freedoms are important, some trivial, some good, and some positively bad" (2003, 33). Regarding her list of the Central Human Capabilities, she first mentioned it in her work *Women and Human Development* which is published in 2000. These ten central capabilities (i.e., (1) life; (2) bodily health; (3) bodily integrity; (4) senses, imagination, and thought; (5) emotions; (6) practical reason; (7) affiliation; (8) other species; (9) play; (10) control over one's environment- for the detailed explanation of each item see Nussbaum (2006, 58-59; 2003, 41-42) are proposed by Nussbaum as essential conditions to be fulfilled in order to live with dignity. Moreover, she mentions that these items can be specified later by the society in question.

In addition to Sen and Nussbaum's contributions and to other works mentioned above on capabilities, in the literature there is a growing body of research on capabilities from both theoretical and application perspectives such as the contributions of Basu (1987), Desai (1995), Gasper (1997), Qizilbash (1996), Stewart (1996), Chiappero-Martinetti (1996, 2000), Saith (2001), Fleurbaey (2002, 2006), Kuklys (2005), Robeyns (2003, 2005a, 2005b, 2005c, 2006a, 2006b, 2008), Fukuda-Parr (2003), Burchardt (2004), Qizilbash and Clark (2005), Wagle

[112] Nussbaum's interest in the partial theory justice is related with her works on one specific list of capabilities (see Nussbaum 2006, Robeyns 2017).
[113] However, Sen mainly employs quantitative approach to studying capabilities.
[114] Nussbaum suggest seeing Chen (1983), and Agarwal (1994) for discussions on women's agency, freedom, and participation.
[115] Related to that, Nussbaum points out the prominence of capabilities approach as a "relevant space of comparison when justice-related issues are considered" (2003, 33).

(2005, 2008, 2009), Anand and van Hees (2006), Anand et al. (2005, 2008, 2009), Chiappero-Martinetti and Roche (2009), Alkire (2002, 2005, 2007a, 2007b, 2008a, 2008b, 2008c, 2015), Alkire and Black (1997), Clark (2005, 2008), Comim et al. (eds.) (2008), Arndt and Volkert (2011), Burchardt and Vizard (2011), Hick (2012, 2014), Burchi et al. (2014), Vizard and Speed (2015), Vizard et al. (2011), Vizard (2006), Rippin (2016), Gangas (2015), Mitchell et al. (2017).

We will finish this section by shortly focusing on the link between capability approach and human development (see Alkire 2002, 2007a, 2007b for different dimensions of human development and missing dimensions of poverty data). This latter plays an important role in the literature of capabilities since the computation of human development based on capability approach was the first quantitative attempt of operationalizing capabilities (see Anand and Sen 1994, 1997; Fukuda-Parr 2003; Fukuda-Parr and Kumar (eds.) 2009, Deneulin and Shahani (eds.) 2009, see also Deneulin and Stewart 2002 for a discussion on the contribution of capabilities on development)[116]. With the initiatives of Mahbub ul Haq and Amartya Sen, the first human development report published in 1990 by the United Nations Development Program (see UNDP 1990, Anand and Sen 1993), presented this challenge. Based on Sen's capability approach, the human development index (HDI onwards) was computed and presented in the report. At the time, development was massively related to income (e.g., gross domestic product-GDP) and economic growth performances of countries, notably supported by the World Bank with its annual World Development Reports (e.g., World Bank 1990, 2001, 2009). Sen and ul Haq's vision of development from social perspective have opened new rooms for conceiving and measuring development in the sense that they promoted social, economic, political rights as the development strategy in opposition to expanding economic growth and basic needs approaches.

Certainly, in somewhere, money is useful for buying some goods and services; however, it does not cover all necessities neither helps to buy everything in life. In opposition to economic approach on development where the money is at the center of development, Sen and ul Haq took it as an instrument among many others (e.g., education, health) for reflecting capabilities of people in different aspects such as economic, social, cultural given human diversity. Herewith, human development approach becomes multidimensional and plural (see Alkire and Deneulin 2009, 28, and see Desai 1991, Kelley 1991 for more comprehensive description of human development and its measurement). Sen (2003, pvii) encapsulates concisely the perspective of human development as "on one side, to a clearer comprehension of how-and in what ways-human lives can go much better and, on the other, to a fuller understanding of how this betterment can be brought about through a strengthening of human agency. I shall call them, respectively, 'the evaluative aspect' and 'the agency aspect' of human development." In this light, the 1990 report defines human development[117] as "both the process of widening people's choices and the level of their achieved well-being[118]" (UNDP 1990, 9), such process valued social development (e.g., education, health), new political reforms in favor of

[116] See also Gasper (2002) for a discussion on the use of capability approach for explaining human development, and Streeten (1994) for a description of human development based on means and ends.

[117] Regarding ul Haq, human development is based on 4 principles: equity, efficiency, participation and empowerment, and sustainability. We don't entre in detail here, see Alkire and Deneulin 2009, 29-30 for more explication.

[118] In the literature, the notion of wellbeing is sometimes expressed as wellbeing or as well-being depending on works. We prefer to use as well-being; but in citations we keep as she is written by authors.

democracy, implementation and maintenance of human rights to provide a life of freedom and dignity to everyone.

The human development index is aimed at measuring development level of countries and rank them (i.e., cross-country comparisons) according to it. Three dimensions which are health, education, and living standards were included in the index. The indicators are as follows: life expectancy at birth, a weighted average of the literacy rate and the combined gross enrollment rate, and per capita GDP adjusted for purchasing power parity (see UNDP 2015 Technical notes for detailed information on the construction of the index, see Acharya and Wall 1994 for the evolution of the HDI, see Streeten 1995 for discussion on the HDI). Of course, the HDI does not include a large variety of capabilities, but she achieved to associate development with education and health as alternative to income approach (however, it is criticized that the HDI has kept GDP as indicator)[119]. Ranis et al. (2006) have proposed additional 11 dimensions of human development: mental wellbeing, empowerment, political freedom, social relations, community well-being, inequality, work conditions, leisure conditions, economic stability, political security and environmental conditions. Few availabilities of internationally comparable data were one of the problems for not having included more indicators in the HDI. On this issue, Alkire (2007b, 348) identified some missing dimensions of wellbeing which are: "employment (particularly informal employment and safety at work), empowerment or agency (the ability to advance goals one values and has reason to value), physical safety (focusing on security from violence to property and person, and perceived violence), the ability to go about without shame (to emphasize the importance of dignity, respect and freedom from humiliation), and psychological and subjective well-being (to emphasize meaning, its determinants, and satisfaction." Moreover, Dasgupta (1990), Tilak (1992), Sagar and Najam (1998), Morse (2006) portrayed the limits of the HDI vis à vis its dimensions (cited in Hirai 2017, 74-75) , Streeten (2000) suggested further research areas in human development, and Ranis and Stewart (2012) summarized achievements and deficits of human development approach.

In this section, we have already introduced multidimensional perspective to poverty based on capabilities. In section 2.4, we will examine multidimensional approach through different measures existing in the literature. Notably, we will see the Multidimensional Poverty Index proposed by the United Nations Development Programme in their annual human development report and a number of national multidimensional measure of poverty. But before, we will discuss whether the capabilities is absolute or relative approach on poverty and what is the link between capabilities and social bonds in the sections 2.2. and 2.3, respectively.

2.2 Capabilities: absolute or relative approach on poverty?

When it comes to measuring poverty, the literature is known for a substantial debate on absolute vs. relative approaches on poverty. World Bank's 1USD, 1.25USD, 2USD per day monetary thresholds are the most popular tools to define and measure poverty from an absolute

[119] Sen (2005, 336-337) mentions that Mahbub ul Huq and his intention was proposing some relevant capabilities but not a fixed list, he says that "the ''human development index'' (the HDI) was based on a very minimal listing of capabilities, with a particular focus on getting at a minimally basic quality of life, calculable from available statistics, in a way that the GNP or GDP failed to capture. Lists of capabilities have to be used for various purposes, and so long as we understand what we are doing (and in particular that we are getting a list for a particular reason, related to a particular assessment, evaluation, or critique), we do not put ourselves against other lists that may be relevant or useful for other purposes".

perspective. There is also the approach of basic needs promoted by the United Nations in the late 1970s to define poverty as the lack of access to some essential elements of life such as drinking water, food, education, health services. United Nations Latin America Economic Commission's (ECLAC) unmeet basic needs approach is also very close to the basic needs and ECLAC's approach is still widely used in the region of Latin America. Absolute poverty and basic needs approaches require the list of essential items, which are considered as the absolute items that every single human being should have wherever she lives. The monetary threshold defined by the World Bank or other similar thresholds such as those of Rowntree (in his study of poverty in city of York, 1901) are defined on the basis of the monetary values of such essential items. In the literature, there is a debate on whether a certain list of basic needs/absolute terms is realistic given the specific country contexts or given the specific society requirements. For instance, Townsend's (1979) famous relative deprivation approach argues for the necessity of defining the essential items of life on the basis of the society's situation or in other words on the basis of social necessities. For instance, social life practices might be different from one country to another, or even personal perception of basic needs might differ as well among individuals within a society or between societies. Sen has also expressed his point of view on this debate. He particularly emphasizes the importance of basic capabilities in measuring poverty in developing countries but at the same time he is not against to consideration of country, community, region contexts in defining the multidimensional measure of poverty. Before showing the ideas of Sen on this issue, first we will present in detail the approaches of absolute poverty, basic needs, and relative deprivation, respectively.

Absolute poverty

Under the influence of pioneer works in subsistence, quantitative methods such as descriptive statistics and socio-economic surveys have become important tools for describing and analyzing poverty. Especially, statistics on daily nutrition, detailed budget statistics including various types of charge and income have become the core of new quantitative poverty measurement wave. Notably, these works have stimulated the analysis of living conditions of labourers. Undoubtedly, one of the famous (urban) quantitative poverty studies is of Charles Booth (see Booth 1902-3 [1889, 1891, 1892, 1892-7]) in the 1880s. Charles Booth is influenced by Henry Mayhew[120] (see Mayhew 2008 [1852]) and Friedrich Engels' (1845)[121] studies on the working poor in the mid-nineteenth century in Britain. Mayhew's nineteenth-century classic titled *London Labour and the London Poor*, has described in detail the living conditions of the poor and of the non-employed people (about 150,000 criminals and 110,000 vagrants). These people were mostly living on charity support and social assistance, and they get used to live as dilly-dallier and to begging. Mayhew has identified three groups of non-workers: those with a 'physical defect', those with an 'intellectual defect' (e.g., lunatics, idiots), and those with 'some moral defect' (e.g., indolent, vagrant, professional mendicant, criminal). Mayhew has also reflected hard and precarious living conditions of the Irish people in multi-aspects. Especially, many of them were farm hands in their country. We also know that in the early 1870s, the United Kingdom faced an influential depression that caused mass unemployment, and a great number of farmers migrated to the cities from Ireland and England countryside. However, when they came to London, they had severe difficulties to find a secure job. For that reason, they had

[120] Originally, Mayhew's book is from 1852. But, in the Bibliography, the recent publication of this book (in 2008) by Wordsworth Editions is given.

[121] It is worth to note that Engels' works take an important place in Marx's ideas on pauperism, which is based on economic perspective.

worked most of time as street vendors. For instance, they were selling fruits, nuts, hazelnut, orange[122]. Generally, they did not achieve in education. Young girls and women did not know how to sew and to do domestic works because they were living in lamentable conditions. Mayhew has explained why the Irish people were become street vendors for four seasons. In his works, he also focused on subsistence, diet habits. In the literature, the notion of subsistence is related to survival- the minimum needs for sustaining the life. Mayhew has collected information regarding the poor's food and beverage consumption, daily expenses. As we can see Mayhew has used both qualitative and basic quantitative methods to show pauperism in London.

Concerning Engels, with his work titled *The condition of the working class in England* (1845, first edition in German), he has focused on the life of industrial workers in terms of their wage and living conditions. His work is based on his own observations and detailed reports. He has found interesting results for instance, high mortality rate in big industrial cities such as Manchester, Liverpool. He has also noticed that the living standards of industrial workers were worse in many perspectives (e.g., low income, health problems) compared to previous years despite the industrial revolution. Similar to Mayhew (in Victorian England, distinguishing the morally '*deserved*' poor from the '*undeserving*' poor) and Engels, Booth has studied living conditions of the poor people in a series of seventeen volumes entitled *Life and labour of the people in London* (1902). His first volume is only based on the data collected in East Est. Then, with his team, for other volumes, he extended the data collection process into different areas in London. Before Booth, the poor was seeing as a *social class* in British society, or as the pauperism. This point of view has been promoted by charity organizations and its defendants, arguing that the poor were morally responsible for their situation, and the poor were unable to lead a normal life due to their social inefficiency (see the theory of social inefficiency, for instance, Bowpitt (2016 [2000], 25). However, with Booth's works, poverty has been seen as a *social problem.*

Booth was particularly affected by poverty and unemployment situation in London. The motivation behind his work is mainly due to social workers and philanthropists' insufficient answers concerning how the poor live on their daily basis. At that time, there was neither comprehensive investigations on poverty nor decomposed analysis in terms of income level and/or occupational group. Booth was searching for the efficient solutions for helping the poor. He was primarily interested in the nature of poverty in general and the current problems surrounding the life of the poor, rather than in a historical or a biographic analysis of the poor. Based on these ideas, he decided to carry out a large-scale survey among the poor for understanding their dissatisfaction and current living conditions. He was also known as a frequent visitor of East Est where he was passing his time by observing behaviours of the poor. Yet he rented a room in the house of a (typical) poor family for more knowing about the poor family's life conditions and for witnessing the family's conversations and struggles (Fried and Elman 1969, 20). As a result of his own observations, he often was contradicting with the opinions and survey results of journalists, charities, or religious. He was believing that the extent of poverty was not as abject as shown by the reports. To prove his hypothesis, he was convinced that only a quantitative approach with statistical techniques would provide accurate, objective, and scientifically pertinent results. With these ideas in his mind, he conducted his

[122] The season of orange was known as the Irish harvest because generally oranges were selling by Irish street vendors.

own survey focusing on the lives and occupations of all the inhabitants in London (but also, he used other sources)[123]. It is worth to note that his analysis lasted seventeen years[124]. In his book, he has used "compendium charts, descriptive statistics, maps, qualitative descriptions of families, homes, streets, conditions of work" (Fried and Elman 1969, 25)

Booth has been the first researcher to use a large-scale survey, a methodological process for putting in evidence the nature and causes of poverty. He has proposed *poverty lines* for determining the number of poor people in London. More specifically, his work consists of descriptive statistics with the aim of showing the reality of the poor in terms of housing, clothing, eating and shopping habits, employment. He has classified the poor by combinations of income and occupation[125], which is about the eight classes labelled from A to H as follows: (A) lowest class of occasional labourers, loafers, and semi-criminals; (B) casual earnings-very poor; (C) intermittent earnings; (D) small regular earnings; (E) regular standard earnings – above the poverty line –; (F) high class labour; (G) lower middle class; (H) upper middle class. People having an income level around 18 to 21shillings per week correspond to the class D, and people who fall below the 18 to 21shillings income threshold are considered as the class C. Booth wrote "my 'poor' may be described as living under a struggle to obtain the necessaries of life and make both ends meet; while the 'very poor' live in a state of chronic want" (Fried and Elman 1969, 55). The classes C and D are considered as tolerable poverty situations, and the classes A and B are seen as the most miserable desperate situations.

Booth's findings are various, they have contributed to the literature from different angles. Through his first book (1889), first of all he has showed that the ratio of employed people living in the acute poverty is dramatically high compared to the ratio computed by the Social Democratic Federation. He has proved the existence of precarity in London. Second, he has showed that poverty is a question of social failure. At the time, among public, media, and government the idea of poverty as a result of lack of personal abilities and of skills was very popular. Poor people were blamed for their behaviours and were seen as deserving misery. According to Booth's statistics, 85 per cent of the situation of the poor is due to unemployment, low pay, health issues, or large family. He has recommended eliminating or reducing significantly the class A, which would make possible to save the class B from starving and to move the classes C and D above the poverty line. Besides he has emphasized the importance of improving the functioning of trade unions and cooperatives, which should welcome all working people and not only the elites of labour class. Booth and his team, in second volume, have more focused on the construction of poverty maps by breakdown of classes. Later, he has investigated the industrial life of the poor. In his latest works, he has also showed that the elderly is the most vulnerable group among the paupers. For that reason, he has recommended the old age pensions for helping the elderly to gain independence and to lead a more pleasant life.

From both methodological and conceptual perspectives, the works of Booth on poverty deserves a particular attention. If we summarize it, we can say that he has made clear that poverty touches various aspects of life such as shelter, clothing, daily calorie intake but also working conditions, health issues, immigration, situation of streets. He has found that 35.2 per

[123] Sources such as census between 1881 and 1891, reports, school administration data, information collected by charities, officers.
[124] All volumes of his research published in one book in 1902 entitled *Life and Labour of the People in London*, 1889-1902, 17 volumes.
[125] He listed 39 occupations.

cent of people living in East Est and 30.7 per cent of the London population were living under the poverty line. As Fried and Elman have mentioned (1969) that thanks to Booth's spectacular work (his observations, findings, recommendations to eliminate poverty), British people have got the opportunity of learning about the concept of poverty in terms of how the poor live in miserable conditions in London. Fried and Elman (1969, 30) wrote that "…Never before had the middle classes been told in such harrowing detail about the effects of moral decay and destitution, about the domestic lives of the poor, about the oppression of work, the condition of women workers, the practice of sweating, about the new immigrants." Despite many acknowledgments, there are also come critics regarding Booth's methodology. For instance, Veit-Wilson (1986), and Spicker (1990) have argued that Booth's concept of poverty is quite relative, and it is based on Booth's subjective perceptions on the basic standards of life in Britain.

Nowadays, one of the most frequent research questions in poverty studies is 'how we can define or identify the poor'. Booth has neither answered explicitly to this question, nor used a specific definition of poverty. For collecting the data, he directly went to the places (e.g., East Est) where there was a visible poverty in terms of street situation, shelter, clothing styles of the indigents, or he referred to what he heard about streets and to what he saw on the news. He has chosen East Est because of its reputation as a place of the poor and he has not used income thresholds[126] to select the location of the study. However, he has settled poverty lines and poverty classes as combinations of occupation and income. We will see later that his poverty line approach has been an inspiration many poverty studies such as those of Rowntree, Townsend. Moreover, Booth's empirical work has largely influenced the opinion of British people and authorities on poverty (for example, Booth's work has had a significant impact on the implementation of the Old Age Pension Act in 1908 and the Minority Report on the Poor Laws in 1909) and he has been acclaimed in the academic and political areas for his sophisticated quantitative methodology to measure poverty. Booth has expressed that "my object has been to attempt to show the numerical relation which poverty, misery, and depravity bear to regular earnings and comparative comfort, and to describe the general conditions under which each class lives…" (Fried and Elman 1969, 48).

A few years later, Seebohm Rowntree (1901[127]) has impressed British society with his book '*A study of poverty in the city of York'* written based on the results of his first investigation on poverty and socio-economic conditions among all working-class families in the city of York. Like Booth, he has carried out a social survey[128] and used quantitative approach for describing in depth the daily life of labourers. It is often mentioned that Rowntree was impressed by Booth's work; however, Rowntree's work is much more comprehensive regarding the data (notably the collection of income data). Booth's data are based mainly on households with school-age children (i.e., data was collected through School Board visitors). Whilst Rowntree and his team have also investigated each working household in the city of York (see Briggs 1961, Hennock 1991). Rowntree has defined two types of poverty: '*primary poverty*'

[126] We will see later that, in contemporary poverty studies, poverty line is the main tool for identifying who is poor compared to the geographical criterion (i.e., reputation of neighborhood, street, zone in a city because of its visible miserable living conditions and destitution). Booth used in his early works on poverty in London.

[127] Later, in 1941, he showed the results of his second survey.

[128] This is a spectacular survey in a sense that all labourers (and their households) in city of York were interviewed. Rowntree launched the survey in January1899.

corresponds to having insufficient family earnings for meeting the "minimum necessaries for the maintenance of merely physical efficiency[129]" (1901, 86), on the other hand '*secondary poverty*' refers to having sufficient family earnings given that some of the earnings are spent for other useless consumptions such as "drink, betting, and gambling, ignorant or careless housekeeping, and other improvident expenditure, the latter often induced by irregularity of income" (1901, 142). Rowntree wanted, therefore, to distinguish families suffering from insufficient income and families living improvidently despite a sufficient amount of family income. Like Booth, he has described the living standards of the poor (e.g., food, housing, shelter, health, family budget, poor relief) that he mentioned as "a picture of the kind of life they live" (1901, 32). To show the misery, he wrote that "…Housed for the most part in sordid streets, frequently under overcrowded and unhealthy condition, compelled very often to earn their bread by monotonous and laborious work, and unable, partly through limited education and partly through overtime and other causes of physical exhaustion to enjoy intellectual recreation…" (1901, 144-145). He has also partitioned the population of York into seven classes[130]. How did he measure primary poverty? First of all, Rowntree has developed a methodology to determine to what exactly correspond (i.e., to which items) the minimum necessaries for the maintenance of merely physical efficiency- Second, he has developed a methodology to measure family earnings. He measured family income/earnings based on wages (see 1901, 29 for more details). At that time, only wages and outdoor relief were existing[131]. He has considered the items such as "food, house rent (including rates), household sundries (such as clothing, light, fuel, etc.)", which he called the minimum necessaries[132], and then he

[129] But we know that previous to Rowntree, in 1886 Henrietta Barnett has provided a comprehensive study on physical efficiency including workhouses and army's dietaries, the necessary weekly expenditure for a family of eight, and the cost of other diet samples (cited in Gillie 1996, 724). Thus, Rowntree's diet is not the first way of operationalizing the physical efficiency in the literature.

[130] From class A to class G respectively: (a) total family income under 18s. for a moderate family, (b) total family income 18s. and under 21s. for a moderate family, (c) total family income 21s. and under 30s. for a moderate family, (d) total family income over 30s. for a moderate family, (e) domestic servants, (f) servant-keeping class, (g) persons in public institutions (see 1901, 28 for details), where a moderate family is composed of a father, a mother, min. 2 or max. 4 children. The class A is the poorest class.

[131] For a provincial city as York, it was not very difficult to gather information on wages paid by the major employers. As the city is not huge as London, employers were known generally. And, at that time there were neither deductions nor benefits, which facilitated the measurement of wages. He has just taken in consideration the short-time allowances. Meanwhile, Rowntree has also tried to collect direct information on wages during the interviews; otherwise, he has estimated the wages of skilled workers. Concerning unskilled workers, he has got information directly from employers.

[132] According to Rowntree, a family living with the minimum necessaries should never buy a train ticket or an omnibus, the family members should never go to campaign or they can only walk around, they should never buy a newspaper, they should never buy a concert ticket, they cannot write the letters to their children left the home because they cannot afford stamps, they should never donate church or chapel or help friends or neighbors, they cannot save the money, they cannot afford a health insurance, they cannot subscribe to a syndicate, the adults cannot give the pocket money to their children for buying sterns, balls, or candies, the father of the family cannot drink a beer or smoke, the mother of the family cannot buy herself and her children beautiful dresses because they do not have enough money for affording all these activities (see Rowntree 1901, 133-34 for details and complete description).The minimum necessaries can be seen as the basic needs according to Rowntree: the basic needs should contain only the must have items and the items necessary for maintaining health. In addition, Rowntree accepts that regarding intake calories per day, he was even more severe than the criteria of the Poor Law. Moreover, Rowntree has updated the list of minimum needs in due course. For instance, in his

has estimated the cost of the items in terms of expense (see 1901, 88-89). He has mentioned that "expenditure needful for the development of the mental, moral, and social sides of human nature will not be taken into account at this stage of the inquiry" (1901, 87). As Booth and other Victorian nutritionists, food was one of the most important items in Rowntree's poverty measurement process. For that, he has gathered calorie information based on the function of food in organism, and the minimum quantity of food necessary for men, women, and children. Once Rowntree has defined calories, then he has computed the cost of diet by men, women, and children (see 1901, 88). He collected the necessary information on food by asking experts, on housing by examining expenditure patterns, and on clothing by checking public opinion. Rowntree has provided a table of select dietaries for men, women, and children (see 1901, 99-102). A small sample of these dietaries, as an example, is as follows:

Table 4. Dietaries by men, women, children, drawn by Rowntree (1901, 99-102)

Day	Breakfast	Dinner	Supper
MEN			
Sunday	Bread, 8 oz. Margarine, ½ oz. Tea, 1 pt.	Boiled bacon, 3 oz. Pease pudding, 12 oz.	Bread, 8 oz. Margarine, ½ oz. Cocoa, 1 pt.
Monday	Bread, 8 oz. Porridge, 1 1/2 pts.	Potatoes with milk, 24 oz. Bread, 2 oz. Cheese, 2 oz.	Bread, 8 oz. Vegetable broth, 1 pt. Cheese, 2 oz.
WOMEN			
Sunday	Bread, 6 oz. Margarine, ½ oz. Tea, 1 pt.	Boiled bacon, 3 oz. Pease pudding, 8 oz.	Bread, 6oz. Margarine, ½ oz. Cocoa, 1 pt.
Monday	Bread, 4 oz. Porridge, 1 pt.	Potatoes with milk, 20 oz. Bread, 2 oz. Cheese, 1 1/2 oz.	Bread, 6 oz. Vegetable broth, 1pt. Cheese, 2 oz.
CHILDREN, AGES 8-16			
Sunday	Bread, 6 oz. Margarine, ½ oz. Tea, ¾ oz.	Boiled bacon, 3 oz. Bread, 3 oz. Potatoes, 8 oz.	Bread, 6 oz. Margarine, ½ oz. Cocoa, ¾ pt.
Monday	Bread, 3 oz. New milk, 1/" pt. Porridge, ¾ pt. Sugar, ½ oz.	Potatoes with milk, 16 oz. Bread, 2 oz. Cheese, 1 ½ oz.	Bread, 6 oz. Vegetable broth, ¾ pt. Cheese, 1 ½ oz.
CHILDREN, AGES 3-8			
Sunday	Bread, 5 oz. Milk, ½ pt.	Boiled bacon, 1 oz. Potatoes, 4 oz. Skim milk, ½ pt.	Bread, 5 oz. New milk, 1/2 pt.
Monday	Bread, 2 oz. New milk, ½ pt. Porridge, ½ pt. Sugar, ½ oz.	Potatoes with milk, 12 oz. Bread, 2 oz. Cheese, ½ oz.	do.

second study, he has included new items such as radio, newspaper, gift for children and holiday. As demonstrated by Veit-Wilson (1986), Rowntree in fact has used relative measures to define absolute needs for leading a pleasant life. See also Alcock (2006 [1993]) for some discussions on that.

The following table shows the weekly cost of a standard dietary, considered as the (weekly) poverty line (1901, 105):

Table 5. The weekly cost of a standard dietary for men, women, and children, computed by Rowntree (1901, 105)

Men	**Women**	**Children, 8-16.**	**Children, 3-8.**	**Children under 3.**
3s. 3d.	2s. 9d.	2s. 7d.	2s. 1d.	2s. 1d.
Average for Adults 3s.		Average for Children 2s. 3d.		

Once Rowntree has gathered all necessary information on wages, and cost of minimum necessaries, then he has compared whether families of different size have enough earnings or for affording minimum necessaries (per week). For instance, Rowntree has determined the minimum amount (in terms of shillings and pence) of income needed for affording food, rent, household sundries for different types of family (see the table below for a sample of minimum necessary expenditure per week, namely the *primary (weekly) poverty line*). The total amount, as the cost of minimum necessaries, is considered as the *absolute poverty line*. Thereby, families having earnings below the absolute poverty line are considered primarily poor[133]. Regarding secondary poverty, Rowntree measured it based on his observations.

Table 6. Minimum necessary expenditure per week for families of different size, called the primary (weekly) poverty line drawn by Rowntree (for a full coverage see Rowntree (1901, 110))

Family	**Food**	**Rent[134]**	**Household Sundries**	**Total**
1 man	3s.	1 s. 6d.	2s. 6d.	7s.
1 woman	3s.	1s. 6d.	2s. 6d.	7s.
1 man and 1 woman	6s.	2s. 6d.	3s. 2d.	11s. 8d.
1 man, 1 woman, 1 child	8s. 3d.	2s. 6d.	3s. 9d	14s. 6d.
1 man, 1 woman, 2 children	10s. 6d.	4s.	4s. 4d.	18s. 10d.

Rowntree is interested mainly in 'what is the percentage of population living in poverty?' (i.e., people under the poverty line). He has estimated that for a working-class family with three children, it won't be possible to escape from primary poverty for almost a decade. He has characterised five specific periods of poverty for a labourer; this process is known as 'cycle of poverty'[135]: "people are more likely to be in poverty in childhood, when they are parents with

[133] He has found that 9.9 per cent of population of the city of York live below the primary poverty line.
[134] Concerning rent, Rowntree has mentioned that "in the calculations, actual rent has in every case been taken as representing the minimum necessary rent expenditure. Numbers in this column are a rough approximation of the average rent paid in York by families of various sizes" (1901, 110).
[135] Rowntree has estimated that, for a labourer, the life is about five successive periods: (1) if the father is not a qualified labourer the child will live in poverty during childhood; (2) when the child and siblings will work, the situation of the family will be better and even the family income may exceed the poverty line. The child will live for a while with parents for saving some money; (3) when the child will get married, he or she can use savings. After marriage, until having children, he or she will live in prosperity; (4) but after having children he or she will return to poverty, which will last approximately a decade

their dependent children, when children leave home and marry, and when they are no longer able to work" (Bradshaw 2000, xxxiv[136] in centennial edition of Seebohm Rowntree's *Poverty: A study of town life).* He has looked for the immediate causes of primary and secondary poverty. He has found that 51.96 per cent of people living with primary poverty work regularly but they perform low waged jobs; 22.16 per cent of people living with primary poverty live in large family (i.e., more than 4 children); and 15.63 per cent of people living with primary poverty has faced death of the chief wage-earner in their family. He particularly has emphasized that the issue of low wage is the most important element of primary poverty. He also has found that high rent is the most important part of household expenses; for instance, he has found that families earning less than 18 shillings per week spend 29 per cent of the family budget for paying rent. Furthermore, the mortality rate amongst the poor living in the slums is dramatically higher compared to wealthier areas in London.

On theoretical level, Rowntree's most important contribution is the concept of 'physical efficiency' and the collection of detailed information on housing, wages, and occupations to explain poverty. Thanks to his fieldworks, he has stimulated evidence-based policies. As Booth, Rowntree also has showed that poverty is a structural problem due to low income/earnings rather than a behavioural problem. In 1936, almost thirty-five years after the first survey, Rowntree has conducted his second research in York entitled *Poverty and progress: A second social survey of York* (1941). In his second survey, he has added new items to the list of absolute needs, such as radio, books, newspapers, beer, tobacco, present and holiday. He then has conducted in 1950[137] a third survey on poverty in York. The results of Rowntree's final poverty study in York (Seebohm and Lavers 1951) have showed a lower per centage of poverty in York. This result demonstrates the positive impact of welfare state reforms and Beveridge report[138] (Beveridge 1942) to reduce poverty in the United Kingdom.

In terms of public policies, the Beveridge Report (1942) has used the updated version of Rowntree's food regime/diet scheme. Moreover, we can still witness his influence in contemporary poverty studies in particular for determining budget standards and food poverty measurement (Parker 1998). Rowntree has argued that family allowance is a very important element for increasing the living conditions of the poor and the allowance makes social and active the poor in society. Bradshaw (2000, xiv in centennial edition of Seebohm Rowntree's *Poverty: A study of town life*) has mentioned that the structure of social assistance in Britain is based on Rowntree's works on poverty. However, Rowntree's work has also been criticized in some views by other scholars. For instance, Bosanquet (1902) has argued that Rowntree's poverty line is based on subjective judgments and the overall measure of poverty depends on Rowntree and his team's impressions during the short house-to-house investigations (cited in

until the first child has fourteen years and starts working. If he or she has more than three children, the difficult phase until the better situation may prolong; (5) late in life, he or she will fall again in poverty because the person could not save enough money in the past and he or she will not be able to work in the old age.

[136] This quotation is from the preface of Jonathan Bradshaw in the centennial edition of Seebohm Rowntree's *Poverty: A study of town life* published by the Policy Press and the Joseph Rowntree Charitable Trust in 2000 (xix-xxxii).

[137] The study is published in 1951.

[138] The Beveridge Report is the fundamental element of the British welfare state. The report has suggested criteria for national insurance (see Gordon and Spicker (eds.) 1999, 17).

Welshman 2006, 46-47; see also Bowpitt 2016 [2000], 29)[139]. Gillie (2000) has made clear that Rowntree's poverty line is certainly not the first one because previously Barnett (1886) and Scott (1889) have already computed poverty lines by employing a nutritional approach and McDougall (1885) has computed poverty lines in the mid-1880s.

As we have seen, in the works of Rowntree, Mayhew, Booth, the focus is on the life of labourers and especially on unemployed. Labourers in industrial cities were used as the object of study in order to describe what poverty is and how poor people face this kind of life in terms of their daily nutritional habits, residential conditions and expenses. In this direction, in 1915, Bowley in collaboration with Burnett-Hurst (and Tawey with his introduction) published their poverty investigation with the aim of describing the economic conditions of working-class households from four middle size cities, namely Northampton, Warrington, Stanley, and Reading[140]. Bowley and Burnett-Hurst have collected data, which is comparable with the data collected by Rowntree in 1899. While Rowntree has investigated each working-class household in York; Bowley and Burnett-Hurst have constructed a random sample of population as "approximately one working-class house in 23 being visited at Northampton, one house in 19 at Warrington, one house in 17 at Stanley, and one house in 21 at Reading" (1915, 11). Given the size of these four cities, of course the work of Bowley and Burnett-Hurst cannot reflect the industrial life in big cities such as London, Liverpool, Manchester. But, with their work, they have introduced the representative method in social surveys in Britain.

The relationship between industrial life and poverty is at the heart of Bowley and Burnett-Hurst' s study. Authors have mentioned that regarding industrial characteristics, there are two types of towns: (i) those dominated by one staple industry, where the minor industries are complementary; (ii) those with a variety of co-existent and independent commerce (1915, 12). These two types of towns[141] are included in their sample. To reflect the living standards of the workers, authors have showed by tables (i.e., including total effective numbers and percentages) the rent of working-class houses, the working-class houses in terms of number of rooms, the number of overcrowded houses, the link between rent and family income in terms of median rent as a percentage of income, total earnings per week by family composition, full-time weekly wages of adult males. In addition, they have compared the results with the results of Rowntree's study in city of York in 1899.

Bowley and Burnett-Hurst have used the absolute poverty line approach (i.e., called *standard*) to find out the percentage of households living above, living at standard, living below the poverty line. They have called 'richer households' for those living above the standard (1915, see Table VIII on page 39 for all numbers by four towns). They have used two types of poverty lines: (i) Rowntree's line (1901); (ii) the poverty line they have adapted from Rowntree's line, which they called *New Standard*. Bowley and Burnett-Hurst have mentioned that "Mr.

[139] Bosanquet has mentioned that "it is obvious that impressions obtained by such flying visits must be almost valueless" (1902, 12-13). For more discussions on that, see Bosanquet (1902) and for Rowntree's answers to Bosanquet, see Rowntree (no publisher or date Briggs (1961) gives 1903 but British Library of Political and Economic Science catalogue gives 1901-cited in Veit-Wilson 1986, 99).

[140] Enquiries were conducted during the summer and autumn of 1913 in Northampton, Warrington, and Stanley. In Reading, the enquiries were conducted in the autumn of 1912.

[141] Northampton is characterized by shoe industry, Stanley is characterized by coal extraction, Warrington has foundries, metallurgic industry, textile factories, tanneries, a soap factory, and breweries, Reading has a large railway center, mechanical industries, prints, important biscuit sector. As we can see Warrington and Reading have a larger variety of economic activities.

Rowntree took the cheapest possible budget, which was largely vegetarian. The New Standard allows for some expenditure on meat and arbitrarily raises the adult male standard by 9d. a week for his purpose." (1915, 37). Herewith, the authors first have calculated the total income for every working-class household and then, they have looked at whether the total income exceeds or falls short of the poverty lines. They have mentioned that households being below the Rowntree's poverty line and below the New Standard at the same time, are the families living in the workhouses or in other Poor Law institutions (i.e., we recall that these families are named the *paupers*). Bowley and Burnett-Hurst have also interested in immediate causes of primary poverty such as the percentage of households below the Rowntree Standard by immediate causes. Their main result is "one-half of the households below the poverty line at Warrington and Reading, nearly one-half at York, and one-third at Northampton, were living in poverty because the wages of the head of the household were so low that he could not support a family of three children or less" (1915, 41). As we can see low wages of the non-qualified workers is the principal cause of primary poverty. This result is valid for the three cities except Stanley where that the chief wage-earner is ill or old is the substantial cause.

Eighty years after the work of Bowley and Burnett-Hurst, the United Nations (1996, 47) during the 1995 World Summit for Social Development in Copenhagen, has defined absolute poverty as "a condition characterised by severe deprivation of basic human needs, including food, safe drinking water, sanitation facilities, health, shelter, education and information. It depends not only on income but also on access to social services." The importance of measuring absolute poverty is highlighted by the United Nations in the Programme of Action of the Copenhagen Summit ((United Nations, 1995) cited in Gordon (2000, 50)) as: "Elaborating at the national level, the measurements, criteria and indicators for determining the extent and distribution of absolute poverty. Each country should develop a precise definition and assessment of absolute poverty, preferably by 1996, the International Year for the Eradication of Poverty." At the nineteenth and early twentieth centuries, the focus was on the three main items, namely, shelter, food and earnings. With this new framework of the United Nations, the list of essential needs has expanded beyond monetary perspective to poverty. Gordon et al. (2000b, 87-88) have defined absolute poverty as follows: "being so poor that you are deprived of basic human needs. In order to avoid absolute poverty, you need enough money, and/or things you grow or receive to cover adequate diet, shelter/housing, warmth/heating, clothing, adequate sanitation/latrine facilities, access to basic health care, access to basic education/literacy programme." Gordon et al. have proposed the two following questions for better analyzing absolute poverty: "what does your family need to keep out of absolute poverty?", and "how far above or below that level of income is your family? 1. A lot above that level of income, 2. A little above, 3. About the same, 4. A little below, 5. A lot below that level of income." Moreover, the Organisation for Economic Co-operation and Development (OECD 1976, 69 cited in Spicker, Leguizamon and Gordon 2007) has defined absolute poverty as "a level of minimum need, below which people are regarded as poor, for the purpose of social and government concern, and which does not change over time."

In next sub-section, we will focus on basic needs approach, which has still a very important place as a concept in contemporary United Nations poverty studies. Basic needs approach (Streeten et al. 1981) is introduced by the United Nations in collaboration with Paul Streeten and his team.

Basic needs

In the literature, there is also another concept of poverty called *basic needs,*[142] (see Streeten et al. 1981; Stewart 1977, 1985; Streeten and Burki 1978; Hicks and Streeten 1979; Streeten 1984, 1981, 1979; Ghai, et al. 1977; ILO 1976; ODI 1978)[143], which is quite similar to Rowntree's minimum necessities concept (1901). For years, economists have focused on economic growth and have considered the gross national product (GNP onwards) per head as the best indicator of welfare and development. GNP is particularly known for facilitating international comparisons. However, GNP is a very poor indicator for reflecting multiple aspects of life (i.e., education, health, discrimination, inequality, social justice). In this perspective, Nussbaum (2008, 111) qualifies GNP as a crude measure, which "did not even take distribution into account, and thus was utterly useless in confronting nations with a lot of poverty and high rates of inequality. And it was actually worse than useless, because it gave high marks to nations that contained huge inequalities, encouraging people to think that such nations (for example, South Africa under apartheid) had done things right". Besides, in the early 1980s despite the economic growth-based development programs imposed to southern countries the Word Bank has estimated approximately 750 million people living below the absolute poverty line[144]- which corresponds to nearly one-third of the southern countries' population. The situation was a disaster for the majority of the world population. The statistical figures have showed that the orthodox development policies such as the economic oriented ones failed to promote human development and to reduce poverty in developing countries (see Streeten 1979, Singh 1979 for discussion on basic needs approach versus the New International Economic Order).

Thus, given the miserable situation of southern countries, some economists have proposed to use a social approach rather than a simple focus on GNP. Particularly in 1973, the World Bank Group president McNamara's famous Nairobi speech was a kind of declaration of war against poverty in southern countries. In his speech, he especially has emphasized the importance of the 'essential human needs' and of the voices of the poor. Eradicating absolute poverty and solving accessibility problems to basic needs such as nutrition, shelter, education, health was put in the agenda of development goals in developing countries. With this initiative, in 1976 during the Word Employment Conference organized by the International Labour Organization in Geneva, the recommendation of the basic needs approach was unanimously adopted by the participants to the conference. One year later in 1977 the report named 'Employment, Growth and Basic Needs' was s published by ILO. This report introduced the concept of basic needs for development purposes. Impressed by this work, McNamara set up a working group directed by Paul Streeten, for working on basic needs.

The basic needs framework from the ILO conference includes "certain minimum requirements of a family for private consumption: adequate food, shelter and clothing, as well as certain household furniture and equipment», «essential services provided by and for the community at large, such as safe drinking water, sanitation, public transport and health, education and cultural facilities", and "it should be placed within a context of a nation's overall economic and social development. In no circumstances should it be taken to mean merely the minimum necessary

[142] It is also known as *First things first.*

[143] The works for developing the basic needs approach were undertaken between 1978 and 1980. However, in 1978 the Word Bank has launched the initiative for defining the poverty reduction programs based on the basic needs approach in Southern countries.

[144] In the literature, some studies used the absolute poverty as the food criterion-based poverty line.

for subsistence; it should be placed within a context of national independence, the dignity of individuals and peoples and their freedom to chart their destiny without hindrance" (ILO 1976, 24-25). In line with McNamara's speech, the basic needs initiative aimed at encouraging people affected by poverty to participate in decision making process (see Ghai et al. 1977, 32).

Townsend (1993, 32) has described the basic needs approach as "an enlargement of the subsistence concept". He has suggested focusing on local community needs in addition to individual's and families' needs. Albeit, he has criticized that the basic needs approach focuses mainly on physical needs rather than including social needs (he also has criticized the concept of subsistence by saying that it "minimises the range and depth of human need", see Townsend 1993, 36). He has recommended sub-analysis for specific poor and disadvantaged groups such as ethnic minorities, women, children, elderly, disabled people. Townsend has mentioned that (1993, 32) "if there is poverty there has to be a social structure of poverty, with sections or groups in the population experiencing higher risk of that poverty, and some recognised to be more severely impoverished than others. This illuminate causes, but also priorities in policy. The links between the two have to be recognised, so that they can be exploited for both scientific and political or social reasons".

Before Streeten et al.'s (1981) basic needs framework, the United Nations Research Institute for Social Development (UNRISD) have worked on the development of social indicators such as composite indicators in the 1960s against the very popular economic indicator GNP. For instance, Drewnowski and Scott (1966) have developed the *Level of Living index*. This index takes in consideration two types of basic needs: physical needs (e.g., nutrition, shelter, health), and cultural needs (e.g., education, leisure, security). Moreover, the Overseas Development Council (ODC) with D.M. Morris have developed the *Physical Quality of Life Index* (PQLI), which corresponds to an index of basic needs (see Morris 1979). As indicators, the life expectancy at age one, infant mortality, and literacy are used for measuring the fulfilment of minimum human needs (Hicks and Streeten 1979, 575-576).

Mahbub ul Haq, one of the authors of the basic needs approach, has mentioned that the basic needs "is about an approach that enables the poor to earn or obtain their *basic needs*" englobing the dimensions of education, health, nutrition, housing, water supply, and sanitation (1981, vii). The main goal of the approach is to "provide the opportunities for the full physical, mental, and social development of the human personality and then derives the ways of achieving this objective" (Streeten et al 1981, 33). In that sense, people suffering from hunger, malnourishment, being ill, or lacking access to education, safe drinking water, sanitation facilities, or decent shelter are considered poor. This approach made way for special programmes for vulnerable, disadvantaged groups such as malnourished children, discriminated ethnic minority groups, children lacking access to school. Moreover Brady (2003, 720-721) has argued that "absolute measures of basic needs are most useful in developing countries vulnerable to famine and underdevelopment".

In practice, Streeten et al. (1981) have suggested 6 essential basic needs (i.e., health, education, food, water supply, sanitation, housing) and 8 indicators for measuring poverty based on the basic needs (see Table 7 below). Streeten (1979, 136) has also proposed some other basic needs such as "self-determination, self-reliance, political freedom and security, participation in decision making, national and cultural identity, and a sense of purpose in life and work".

Table 7. The list of basic needs and their core indicators, source: Hicks and Streeten (1979, 578)

Basic need	**Indicator**
Health	Life expectancy at birth
Education	Literacy
	Primary school enrolment (as a percentage of the population aged 5 to 14)
Food	Calorie supply per head or calorie supply as a percentage of requirements
Water supply	Infant mortality (per thousand births)[145]
	Percentage of population with access to potable water
Sanitation	Infant mortality (per thousand birth)
	Percentage of population with access to sanitation facilities
Housing	None

As we can see from Table 7, Streeten et al. (1981) have determined one indicator for the dimensions of *health*, *food*, two indicators for the dimensions of *education*, *water supply*, *sanitation*. However, for the dimension *housing* there is no indicator. The authors have mentioned that there is not any specific need of defining only one indicator per dimension; more than one indicator can measure progress in a dimension, or one indicator can reflect the progress for more than one dimension. Concerning the *housing* dimension, there was not any satisfactory measure of housing needs. Streeten et al. (1981) have added that the indicator *people per room* could be used, but the indicator was more about reflecting overcrowding than the quality of housing.

Concerning how the core indicators can be used technically, Streeten et al. (1981) have highlighted that they did not find an adequate weighting (i.e., relative importance of an indicator) design for the construction of a composite basic needs index (where the scores of all core indicators are aggregated together by taking in consideration the relative importance of each indicator for obtaining a single number of poverty). Moreover, the authors have pointed out that "…basic needs in fact refer to the inputs rather than to ultimate goals" (1981, 94). For instance, the authors have argued that the health indicator *life expectancy at birth* may already be a good indicator alone to reflect basic needs. According to authors, the basic needs indicators are multi-purpose in the sense that they: (i) facilitate gathering "adequate, standardized, comparable international statistics on basic needs" (Streeten et al 1981, 92); (ii) shape development policy analysis; (iii) direct the content and level of aids to countries; (iv) put in evidence the gap between developed and developing countries,;(v) evaluate performance developing countries for meeting the basic needs of their citizens; and (vi) monitor the relationship between the basic needs policies and other economic policies such as trade, investment. However, the quality of results and/or the quality of application process of the basic needs depend on the fulfilment of three following conditions: optimal amount of supply of essential goods and services in the market, optimal demand by consumers (i.e., the poor must have a sufficient level of purchasing power for buying essential goods and services in the market), optimal functioning of the institutions for meeting basic needs in the market. As we can see, the basic needs approach is framed from economic perspective, there is a very few

[145] As a result of discussions between the authors and the United Nations Research Institute for Social Development, the indicator 'infant mortality' was considered as the indicator of the availability of sanitation and clean water facilities (Streeten et al. 1981, 93).

sociological justifications. In addition, there is no distinction between people who did not use the basic items by their own choice and people who do not have access to basic needs due to financial deprivation, social exclusion, or any other reasons.

Streeten et al. (1981, 27) particularly wanted to make operational their concept for policy makers. The concept was used for poverty eradication projects in developing countries. They have provided four possible way of implementation of the process of meeting the basic needs: (i) apply the method called "count, cost, and deliver", which consists of identifying people who do not have the basic items (i.e., basic needs), of counting the number of people in deprivation situation, of computing the cost of providing the basic needs to the target group and of designing the process of delivering the target group the basic items; (ii) create new opportunities for increasing earnings of the poor such as how to make better the production of goods and services, the access to sources and markets; (iii) improve the functioning of central and local institutions for making enable the access to basic needs for everyone ; (iv) make enable the poor to manifest in the design of poverty eradication programs (i.e., help the poor to show their voices and their ideas, experiences for better national poverty eradication policies).

In practice, the basic needs approach was implemented in some developing countries (see Streeten et al. 1981, 168-173). For instance, India's new draft plan for 1978-83, India' revised minimum needs program[146], the Sixth Plan of Nepal (1980-85), or Indonesia's third plan (1979-84). India's new draft was including three main objectives: abolishment of unemployment and underemployment, rising the living standards of the poor, and providing essential needs and/or services such as "drinking water, literacy, elementary education, health care, rural roads, rural housing, and minimum services in urban slums" to citizens. India's minimum needs programme was designed for increasing "allocations for water supply, basic education, rural roads, and other essential needs". The sixth plan of Nepal had two objectives: focus on creating employment opportunities for eradicating absolute poverty, and fulfillment of basic needs, namely, "potable water, minimum health care, primary and skill-oriented education, family planning and maternity and child-health care services, irrigation facilities". The third plan of Indonesia focused on welfare justice with the aim of providing "access to means of fulfillment basic human needs, especially food, clothing and shelter" (Streeten et al. 1981, 170-171). Moreover, at international level, the basic needs approach has played an important role in the elaboration of the UNESCO (United Nations Educational, Scientific and Cultural Organization) 1978 'Study in Depth on the concept of Basic Human Needs in Relation to Various Ways of life and its possible Implications for the Organisations'[147], and of the Brandt 1980 Reports (i.e., The Brandt Report is written in 1980 by the Independent Commission on International Development, pioneered by Willy Brandt, the former German Chancellor. The aim was to put in evidence the differences between Northern and Southern countries in terms of economic development. The report suggested the transfer of some resources from developed (Northern) to developing (Southern) countries. The famous Brandt Line, proposed by W. Brandt, distinguishes Northern and Southern countries by a line in the world map, based on countries' GDP per capita scores)[148].

[146] Authors did not mention the date.
[147] Please see the link below for the document (accessed 14 April 2020):
https://unesdoc.unesco.org/ark:/48223/pf0000028550
[148] Please see the link below for the briefing paper of the Brandt commission in 1980 for the ODI

Moreover, in the late 1970s, the basic needs approach is adapted by the United Nation's Economic Commission for Latin America (ECLAC- Comisión Económica para América Latina y el Caribe in Spanish-CEPAL) to Latin America context under the name of '*unmet/unsatisfied basic needs*' (UBN hereafter or UBNS) to measure poverty based on population census (see Feres and Mancero 2001; Schuschny y Gallopin 2004; ECLAC 2009)[149]. This method is massively used in Latin America and governments adapted indicators and thresholds for country context Even if deprivation indicators may vary among countries; for instance, housing, overcrowding, services, education, and subsistence capacity are used as dimensions in Argentina, Colombia, Peru and Venezuela (Rio Group 2006, 107). The indicators of UBN are considered reflecting '*means*' instead of traditional '*ends'*. The UBN is recognized as a direct method to measure poverty since it focuses on whether the need is meet or not for the household. Generally, in each dimension one indicator is assessed. To distinguish if the household is deprived or not regarding the dimension, the absolute threshold is used. If the household is deprived regarding at least one item, then the household is considered deprived (see Santos et al. 2010). Instead of an aggregate/composite index, generally the results are provided in terms of deprivation per each item (i.e., dashboard) [150]. As distinct from other poverty studies at that time, the UBN reported distributions of different joint deprivations (experiencing concurrent deprivations).

Most recently, Santos et al. (2010) have proposed three methodological modifications to unmet basic needs approach to measure multidimensional poverty and tested it for six Latin American countries between 1992 and 2006 in a comparative perspective. The three methodological propositions are: (i) measuring poverty based on a hybrid approach including both income and UBN items; (ii) employing Alkire and Foster (2007, rev. 2009b, 2011a) counting methodology (computing two main indices_ headcount ratio and intensity of poverty): (iii) using alternative weighting system. Slightly different from the UBN, Santos et al. (2010) used the dimensions of 'children at school', 'education of household head', 'water', 'sanitation', 'shelter' and 'income' (each dimension has one indicator). However, the basic needs is perceived often as commodity bundles and it did not have a widespread and influential impact on development, social policies, improvement of public services for reducing poverty in southern countries. This result is due to the structural adjustment, economic growth, macroeconomic stabilization-based development programs and policies, which are suggested by Bretton Woods institutions in the 1980s to developing countries (see Streeten 2003 for more information on misinterpretations and

(Overseas Development Institute: https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/6638.pdf (accessed 25 June 2018). Please see the link below for Frances Stewart's article on the Brandt Report:
https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/10735/IDSB_12_2_10.1111-j.1759-5436.1981.mp12002010.x.pdf;jsessionid=B4DC8607AADBA9B14A64D861D7A5738E?sequence=1

[149] Certainly, census data is limited in terms of available indicators for measuring UBN. National surveys (micro-level data set) such as general household, socio-economic, living conditions survey bring more information due to more available questions related to quality of life. But census data was considered as the best way to measure poverty in multiple aspects in Latin America due to data availability and comparability issues.

[150] However, we note that the UBN index is provided in some studies. This index is based on counting approach and provides the proportion of people living in households experiencing one or more unmeet basic needs. But the main goal was to report deprivation in at least one basic need (i.e., the union approach). Moreover, each indicator is weighted equally, which causes disproportionally weighting of some dimensions when more than one indicator is considered in the dimension.

misconceptions of the basic needs and see Boltvinik (1992) for the combination of the UBN with income). As we have seen the basic needs approach has been promoted by social economists and international organizations to measure poverty in developing countries. However still to date, economic approach on the basis of income, expenditure, consumption patterns continue to have a significant impact on poverty measurement and policies. In the appendix 9.1, we will explore this financial aspect, commonly known as monetary poverty approach.

Relative deprivation

As we have seen previously in the basic needs approach, the focus is on a potential list of ultimate necessary items for living a decent life. In the literature there is an ongoing discussion about whether it is necessary to define some basic needs, or whether the basic items should be determined from an absolute perspective or rather from a relative perspective. From a relative perspective, the list of essential living items can be determined as a function of society's situation, which means that the list might change from one country to another or even from one society to another. Basic needs and especially the approaches of subsistence and absolute poverty are criticized because they mainly consider individuals as consumers of goods and services. However, individuals have also a social life participating in various social activities and taking different social roles as family member, parent, husband/wife, partner, friend, worker, citizen, neighbor (see Rein 1970; Townsend 1954, 1962, 1970, 1979; Abel-Smith and Townsend 1965; Lister 1990, 2004). In this perspective individuals do not only need subsistence items such as food, shelter, clothing, sundries, but they do have also social needs. British sociologist Peter Townsend (1993) has emphasized that social needs concern all people in the world, citizens of both rich industrialized and developing countries. In this direction, Townsend (1970, 1979, 1981, 1983) has promoted the relative deprivation approach[151] – notably, the concept of social need and social participation to conceptualize and measure poverty from multiple aspects. By using the term *deprivation*, he defines two main goals: (i) include societal (e.g., family, country) and local community perspectives in poverty conceptualization; (ii) combine indicators of resources (e.g., income, subsistence) and conditions (e.g., social, environmental conditions) for specific groups and society as a whole in poverty measurement process.

Townsend defines deprivation as "a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belongs" (1993, 79)[152]. Townsend has argued that deprivation emerges in different forms (i.e.,

[151] Before Townsend, Runciman has proposed the theory of relative deprivation in 1966; in his work entitled *Relative Deprivation and Social Justice* he has concerned mainly with inequalities and social justice in the twentieth century in England. However, Runciman clearly asserts that originally the term 'relative deprivation' is employed by Stouffer et al. (1949) in their work entitled *The America Soldier, Adjustment During Army Life*. Runciman assumes that inequalities exist in every society and he examines relative deprivation on the basis of subjective evaluation regarding two aspects: "a person's sense of relative deprivation will be affected not only by which of several membership reference groups is the basis for his chosen comparison; it will also be affected by what he feels about its relation to his comparative group" (1966, 31). He employs, therefore, two criteria of relative deprivation: *relatively deprived because of group's position in society* vs. *relatively deprived because of own position as member of group*. Notably, the inclusion of the terms '*reference group'* and *'membership group (comparative group)'* forms the fundamental point Runciman's relative deprivation study.

[152]Townsend considers a person *deprived* if he or she "lacks the types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities

multiple deprivation vs. a single form of deprivation), and in combination with prosperity. The latter point emphasizes variety and complexity of deprivation patterns. An individual can experience some deprivations regarding some aspects of life but also some prosperous situations regarding other aspects of life (e.g., being prosperous regarding housing conditions but deprived regarding working conditions). To measure poverty as multiple deprivations, Townsend has proposed to employ new social indicators in addition to income poverty. Townsend has pointed out that "the shift is not simply to a much larger set of indicators of demonstrable or objective material and social deprivation and their links with income. It is also to the changing relationship of deprivation and income through time and across communities occupying different territories" (1993, 34). Under the relative deprivation framework, Townsend, therefore, has emphasized that the list of poverty items should be updated as a function of years and social conditions of countries. For instance, he has mentioned that "people living in the present may not be subject to the laws and obligations as well as the customs which applied to a previous era. In the United States, the UK, Germany, Eastern Europe and the 15 republics created after the collapse of the Soviet Union, or, to take different examples, in Ethiopia, Indonesia, Bangladesh, Iran, Cuba and Chine, there have been sharp changes in living standards, customs and patterns of association within the single decade of the 1980s" (1993, 34).

In 1979 Townsend published the results of his study *Poverty in the United Kingdom* (In 1968-9 he carried out fieldwork during which he interviewed 3,260 households and 10, 048 individuals). He has considered the extent of poverty as a multidimensional issue and as a national phenomenon in Britain. In this work, Townsend (1979, 915) has defined *being deprived* as "the absence or inadequacy of those diets, amenities, standards, services, and activities which are common or customary in society. People are deprived of the conditions of life which ordinarily define membership of society". Herewith he has identified the poor as people who "lack or are denied resources to obtain access to these conditions of life, and so fulfil membership of society" (1979, 915)[153]. In this regard, his general theory is that "individuals and families whose resources, over time, fall seriously short of the resources commanded by the average individual or family in the community in which they live, whether that community is a local, national or international one, are in poverty" (Townsend 1962, 225). Through these two definitions, we can see that Townsend clearly distinguishes the two notions 'deprivation' and 'poverty'.

For Townsend (1979, 249), social life involves different spheres such as "at work or school, in the home, in the immediate vicinity of the home, and elsewhere in the community, and in all of these spheres the individual's diet, health, welfare, occupation and recreation are defined.." He

---

and facilities which are customary, or at least widely encouraged and approved, in the societies to which they belong" (1979, 413).

[153] Or, in another work (Townsend 1993, 36) he has defined being *relatively deprived* as "…cannot obtain, at all or sufficiently, the conditions of life-that is, the diets, amenities, standards and services-which allow them to play the roles, participate in the relationships and follow the customary behavior which is expected of them by virtue of their membership of society. If they lack or are denied resources to obtain access to these conditions of life and so fulfil membership of society, they may be said to be in poverty." In another work (Townsend 1962, 210) called *The Meaning of Poverty*, Townsend has emphasized that poverty and subsistence are relative, dynamic concepts and that "they can only be defined in relation to the material and emotional resources available at a particular tome to the members either of a particular society or different societies".

also emphasised that the types of need change from one society to another and that the list of needs should not be fixed once and for all. According to him, one country's poverty definition will not necessarily match with those of another country. For instance, he mentions that "the amount and cost of the food which is eaten depends on the social roles people play and the dietary customs observed as well as the kinds of foods made available socially through production and availability in the markets" (Townsend 1993, 31).

Similar to many anthropologists, Townsend has adopted a relativist perspective and did not want to endorse one particular list of absolute or basic needs for measuring poverty in developing countries. Townsend's relativist perspective was partially developed by Smith in 1812, by Marx in 1847 (Marx has expressed that "our needs and enjoyments spring from society; we measure them, therefore, by society and not by the objects of their satisfaction. Because they are of a social nature, they are of a relative nature"- see Marx 1847, Sec. IV, 37), and by Booth in the 1880s (Booth 1889, 1891, 1892-7). Smith[154] gave the example of the linen shirt. He has argued that for the Greeks and Romans the linen shirt was not anymore a necessity as before: "A linen shirt, for example is, strictly speaking, not a necessary of life. The Greeks and Romans lived, I suppose, very comfortably though they had no linen. Not in the present times, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without linen shirt, the want of which would be supposed to denote that disgraceful degree of poverty which, it is presumed, nobody can well fall into without extreme bad conduct. Custom, in the same manner, had rendered leather shoes a necessary of life in England. The poorest creditable person of either sex would be ashamed to appear in public without them" (see Smith 1776, Vol. 2., Book V., ch.2). In the same vein, Marx has stated that types of needs and enjoyments depend on society and Booth has defined poverty as the inability to have sufficient resources to live a decent life according to the British society's standards. Townsend has argued that a standardisation of (the criterion of) poverty won't be an objective process for measuring poverty due to variations in social needs and in living conditions among societies but also within society. For instance, he has pointed out that "the necessities of life are not fixed. They are continuously being adapted and augmented as changes take place in society. Increasing stratification and the division of labour, as well as the growth of powerful new organizations, create, as reconstitute, 'need'. In particular, the rich set fashions of consumption which gradually become diffused." (1979, 915).

Townsend (1979, 915) has mentioned the limit of subsistence-based poverty approach in dealing with contemporary needs of societies. Related to this, he wrote that people "are encouraged to believe that 'subsistence' represents the limit of basic human needs, and this tends to restrict their assessment of what individual rights or entitlements could be introduced and guaranteed. A limited definition of need leads to a limited appreciation of rights. On the other hand, needs other than those included in the conception of 'subsistence' are denied full acknowledgment. These are goods, amenities and services which men and women are impelled to seek and do seek, and which by the tests of both subjective choice and behaviour are therefore social necessities, that have traditionally been excluded from consideration in devising poverty standards. People do not live by bread alone, and sometimes they are prepared to forego bread to meet a more pressing social need." Townsend has also criticized that subsistence or 'a physically efficient diet' had used for releasing the income maintenance system in Britain.

[154] We don't enter in detail about Smith's perspective on poverty through his works. But Gilbert (1997) has provided a comprehensive review of Smith's ideas on the nature and causes of poverty.

However, in his opinion, subsistence should be taken as a social issue. He has pointed out that "this preoccupation with nutritional deprivation as the centrally evident problem of meeting need in society has, first, to be extended logically to *dietary* deprivation, thereby putting stress on the kind of food and drink which people actually consume (and the distribution of the budgets from which they purchase it), as well as the amount and quality of nutrients which they absorb, so acknowledging the social definition of dietary need." (1979, 915).

To capture better the real needs of people in Britain, Townsend created an exhaustive questionnaire composed of thirty-nine pages and nine sections on housing and living facilities, employment, occupational facilities and fringe benefits, cash income, assets and savings, health and disability, social services, private income in kind and style of living. Townsend has developed two objective measures of poverty: *relative income standard*, and *relative deprivation standard*. However, one of the most interesting points in his work is that during the interviews people were also asked about the perception of poverty (i.e., how poverty can be defined). More specifically, the chief wage-earner or head of the household was asked the following question: "There has been a lot of talk about poverty. Do you think there's such a thing as real poverty these days?" (1979, 237)[155]. Among participants, sixty-two per cent agreed on the existence of real poverty in Britain. For instance, a 40-year-old manual worker with three children described poverty as "not having any money and not being able to earn any. Hearing the babies cry because they are hungry or cold-that's poverty." (1979, 236).

As for the monetary dimension, Townsend has not only considered earned income, but he has also collected information on unearned incomes, capital assets, benefits in kind, public services in kind, and private income in kind. On the other hand, concerning social perspective, he has measured *relative deprivation* given community or society's style of living. To measure relative deprivation, he has used sixty indicators for representing the 'style of living' including "the major areas of personal, household and social life" of the British people (1979, 251). These indicators reflect different forms *of deprivation* or *multiple deprivation*, such as "diet, clothing, fuel and light, home amenities, housing and housing facilities, the immediate environment of the home, the characteristics, security, general conditions and welfare benefits at work, family support, recreation, education, health, and social relations" (1979, 249, and see Appendix thirteen on page 1173 in Townsend 1979). These indicators are selected to indicate deprivation in terms of "lacking that amenity of not participating in that activity" (1979, 251). For illustrative purposes, in his book, Townsend created a *summary* 'deprivation index' (composed of "major aspects of dietary, household, familial, recreational and social deprivation[156]") on the basis of twelve items (1979, 250): "1) Has not had a week's holiday away from home in last 12 months, 2) *Adults only*. Has not had a relative or friend to the home for a meal or snack in the last 4 weeks, 3) *Adults only*. Has not been out in the last 4 weeks to a relative or friend for a meal or snack, 4) *Children only* (under 15). Has not a friend to play or to tea in the last 4 weeks, 5) *Children only*, Did not have party on last birthday, 6) Has not had an afternoon or evening out for entertainment in the last two weeks, 7) Does not have fresh meat (including meals out)

[155] It is also asked subjective description of poverty.

[156] On social deprivation issue, Holman (1978) gives excellent explanations. He mentions that several findings in the literature show that often a person is not being deprived in only one aspect but simultaneously in several aspects. For instance, Holman yields some cases such as living in poor housing conditions and being ill; or, having low income and low educational attainment; or, living in overcrowded household and not being successful at school, etc. Thus, he has emphasized the importance of examining 'socially deprived areas' and different categories of poor.

as many as four days a week, 8) Has gone through one or more days in the past fortnight without a cooked meal, 9) Has not had a cooked breakfast most days of the week, 10) Household does not have a refrigerator, 11) Households does not usually have a Sunday joint (3 in 4 times), 12) Household does not have sole use of four amenities indoors (flush WC; sink or washbasin and cold-water tap; fixed bath or shower, and gas or electric cooker)." Given that the maximum total deprivation score is 12 and the minimum deprivation score is 0 for an individual, Townsend found 3.5 as the average score of deprivation for the entire sample, 3.4 as the average score of deprivation for children, 3.5 for adults, 3.6 for housewives, 3.0 for young adults (15-44), 3.5 for older adults (50-59), 4.1 for people aged 60-69, and 5.1 for people aged 70 and over. The higher the total deprivation score the lower the participation (1979, 251). Moreover, Townsend did computations by a range of income (i.e., net household income last year in pounds) and by household size (i.e., for households composed of more than one individual, the average of scores of household members gives the household total deprivation score).

In summary, Townsend has provided three measures of poverty: (i) the *supplementary benefit standard*, which is the official measure of the government; (ii) the *relative income standard*; and (iii) the *deprivation standard*. On the one hand, the *relative income standard* is defined in terms of types of household, in terms of 50 per cent (very low) and 80 per cent (low) of the mean income for each type of household (see 1979, 248). This measure helps to detect the population with the lowest income. On the other hand, the *deprivation standard* is computed based on the deprivation index (see previous paragraphs for the deprivation index), and it corresponds to "a standard of income below which people experience deprivation disproportionally to income" (1979, 270). The thresholds Townsend used for the three measures of poverty were higher than the threshold used by the government for measuring poverty- especially for households with children and for households with disabled people. He has also examined correlations between different forms of deprivation (i.e., dietary, clothing, fuel and light, household facilities, housing conditions and amenities, working conditions, health, education, environment, family, recreation, social relationship)[157] and income.

Townsend's work includes various estimations as regards to numerous aspects of living conditions, to the number of poor people, to the people living on the margins of poverty, and to the characteristics of people living in poverty. Across these investigations, the goal of Townsend is to explain comprehensively the poverty in Britain. In his book, in addition to the sections dealing with the incidence of poverty, long-term vs. short-term poverty, regional poverty, the impact of poverty , there are fifteen sections focusing on the analysis of specific groups or subjects such as social class, styles of living, objective and subjective deprivation, deprivation at work, deprivation in housing, deprivation regarding environment, the problems of poor areas, social minorities, the unemployed , the low paid worker, disabled people, the long-term sick people, handicapped children, one-parent families, elderly persons. As we can see, Townsend's investigation is very detailed in terms of sub-analysis compared to previous empirical poverty studies carried out in Britain by Booth, Rowntree, and Bowley and Burnett-Hurst.

The results of Townsend's investigation show that the relative deprivation approach captures more people in poverty compared to Britain's national standard of poverty (e.g., British government's supplementary benefit standard commission's scale rates ignore some vulnerable

[157] In total there are sixty indicators, see Townsend (1979, 1173-1176).

people). The government's computations have showed that 6.1 per cent of population are in poverty and 21.8 are on margins of poverty. Regarding the relative income standard 9.2 per cent of population are in poverty and 29.6 are on margins of poverty. Regarding the deprivation standard, 22.9 per cent of population are living in poverty. Amongst other findings, some of them are as follows: (i) in terms of household composition, one-person households and households with at least five members are most likely to be in poverty compared to other households; (ii) women compared to men, self-employees compared to employees, dependants compared to employed people are likely to be in poverty; (iii) children and elderly people are significantly vulnerable groups and more than half of the retired people live in poverty or live on the margins of poverty. Townsend has also found that poverty is correlated with social class; for instance, the incidence of poverty reached the highest level in the households of unskilled manual workers living with dependants.

Townsend has showed dynamics of poverty explaining that society continuously generates new obligations to its members and therefore society produces new needs. Households, families, individuals, groups' transitions such as move in and move out of poverty can be observed over a certain period. For instance, he mentioned that "at any particular time there are households, who because of demotion, unemployment, sickness, disablement, retirement or increase in dependency, have recently fallen into poverty, just as there are households who, because of promotion, engagement or re-engagement at work, recovery from sickness or decrease in dependency, have just emerged from poverty" (1979, 916). Similar to Booth, he has pointed out the link between life cycle and poverty: "during the early stages of the life-cycle, incomes and other resources are low; children put claims on the resources of the young adults who are their parents; the costs of housing and establishing a home are considerable; and the parents have jobs which are relatively insecure and paid below average. During middle or late middle life, a peak of prosperity is reached; the numbers of dependants and housing costs diminish just when resources actually increase or are at least maintained; earnings for a standard number of hours often increase because of promotion or seniority, and assets are accumulated or invested. In the later stages of life, there is a descent into austerity or poverty" (1979, 916).

In the 1980s, the Breadline Britain Survey[158] has been carried out according to Mack and Lansley's social indicator methodology (Bradshaw 2000, lii[159]) in order to investigate the extent of poverty and to update Townsend's work (because at that time since Townsend's survey, a comprehensive study of poverty and living conditions among British people has not been done yet). Mack and Lansley (1985, 17) have criticized absolute poverty line approach (e.g., Rowntree's approach, which mainly takes in consideration food, clothes, and shelter) because it does not take into account "the maintenance of mental, moral or social sides of human nature". As the alternative, they have proposed a new approach called *consensual*[160]. The authors do not

[158] The company Market and Opinion Research International (MORI) designed and conducted this survey in February 1983. 1,174 people were interviewed (see Appendix A in Mack and Lansley 1984, 287-290 for survey sampling details). The latest version of Breadline Britain is published in 2005, written by Lansley and Mack.

[159] This information is from the preface of Jonathan Bradshaw in the centennial edition of Seebohm Rowntree's *Poverty: A study of town life* book by the Policy Press and the Joseph Rowntree Charitable Trust in 2000 (xix-lxxxii).

[160] The authors have mentioned that their methodology suits with Sen's idea on objectivity in measuring poverty (1985, 46). Sen emphasized the importance of "an objective diagnosis of conditions" based on "an objective understanding of feelings" (cited in Mack and Lansley 1985, 46; Sen 1981, 16)

define needs, the poverty line is based on their personal value judgment, but they consider the "consensual judgment of society at large about people's needs" -(1985, 46). Mack and Lansley have focused on social perception of necessities. Because for them, the life necessities can only be defined by public's opinion, and not by an expert. The authors have also argued that the indicators of starvation, hunger, malnutrition, life expectancy are certainly important for many developing countries, but they do correspond more to the "concept of good health." (1985, 21). Rather, Mack and Lansley have proposed to consider a large variety of deprivation items based on social perception of needs.

Mack and Lansley (1985) have given insight into the link between health and poverty. In general, health indicators show that the poor are likely to have serious health problems, but on the other side the ill-health may be itself a cause of poverty. Life expectancy or mortality rate as one the main indicators of health do not really help distinguishing the poor and the non-poor because the middle- income class people situated above the poverty line are also likely to have health problems as the poor for the similar reasons and conditions. For that reason, Mack and Lansley (1985) do not suggest using only the health indicators for setting a poverty line. They have also criticized Rowntree's (1901) dietary criterion for drawing up a poverty line based on '*health efficiency'*[161]. They have argued that the criterion of absolute poverty is based on the judgements on living standards of the eighteenth century and that the new standards should be adopted for the nineteenth century as a function of new needs, expectations, norms of societies. Life expectancy and health are priorities for some people, but not for everyone. Sometimes, poor people prefer spending their money to goods and services that are not the surviving items, but people want to get them to feel socially included just because of social customs [162]. For that reason, Mack and Lansley have advocated that poverty lines should be rather specific to the situation of each country, or society (called as '*society-approved minimum'*-see 1985, 46) and even to each generation. They have also pointed out that it will be more convenient if the poor can auto evaluate their situation compared to the living conditions of others.

In this light, the goal of the Breadline survey was multiple as follows: find out how the poor face poverty (i.e., including inequality in terms of unfair distribution of resources in society[163]); what British people really need in order to live decently; what are the actual living conditions and public opinions about State benefits in Britain. Mack and Lansley (1985, 9) have mentioned that "in particular, the study attempts to measure the extent of poverty not in terms of some arbitrary income level but in terms of the extent to which the poor are excluded from the way of living that is expected and customary in society today. This requires a survey not just of the poor but of the rest of society as well." Poverty is, therefore, defined as "*an enforced lack of socially perceived necessities*" (Mack and Lansley 1985, 9) and it is operationalised through

[161] Nevertheless, Rowntree used the items of radio, books, newspapers, beer, tobacco, presents, and holidays in his 1936 poverty study (see Mack and Lansley 1985, 27).

[162] For instance, Mack and Lansley (1985, 25) give an example: "Tricia is a single parent with two school-aged children. At Christmas, she gave her son a bicycle worth around 30 pounds. This was his only present of the year and to afford this Tricia had saved all year, putting aside a small sum each week. To do this, she cut back on food for herself: she usually misses breakfast; at lunchtime she just has a cup of coffee; in the evenings, she has a small meal with the children, something like eggs on toast or beans on toast. Recently, her health has been bad."

[163] Mack and Lansley (1985, 40) do not explore the distribution of resources in society, but rather they focus on entitlements.

35 items representing the standard (minimum) decent life in Britain[164]. Hence, "individuals sharing the social perception will feel deprived when they lack the items defined by society generally as 'necessities'" (1985, 39). The areas included in the survey are food, heating, clothing, consumer durables, entertainment, leisure activities, holidays, social occasions and activities (1985, 44). However, we note that the authors have taken in consideration the "aspects of life facilitated by access to money" (1985, 44). In addition, Mack and Lansley have qualified their work as "a new basis from which to evaluate the effectiveness of the welfare state" (1985, 9)[165].

There are two important contributions in this research.[166] The first contribution is that public opinion/opinion polls have defined the list of necessities. The authors wanted to discover the standard of living in Britain in the 1980s, via asking directly to people which items they consider 'necessities' for pursuing a decent life (see 1985, 43-44). In other words, for the first time, the minimum standard of living is found by using the consensual approach. In previous poverty studies, the focus was on the creation of one particular list of minimum standards such as minimum nutritional diet set by experts and on household's expenditure patterns. Once the list was determined, its costs was calculated as 'minimum costs' that was converted into 'minimum income' level called the poverty line. After, the 'minimum income' (i.e., the poverty line) was compared with the income level of households (i.e., households having income below the poverty line are considered poor). But, in the new approach of Mack and Lansley, the "minimum acceptable way of life" is determined as regards to the "views of society as a whole[167]" (1985, 42). Regarding Mack and Lansley (1985, 42): experts cannot be expert in all domains; the expenditure patterns cannot reflect directly needs but they can only give insight into the financial situation of households; the standards of living may vary between people having the same income level. The second contribution concerns the question of 'whether an item is missing because of personal choice or because of another reason'. On the basis of answers to this question, only the items that could not be afforded are included in the list of missing items (i.e., only '*enforced lack of necessities*' are taken in consideration for the computation of the number of poor). If more than 50% of the population consider an item important, then the item is qualified as a '*socially perceived necessity*'.

The Breadline Britain methodology has covered three steps (see Gordon 2000, 63): first, identify social necessities based on public opinion; second, identify individuals deprived of

[164] Mack and Lansley (1985) give insight into the goal of Breadline Britain as: "this study tackles the question how poor is too poor? By identifying the minimum acceptable way of life for Britain in the 1980s. Those who have no choice but to fall below this minimum level can be said to be in poverty. This concept is developed in term of those who have an enforced lack of socially perceived necessities. This means that the necessities of life are identified by public opinion and not by the views of experts or, on the other hand, the norms of behaviour per se."

[165] For instance, Mack and Lansley give an example of the inefficiency of supplementary benefit (1985, 171): "It's a well-known fact that people on supplementary benefit get about one pound per day per child. That's to clothe and feed them on. And it's ridiculous because nobody can clothe and feed a kid on one pound a day. [A single parent with two school-aged children]."

[166] Mack and Lansley (1985, 11) have mentioned that at that time in Britain, the official national Family Expenditure Survey and the General Household Survey were not yielding necessary information on the needs of British people.

[167] The authors have mentioned that the public opinion-based process was previously used by Rowntree to identify clothing needs, and by the European Union Commission in 1981 to identify a minimum income level necessary for pursing a minimum acceptable life.

socially perceived necessities; third, identify household income levels of the individuals at greater risk of deprivation regarding social perceived necessities. Each interviewee is asked which items in the list seem them necessary for leading a decent life in Britain (i.e., there are four response options: have, do not have but don't want, do not have and can't afford, does not apply). The methodology, therefore, defines (i) among the items perceived as necessities, which of them they have; (ii) among the items perceived as necessities, which of them they would like to have but they cannot afford. The deprivation index is constructed based on the items the majority of population perceives as social necessities, but they cannot afford because of lack of income. Mack and Lansley's work later was applied by Later, in 1990, the MORI- Market and Opinion Research International has applied Mack and Lansley's methodology in Britain (see Gordon and Pantazis 1997), and the Beaufort Research has employed in Wales in 1995 (see Gordon 1995).

The thirty-five items included in Mack and Lansley's study are ranked as a function of the proportion of respondents selecting the item as a necessity. "The rankings show that there is a considerable degree of social consensus. Over nine in ten people are agreed about the importance of the following basic living conditions in the home: heating (97 per cent), a not shared indoor toilet (96 per cent), a damp-free home (96 per cent), a not shared bath (94 per cent), and beds for everyone (94 per cent)" (1985, 53). On the other hand, there are nine items classified as necessities by less than 50 per cent of the respondents : a best outfit for special occasions (48 per cent), a telephone (43 per cent), an outing for children once a week (40 per cent), a dressing gown (38 per cent), children's friends round for tea/a snack once a fortnight (37 per cent), a night out once a fortnight (adults) (36 per cent), friends/family round for a meal once a month (32 per cent), a car (22 per cent), and a packet of cigarettes every other day (14 per cent). Authors have showed the importance of having opinion of people about basic necessities (1985, 55-56): for instance, in 1899 when Rowntree carried out his first survey, he characterized a poor family as "they must never spend a penny on railway fare or omnibus." But Mack and Lansley's survey results have showed that nearly nine in ten of British people do care of spending money on railway fare or omnibus as a necessity for living in the 1980s of Britain. Another example given by the authors is during the Victorian era a refrigerator or a washing machine were unknown, but in the 1980s these items have become social necessities. The authors have noted that new additional items can become necessary in other time because. But also, regarding van den Bosch (2001, 54), some items become necessities because of "their symbolic significance, or because of prevailing social norms and standards."

Once the ranking of items as a function of social necessities is found, a large variety of tables is provided by Mack and Lansley (1985) in their book. To mention some of them, the tables include (i) the distribution of items in rank order by social class (i.e., professional and managerial occupations, other non-manual workers, skilled manual workers, semi-skilled and unskilled manual workers, and social security recipients), age groups, household type (i.e., pensioners, single parents, families with children, households without children, and single people), political views, percentage of population classing the item as necessity; (ii) the distribution of items in rank order by possession of items (i.e., have/could not do without, have/could do without, don't have/don't want, don't have/can't afford), percentage of population classing the item as a necessity; (iii) the distribution of the 'necessities' items in rank order by net equivalent household income (i.e., poorest, middle, and top), percentage of population not having the item by choice; (iv) the distribution of lacked adult necessities by income characteristics such as households by income deciles, households in bottom 40% of

income range, in terms of percentage of not having the items by choice; (v) the distribution of the 'necessities' items in rank order by net equivalent household income, households on supplementary benefit, in terms of percentage of people not having the item because they can't afford it; (vi) the distribution of satisfaction with living standards (i.e., satisfied, neither satisfied nor dissatisfied, and dissatisfied) by lack of necessities (see Mack and Lansley 1985 for full tables).

Besides Britain studies, the consensual approach developed by Mack and Lansley (1985) was also used in other regions, countries such as[168] in Europe by Eurostat (Ramprakash 1994, Vogel 1997), in Sweden, Finland by Halleröd (1994, 1995a, 1995b, 1998), in Denmark by Mack and Lansley (1985), in Ireland by Callan et al. (1989), and Nolan and Whelan (1996), in Belgium by Deleeck et al. (1991), and van den Bosch (1998), in the Netherlands by Muffels et al. (1990), in Europe with a comparison between Britain, Denmark, Ireland, Belgium by van den Bosch (2002), and van den Bosch et al. (1993), most recently across Europe by Eurostat based on material deprivation index (2012), and in Benin by Nandy and Pomati (2015) in measuring multidimensional poverty.

The council of Europe (see European Union Commission 1981; cited in Gordon 2006, 30) has defined poverty, in 1975, as "individuals or families whose resources are so small as to exclude them from the minimum acceptable way of life of the Member State in which they live", where resource includes "goods, cash income, plus services from public and private resources". Later, in 1984, the European Union Commission (see EEC 1985; cited in Gordon 2006, 30) has used an extended definition of poverty as "…the poor shall be taken to mean persons, families and groups of persons whose resources (material, cultural and social) are so limited as to exclude them from the minimum acceptable way of life in the Member State in which they live." It is important to note that European Union's poverty definitions include social exclusion aspect as well (see the sub-section of social exclusion). In European Union's poverty definitions, the relative deprivation approach connects essential needs (for leading a minimum acceptable way of life) to social exclusion. Similarly, Atkinson et al. (2002) have argued that poverty is a relative concept that should include both the perception of individuals on socially minimum acceptable lifestyle and the socio-economic level of society. They have mentioned that "it is senseless to search for an absolute standard whereby the content of poverty can be determined accurately for all countries and for any moment in time" (2002, 78). They have also added that poverty is a graduated concept and the distinction needs to be made between 'subsistence insecurity' and 'far-reaching insufficiency of resources' (2002, 79): regarding first group, people have difficulties in affording (financially) some needs, but they are not in a position of being socially excluded from the mainstream; whilst the second group includes people who are definitely excluded from the society such as the homeless.

In practice, European Statistical Office (Eurostat) and the Organisation for Economic Co-operation and Development (OECD) have developed a monetary indicator (i.e., relative standard of income) for measuring relative deprivation and comparing it across countries. Regarding this measure, households having an annual income level below sixty per cent of median income are considered poor. This method is closely related to income inequality, as it refers to the economic welfare of an individual in relation to that of others in the same society[169].

[168] These studies are cited in Gordon (2000) and Gordon et al. (2000b, 93).
[169] The relative threshold can be defined according to median income or income distribution quintiles.

However, this monetary based poverty measure can be considered problematic. Recent poverty studies argue that poverty is beyond lack of sufficient money. For instance, Gordon et al. (2000b, 93) have mentioned that the poverty threshold and the modified OECD scale used by Eurostat are arbitrary (i.e., modified OECD scale allows to adjust household income regarding household composition: 1 for the first adult of the household, 0.5 for the other adults and 0.3 for children aged under 14) because these measures do not capture participation in "economic, social, cultural and political life" in society. Gordon et al. suggest that Eurostat should compute low-income level, for instance by using Budget Standards approach pioneered by Rowntree in 1901 (see Oldfield and Yu 1993) or Family Expenditure Survey (FES) or other similar surveys such as European Statistics on Income and Living Standards (SILC panel data), national Household Budget surveys. Herewith, monetary aspect can be seen as a dimension/component of (multidimensional) poverty. For instance, Townsend et al. (2000) have proposed a combination of income and living standards for measuring poverty. Regarding them, "standard of living includes both the material and social conditions in which people live and their participation in the economic, social, cultural and political life of the country" (Gordon et al. 2000, 91; Gordon and Pantazis 1997).

In addition to relative income approach, another practice of measuring poverty based on deprivation concept is the *material deprivation,* which is often considered as a complementary tool to monetary approach (see Boarini and d'Ercole 2006 for material deprivation measurement in OECD countries, Whelan and Maître 2012 for European comparison of material deprivation). Regarding this aspect, material deprivation indices are the most common non-monetary poverty measures especially in Europe (see for instance, EU-2020 poverty and social exclusion measure, The English Indices of Deprivation (Smith et al. 2015))[170]. As one of the prominent pioneers in this domain, in 2012 Eurostat has proposed the index of non-monetary poverty (i.e., *pauvreté d'existence*) to measure material deprivation across Europe (cited in Alkire et al., 2015, 132). This index consists of identifying who is deprived given the set of material items defined by Eurostat. Currently, there are 9 items: i) coping with unexpected expenses: ii) having annual holiday away from home, iii) avoiding arrears (mortgage or rent, utility bills or purchase instalments): iv) having a meal with meat, chicken, fish or vegetarian equivalent every second day: v) keeping home adequately warm: vi) having a washing machine; vii) having a colour TV: viii) having a telephone: ix) having a personal car (see Guio et al. 2009, Maître et al. 2014). If the household is deprived regarding at least 3 of 9 items, then individuals living in the household are considered materially deprived. But, if the household is deprived regarding at least 4 items, then individuals living in the household are considered severely materially deprived. Concerning English multiple deprivation index (Smith et al. 2015), it includes 37 indicators and 7 dimensions, which are: income, employment, health and disability, education, skills, and training, barriers to housing and services, living environment, and crime. Niemietz (2010) has criticized the use of median income as a relative poverty measure. He has argued that this method does not reflect properly the social context. Instead, he has proposed to combine budget standard approach with consensual material deprivation indicator to measure poverty from relativist perspective. There are also numerous academic papers that have adapted Townsend's relative deprivation approach such as Luzzi et al. (2008), Halleröd (1995a, 1995b), Suter and Paris (2002), Suter and Iglesias (2005).

[170] For overseas examples, see Saunders et al. (2008), Jensen et al. (2002), Fergusson et al. (2001).

Sen's view on absolute versus relative approaches

Townsend's marvellous work on poverty in the United Kingdom has significantly contributed to poverty studies in the twentieth century. However, his work has also received some critics. Piachaud (1981) has drawn attention on four issues. *First*, he has argued that patterns of living are diverse and tastes for living are relative. Townsend has stated that life is about the mix of deprivations and non-deprivations. Piachaud has found ambiguous Townsend's attempt to define a threshold for distinguishing the poor and the rest of society given that consumer's preferences cannot be evaluated objectively and that the living patterns are composed of deprivations in some aspects and of non-deprivations in other aspects. Similar to Piachaud (1981, 1987), Desai (1986) has also criticised Townsend's deprivation indicators saying that Townsend cannot decide him-self as an expert what people need their life to pursue a non-poor life. For instance, vegetarians won't need a joint of meat on Sunday. *Second*, Piachaud (1981) has pointed out that some social aspects are not included in Townsend's deprivation index, for instance, working conditions, environmental issues, and accessibility problems to public services. *Third*, Piachaud (1981) has argued that in relative perspective even if all income inequalities are eliminated in society, still no one can be identified as deprived given the diversity of people's behaviours and preferences. For instance, he has mentioned that "it is no indicator of deprivation if someone chooses to stay at home, eating salads and uncooked breakfasts. But all these personal choices are 'extraordinary', and so add to the score on Townsend's deprivation index" (cited in Gordon 2010, 232). *Fourth*, Piachaud (1981) has emphasized that the term of 'ordinary membership of society', which is used by Townsend to explain that poverty is not having necessary conditions to fulfil membership in society (e.g., being not able to practice society's common life style in terms of customs, activities, diets), is not coherent with diversity of life.

Piachaud (1981) has queried how can be possible to define objectively the features of 'ordinary membership of society' considering that life choices, living styles are relative and no one is obligated to follow in same manner the life, to accept the same values, customs, or to enjoy the same activities, or to employ the same diets. In other words, if someone does not prefer performing 'ordinary membership of society', it does not mean that this person is an outcast. In line with these critics, Piachaud has suggested focusing on "the choice a person has, and the constraints he or she faces" (cited in Gordon 2010, 232). He has argued that the preferences of persons can be interesting for a sociologist, but it is not directly related to poverty. Poverty, rather, is related to living conditions of persons; for instance, if the person is able to achieve what he or she would like to do[171]. In this sense, Piachaud (1981) does not find logic to determine whether a person is deprived or not regarding an aspect of life; for him, it makes more sense to explore the living conditions comprising constraints, opportunities in the life of

[171] The behaviors of individuals regarding their opportunities, constraints, sources can also be analyzed. But this ex-post approach to poverty risks to lead to behavioristic explanations of poverty rather than structural explanations when it comes to find out the causes of poverty. Several studies in the U.S. literature have showed that the poor are massively labelled as who do not know how to behave correctly to achieve success in life such as they do not know how to take and benefit opportunities (see underclass, culture of poverty in the appendix 9.1). Thus, it becomes important to explore in a comprehensive way the living conditions of people to reveal different forms of poverty. Piachaud (1981) has mentioned that "the study of styles of living is essentially about outcomes-how people choose to behave given the choices open to them. As Townsend found, it reveals a wide diversity of behavior. But what is of much more importance is the level of resources a person has, and the opportunities this affords" (cited in Gordon 2010, 232).

the person. He has summarized that "the reason for tackling poverty is not to create uniformity, but to push back the constraints and increase choice and freedom" (cited in Gordon 2010, 232). In short, he has stated that poverty cannot be analysed *objectively* based on relative deprivation concept[172].

In line with Townsend, some researchers (see Gordon and Spicker 1999) have opposed absolute poverty criteria, which are based on external experts' opinions about people's minimum needs. Townsend has argued that there are some essential (absolute) human needs in each society but the most important is that these needs must be defined by the society itself. Generally, undernutrition and starvation are recognized as the main indicators of absolute poverty. For instance, Sen (1983, 153) has expressed that "if there is starvation and hunger then, no matter what the relative picture looks like-there clearly is poverty." For Sen, malnutrition and hunger are two absolute elements of poverty, which should be priority of poverty elimination programmes in southern countries as he wrote (1981, 14) "while it can hardly be denied that malnutrition captures only one aspect of our idea of poverty, it is an important aspect, and one that is particularly important for many developing countries. It seems clear that malnutrition must have a central place in the conception of poverty." This opinion led him to argue with Peter Townsend, who was defending the opposite idea 'relative deprivation'. For Sen (1978) it was clear that malnutrition and hunger were absolute deprivations that developing countries were facing acute absolute poverty, and that Townsend's relative deprivation approach could only be a complementary perspective to the absolute poverty approach. In this issue, Sen (1978, cited in Mack and Lansley 1985, 20) has mentioned that "there is an irreducible core of absolute deprivation in our idea of poverty which translates report of starvation, malnutrition and visible hardship into a diagnosis of poverty without having to ascertain first the relative picture. The approach of relative deprivation supplements rather than competes with this concern with absolute dispossession." Similarly, Piachaud (1981) has argued that some deprivation items of Townsend were not of a great interest to understand the situation of the poor in southern countries, and that the research should not focus on the fact of being deprived or not, but, rather, on the reasons why the poor lack of opportunities to have access to some resources.

In fact, Sen was not against to relative perspective. He has argued that since all countries do not have necessarily the same living standards, the relative importance of capabilities items may change from one country to another. For instance, he gives the example of "being relatively poor in a rich country can be a great capability handicap, even when one's absolute income is high in world standards. In a generally opulent country, more income is needed to buy enough commodities to achieve the same *social functioning*"[173] (2006, 36). Sen has mainly argued for the importance of basic necessities in developing countries, but he was not against to the fact that the basic necessities can be determined in relative to country standards. However, he strongly emphasized the evidence of hunger, starvation, and the lack of basic needs of human life such as no access to sanitation, to health facilities, to electricity, to drinking water, to shelter, as the main most urgent concerns of developing countries[174]. In one sense it is clear that

[172] Townsend wrote that "poverty can be defined objectively and applied consistently only in terms of the concept of relative deprivation" (1979, 31). Piachaud does not agree with Townsend regarding this issue.

[173] Sen associates this relative deprivation issue to *social exclusion* which he considers one of the contemporary preoccupations.

[174] Sen (1996, 57-58) mentioned that "in the context of some types of social analysis, e.g., in dealing with extreme poverty in developing economies, we may be able to concentrate to a great extent on a

generally citizens of developed countries are more likely to encounter the social exclusion problems, whilst the citizens of developing countries are more likely to struggle for survival due to conflicts, war, starvation, hunger, and to being deprived of basic capabilities such as democracy, human rights, education, health facilities. On the other hand, Townsend has argued that nutritional requirements depend on cultural effects, diseases depend on the level of medical technology, shelter depends on what society uses as shelter (see Gordon 2000, 51). Moreover, Sen has emphasized that the conditions of absolute poverty may change over time and absolute poverty is not only about hunger and malnutrition (see Sen 1984, 1985, 1992, 1999). In particular, Sen has used the term *absolute* for drawing attention to alarming poverty situation in the world such as lacking access to food, water, education, shelter, sanitation facilities, health services.

Certainly adequate daily nutrition is the essential for surviving, and undernutrition, hunger, starvation are undeniably the most catastrophic forms of poverty. However, Townsend (1979) wanted to draw attention to other components of poverty, such as social exclusion, lack of social life. The main idea was that citizens can determine themselves the basic needs. Townsend's relative approach has contributed significantly to the analysis of poverty in developed countries. Nevertheless, for developing countries, his approach can be considered problematic regarding two aspects: (i) the link between cultural values and human rights; (ii) adaptation[175] to situations (see Clark 2012, Neff 2012, Qizilbash 2006, Teschl and Comim 2005, Crettaz and Suter 2013 for the concept of adaptation, and adaptation to poverty). In some developing countries, a number of violations of human rights are accepted by the majority of society as part of traditions and customs, for instance, different configurations of social pressure on women. Nussbaum (1995, 1999, 2000) has portrayed the lives of women who did not realize or who did not value enough the 'central human capabilities'[176], which these women later valued and approved. Another example is: in a community where the idea of girls should not go to school is part of traditions, the person who does not follow this practice will be considered relatively deprived according to Townsend's approach despite she just wants to achieve her right to go to school. From this perspective, relative deprivation can be dangereous because it can encourage some harmful conservative policies regarding minority and vulnerable groups. Which items are the most important can be asked to people, however it is important to make sure that people are conscious about human rights. The relative deprivation approach is, therefore, more suitable for the countries where democracy has reached a good level. Furthermore, several studies have showed that people can adapt themselves to negative situations for many reasons. For instance, Halleröd (2006), Droz (1998) have showed that some disadvantaged groups adjust their preferences, expectations to their monetary situation, or regarding Bourdieu (1979) the working-class population adjust their expectations and judgements to their living conditions.

relatively small number of centrally important functionings and the corresponding basic capabilities (e.g. the freedom to be well nourished, well sheltered, and in good general health, the capability of escaping avoidable morbidity and premature mortality, the ability to move about freely, and so forth). In other contexts, the list may have to be longer and more diverse".

[175] "Adaptation, in its broadest sense, refers to any action, process or mechanism that reduces the effects (perceptual, physiological, attentional, motivational, hedonic, and so on) of a constant repeated stimulus" (Frederick and Loewenstein 1999, 302).

[176] The concept is developed by Nussbaum (1995, 1997, 1999, 2000).

The relative deprivation approach is important for understanding different aspects of poverty and especially for understanding the extent of social needs (beyond subsistence) but it is not entirely correct to say that poverty is a relative issue. As Sen (1999, 1992, 1987a, 1987b, 1987c, 1984, 1981) has mentioned, poverty has also an absolute feature regardless of development level of country. Sen has pointed out that the lack of freedom, opportunities, possibilities are the absolute elements of poverty. For instance, taking vacation (e.g., travelling to another country may cost expensive due to transport and accommodation fees in U.S. dollar) is the human right of both Northern and Southern citizens even if Northern citizens have more opportunities, financial sources for travelling foreign countries compared to Southern citizens. Even if non-vacation becomes normal in the eyes of Southern citizens due to lack of money, of information or of other necessary resources and means, we cannot consider that non-vacation is not an item of poverty in Southern countries. At this point, for measuring poverty, the focus should not be on the list of relative deprivation items or of social necessities, but the focus should be on real possibilities, real opportunities people have to achieve what they value.

Absolute poverty approach influenced by nutritionists is still an important criterion for measuring poverty in developing countries where hunger and starvation continue to threaten many people. But we also know that people have different needs in their life from different perspectives such as educational attainment, access to health facilities, decent and safe working conditions, security in where they live. If we would like to make a list of needs, this can be a very long process and the list can go to infinity. But at least by being in consciousness of multiple needs, poverty studies attempt to extend beyond nutritional perspective including both social and economic aspects. Often, under this view, some items of dwelling conditions (e.g., toilet, bath, water, heat, kitchen, room, roof, sundries) are included in multiple deprivation-based poverty studies. In fact, Rowntree's absolute poverty line, and some other subsistence studies covered three main subjects that are nutrition, shelter and clothes. An intensive debate on whether these items should be considered as basic needs or not, exist in the literature. Basic needs defenders argue that there are some items that every single person must have in his or her life for surviving and/or for being able to start a good life. To that list, some other important elements of life such as access to education, access to health facilities, and drinking water are also included in development programmes for the Global South promoted and supported by international, non-governmental agencies and aid campaigns/programs. In contrast to absolute and basic needs approaches, relative deprivation perspective flourished new discussions highlighting various social necessities such as social participation, which may vary from one society to another (e.g., among countries, communities). The concept of relative deprivation attracted more attention of developed countries (advanced economy, affluent societies) in opposition to absolute poverty and basic needs approaches, which seemed to them unnecessary as the share of population suffering from hunger, starvation, accessibility problems to water, electricity, shelter, clothes is reported almost at negligible level in the Global North. The citizens of developed countries were rather reported suffering regarding precarious working conditions, weak social relationships, difficulties of balancing work and family time. In this regard, the idea of employing relative approach for measuring poverty in developed societies and employing absolute approach for measuring poverty in developing countries is widely agreed in the literature.

We can say that given the nature of poverty in developing countries, which still witness accessibility problems to electricity, water, drinking water, adequate shelter, public health and education services, the basic needs approach has remained a prominent concept to measure the

extent of poverty in the Global South. But there are two main disadvantages: first, even if developing countries are in significant delay regarding many aspects of development, various combinations of deprivation and prosperity emerge also regarding different social groups. For instance, some needs of upper and upper-middle class people will be different than those of middle class that will be different than those of lower-class no matter the overall level of development in the country. The needs should be updated and comprehensively determined in surveys regarding technological, social, political, economic conditions for different classes in society, including the least developed and developing countries. Second, in these two approaches, generally the value of 'must have items' is converted to a monetary metric such as what a level of minimum income the household/individual needs for meeting basic needs. Monetary perspective has always attracted attention of people because of its simple idea of only focusing on money. It is true that people need money for buying goods and services; but life is not only about money. If we think a little deeper, we can see that for earning money first of all people need a proper, safe and adequately paid job. But still having a good job is not sufficient in case when the person becomes gravely sick. Life is about various unexpected events and risks; herewith, social security and different types of insurance may emerge as quite important. People have also emotions. For instance, have a dignity, being self-confidence, having freedom, being respected, having family, friends, find a job that makes the person happy are the issues far from solved by money. A non-monetary approach requires a comprehensive analytic reflection towards life. This reflection, unfortunately, stays limited in the basic needs and absolute approaches. In parallel, monetary approach suffers of not taking in consideration social aspects of life (by reducing the function of the person to money). The minimum income levels are generally settled at 1 or 2 USD per day that may be still a very important amount for a citizen of a developing country, but still, it won't be sufficient to buy all goods and services both at national and international level. So, why to think that citizens of rich countries deserve more money and those of developing countries not?

When it comes to relative approach, even it seems that it covers more items and dimensions compared to absolute and basic needs, there are three main disadvantages: first, in practice, relative deprivation is often measured in terms of monetary metric. However, originality of this approach reposes on its social dimension. As we have explained in previous paragraph, an income level won't be able to detect directly the value of friendship, social relationships, emotions, family bonds. Second, relative deprivation is based on social necessities, which are determined as a function of values, customs, traditions, habits of society. But does everyone need to live in same way? or social participation should be the same for everyone? The answer is no. At this point, subjective poverty line seems a more adequate tool for determining social necessities. Relativity will vary as a function of individual perception and expectation. But, taking an average or a majority vote of public opinions on social necessities risks to hidden various individual point of views. The fulfilment of criterion of relativity requires a full coverage of subjective evaluation in concerned population regarding different items of life. However, this process appears complex and costly, notably for high populated societies and developing countries. Third, in many developing countries, the maintenance of democracy and human rights appears problematic. In such context, many people no matter their income level find themselves without voice to express their ideas and to live freely the way that they wish to live. Social rules may become oppressive threating freedom of citizens. Adaptation to poverty do often emerge as well among people living in hardship for long period. Chronic poverty, political conflicts, war, insecurity, repetitive fluctuations, high vulnerability to economic,

environmental shocks may lead people to adapt themselves to their disadvantaged situation and to try to be happy with what they have. To justify, we will give two examples. First example is: suppose that in a country where women cannot go to school or women cannot work or can work only in specific jobs; if we consider that this is the way that women get used to live as part of their tradition or social values, then being not able to go to school or being not able to work won't be considered as deprivation items in relative poverty perspective. Second example is: suppose that in a least-developed country where the access to private sanitation (e.g., toilet, bath) is shared by many dwellings or it is not adequately provided in their own dwelling; if people get used to live in that way, it is possible that for them having a private bath and a private toilet would not be a necessity. Of course, not everyone should live in the same way in the world, or should have the same values, or have to adopt their life to the technology-based luxury life or have to have the same materials. But the red line in relative approach to poverty is that the level of human rights, democracy and how the traditions, social values, social rules are defined become crucial in order to correctly distinguish between the social necessities determined in a free, democratic environment and the social necessities determined in a destiny, social pressure oriented non-democratic environment.

Given the conceptual problems emerging in absolute, relative poverty and basic needs approaches, the capabilities approach brings a new perspective to poverty analysis by emphasizing the importance of real freedom in defining and measuring poverty: to distinguish poor vs. non-poor in a society, the focus is shifted to from basic vs. relative terms to which items should be considered as deprivation indicators for measuring how capable is the person to live the way he or she would like to live without damaging the life of others. However, this approach is criticized mainly regarding two points: (i) there is no precise measurement framework; at first glance the theory seems vague, implicit, philosophical for applying it in a real word; (ii) theoretical framework promotes the notion of basic capabilities for measuring poverty in developing countries arguing that some functionings are fundamental for living a decent life (such point of view appears close to absolute poverty). Sen has argued that indicators, dimensions (i.e., items) are not defined in absolute terms (i.e., they may vary among societies), but the principle of capabilities is an absolute term. In fact, the capabilities approach attempts to take a universal level by focalizing on capabilities rather than on specific items of deprivation. Hence, instead of finding out whether a person is deprived or not in which items, the analysis of social, economic, political, cultural environment where the person lives (i.e., living situation of the person explained in terms of functionings including his or her opportunities, sources, social bonds, abilities, social and economic conditions, interactions with others, impact of external factors) becomes the heart of poverty analysis.

2.3 Capabilities and social bonds

As we have seen in previous section, Sen considers real liberty, capabilities as the absolute term and suggests updating the list of basic needs or the list of poverty items as a function of the social, economic, cultural situation of the society. In the literature, there are also 3 other approaches, namely social exclusion, social disqualification, marginalidad, which take in consideration social relationships and some aspects of relative deprivation. In absolute poverty, there is a list of some elements that are considered must-have/must-fulfilled for every single person in the society. So that, in general this list includes access to drinking water, food, electricity, heating, clothes, education, health services. On the other hand, in the following 3 approaches we will see in next paragraphs there is this idea of social necessities and poverty is

seen as the situation of not having social necessities or as being excluded from society due to the lack of socially perceived items or as being outside of the mainstream or as being marginalized from the mainstream of society. We will explain social exclusion, social disqualification, marginalidad approaches respectively and we will finalize this section by explicating the link between capabilities and social bonds, and therefore how the capabilities, social exclusion, social disqualification, and marginalidad approaches are connected.

Social exclusion, social disqualification, marginalidad

Originally, the notion of social exclusion comes from the French literature (Lenoir[177] 1974, Klanfer 1965) and it was used to refer in particular to people who were excluded from the French Bismarckian[178] social insurance system such as disabled, lone parents, uninsured unemployed people (Evans 1998, Duffy 1995, Silver 1994). Lenoir (1974) has evoked the notion of *social maladjustment* of some group of people into the mainstream of society such as criminals, young drug addicts, alcoholics, the mentally ill and other vulnerable groups (cited in Paugam (ed.) 1996, 10) not due to their personal traits and behaviours but due to functioning of society (e.g., income inequality, inequality regarding access to adequate education and health services, geographical segregation in metropolitan areas). The focus, therefore, was on how to reinforce again social bonds, especially how to ameliorate social integration and solidarity in French society. In this framework, a number of insertion programmes such as the Revenu Minimum d'Insertion (RMI)-guaranteed minimum income (i.e., income given to individuals who have no earnings or income given to individuals who have low income regarding the official line set by decree) was approved by French Government in 1988 (but RMI is removed later in May 2009). However, social exclusion studies have started to take new shape from the 1980s in France focusing on migrants, youth people, and especially urban ghettos of France called *banlieues* (Dubet 1987, Fassin 1996). This movement has led to the emergence of 'new poverty' approach. Notably, new policies were developed with the aim of improving social integration of second-generation migrants.

Among contemporary studies, different definitions of social exclusion are proposed by scholars focusing primarily on social participation in relative and dynamic aspects (see Paugam (ed.) 1996, Robinson and Oppenheim 1998, Burchardt 2000, Burchardt et al. 2002). For instance, Duffy (1995) mentions "inability to participate effectively in economic, social, political and cultural life, alienation and distance from the mainstream society", Phillips (2006, 116) employs "detachment from the mainstream society, dissociation from social milieu, social polarisation, lack of social rights" to define social exclusion. Similarly, but emphasizing *process and time*, Walker and Walker (1997, 8) conceptualize social exclusion as "the dynamic process of being shut out…from any of the social, economic, political and cultural systems which determine the social integration of a person in society." Atkinson (1998, 10) settle three characters of social exclusion: relativity (being socially excluded depending on social relations and functioning of society), agency (actions of who exclude others define who will be excluded), dynamics (social exclusion evolve over time and from generation to generation). Ruggeri-Laderchi et al. (2003, 257) describe social exclusion as "the processes of marginalization and deprivation that can arise even in rich countries with comprehensive welfare provisions". At international platform, especially thanks to the initiatives and works of United Nations Development Programmes,

[177] René Lenoir was the Secretary of State for Social Action in the d'Estaing Government (1974-78).
[178] The term comes from the German national insurance framework implemented under the administration of Bismarck (see Gordon and Spicker (eds.) 1999, 18).

social exclusion is also evoked as a "lack of recognition of basic rights, or where that recognition existed, lack of access to political and legal systems necessary to make those rights a reality" (Burchardt et al. 2002, 3).

Levitas (2000 [1998], 359) proposes three different approaches to define social exclusion. First approach is the *redistributive discourse* (RED) that determines social exclusion as a result of poverty. For defending this approach, Levitas refers to Townsend's relative deprivation concept that suggests "individuals, families and groups can be said to be in poverty when…their ressources are so seriously below those commanded by the average individual or family that they are, in effect, excluded from ordinary living patterns, customs and activities" (1979, 32). Levitas (2000 [1998], 359) accepts that "within this broad RED framework, social exclusion is understood to means something more complex than is colloquially understood by poverty, in that it is dynamic, processual, multidimensional, and relational, and it allows space for the understanding that discriminatory and exclusionary practices may be causes poverty. Nevertheless, poverty remains at the core." Based on Levitas' arguments, unequal distribution of resources, means, and opportunities within society ascertains poverty as a relative deprivation, which generates also different types of social exclusion. From this perspective, we can consider that social exclusion is one of the forms of (multidimensional) poverty. However, it is important to note that not all poor people should feel socially excluded or they should be considered as socially excluded. Second approach is *social integration* discourse (SID) "in which paid work is represented as the primary or sole legitimate means of integrating individuals of working age into society. The excluded are those who are workless or, in the case of young people, at risk of becoming so." (Levitas 2000 [1998], 359-360). According to Levitas (2000 [1998], 360), on the one hand, low income is the major indicator of social exclusion regarding redistributive discourse, on the other hand, labour force participation[179] (e.g., unemployment, unpaid work) is the major indicator of social exclusion regarding social integration discourse. In this second approach, unemployment or at risk of unemployment are seen as the main factors inhibiting social inclusion. And third approach is *moral underclass discourse* (MUD) "which emphasizes moral and cultural causes of poverty, and which is centrally concerned with the moral hazard of dependency, and thus with workless households rather than individual labour-market attachment" (Levitas 2000 [1998], 360).

In practice, for measuring social exclusion, various indicators were developed by researchers in the literature. The Institute for Public Policy Research (IPPR 1998) define four dimensions and fifteen indicators. The dimensions are: income poverty, exclusion from the labour market, exclusion in education, health. All indicators are objective. For instance,[180] for the dimension

[179] Levitas (2000 [¦998], 360) uses rather the term "unemployment or economic inactivity" as the lead indicators of SID approach to social exclusion because of unpaid, non-market work. But people have a freedom to choose to be economically inactive as well. This is why, we suggest that it is better to use the term labour force participation, or more specifically unemployment, unpaid work, informal work.

[180] Income poverty indicators: proportion of the population falling below 50% of average household income, change in average real income for people in the bottom three deciles, length of time spent in the bottom three deciles, incidence and length of time on income support and/or housing benefit (area based). Exclusion in education indicators: proportion of 16 years old failing to get at least 20 GCSE points, average GCSE points score in school league tables. Health indicators: standard mortality ratio in social class IV/V relative to other classes, for men and women, adding an equity dimension, such as social class, socioeconomic group or deprivation, to the targets for heart disease, stroke, accidents and

of labour market, indicators are: unemployment rate defined by the International Labour Organization, employment-population ratio, proportion of workless households, duration of unemployment, claimant unemployment rate to track changes at local level. These indicators are generally provided in national labour force survey reports, and some of them are also used in multidimensional poverty measurement framework. Robinson and Oppenheim (1998, ii) suggest employing additional indicators related to "housing, crime, environment, family breakdown, social and political exclusion, social capital, participation to civic organization, and social support networks". In British case, the *Breadline Britain* report (see Gordon and Pantazis 1997, Bradshaw et al. 1998) include both the question of whether a person has or not the item and the question of whether the person sees the item as necessary in her/his life. In this methodology, social exclusion is conceptualized from redistributive discourse (RED, see Levitas 2000 [1998]) perspective. Herewith, this report provides a comprehensive picture of social participation regarding activities and consumption. Reader can see easily the average tendency of consumption and activities in British society, and the percentage of people who cannot have access to goods and services that they would like to obtain.

The *Poverty and Social Exclusion Survey-PSE*[181] *of Britain* (see Gordon et al. 2000, Townsend and Gordon 1991, Gordon and Pantazis 1997, Pantazis et al. (eds.) 2006) initiated at 1999 and funded by Joseph Rowntree Foundation is also in same line with the Breadline Britain survey (Mack and Lansley 1985, Gordon and Pantazis 1997) but PSE is more extended in terms of questionnaire including both the questions of deprivation and social exclusion such as familial and friendships networks. The authors, in PSE, identified four domains of social exclusion: financial exclusion (inadequate income), labour market exclusion, service exclusion, exclusion from social relations. Gordon (2000) qualifies this survey as very comprehensive (i.e., including in-depth analysis of social exclusion, the incidence and severity of poverty, the link between poverty and social exclusion) and scientifically rigorous (i.e., PSE is based on non-arbitrary deprivation calculations, public perception of poverty and subjective poverty line). PSE asks to interviewees what are the life necessities that every adult and child should be fulfilled with. People deprived of *socially perceived necessities* (i.e., these are the needs considered as the basic needs for living a decent life by most of the sample) are considered poor. As in other studies such as of Rowntree, or of Booth, the PSE method has also defined a poverty line, which is subjective[182] here because it is defined by society itself, and not by the experts. In PSE, as recommended by the Copenhagen approach[183], the indicators of standard of living, absolute

cancer set in our healthier nation (DoH 1998), tracking the incidence of depression with an equity dimensions, indicators to track prevention strategies and access to health care, with an equity dimension.

[181] See http://www.bristol.ac.uk/poverty/pse. The survey was conducted by researchers from the Universities of Bristol, York, Loughborough and Heriot-Watt in collaboration with the Office for National Statistics (ONS).

[182] Gordon (2000, 59) mentions that subjective poverty lines have been already used in other country contexts such as by "Hagennaars (1986), Hagennaars and DeVos (1988) in Netherlands, Danziger et al. (1984), Colsanto et al. (1984) in the United States, Callan et al. (1989), Nolan and Whelan (1996) in Ireland, Townsend and Gordon (1991), Townsend et al. (1997), Gordon et al. (2000a) in Britain, Saunders and Matheson (1992) in Australia, and Halleröd (1995a, 1995b) in Sweden."

[183] The World Summit for Social Development has taken place in Copenhagen in 1995, by the United Nations. The report of the conference has highlighted the need of substantial reduction of overall poverty and the eradication of absolute poverty (UN (1995, 60-61)).

poverty and overall poverty are selected. Absolute poverty and overall poverty[184] include five thematic: adult items, adult activities, children's items, children's activities, and other. Adult activities include familial and friendships networks, the indicators are: visits to friends/family, visiting friends/family in hospital, attending weddings, funerals and other such occasions. Concerning child activities, the indicator is celebrations of special occasions. These indicators reflect social exclusion[185].

Moreover, Burchardt (2000, 388) mentions that "an individual is socially excluded if he or she does not participate to a reasonable degree over time in certain key activities of his or her society, and (a) this is for reasons beyond his or her control, and (b) he or she would like to participate.[186]" In a broad sense, it is a very clear definition; however, the terms 'reasonable degree' and 'certain key activities' cause uncertainty notably for the measurement process. Based on this definition, it is important to follow the same individuals over time for observing their behaviors (i.e., longitudinal/panel data is needed). Concerning key activities, it will be important to ask directly to interviewed people what do they consider themselves as key activities (e.g., Breadline Britain). Burchardt (2000) accepts that her definition is intentionally set vague for facilitating operationalization of the concept. For testing her definition of social exclusion, Burchardt used British Household Panel Survey (BHPS) from 1991 to 1995 (wave 1 to wave 5)[187]. Four dimensions and one indicator per dimension are chosen: *consumption* measured by 'low income', *production* measured by 'not engaged in a socially valued activity' (e.g., paid work, volunteering), *political engagement* measured by 'does not vote and not part of any campaigning organization', *social interaction* measured by 'lacks emotional support' (e.g., relationships with friends or family). Further, Burchardt et al. (2002) by using the same methodology, has measured social exclusion in a dynamic and multidimensional perspective (see also Burchardt et al. 1999) based on BHPS for the years 1991 to 1998. They define four domains of social exclusion: consumption, production, political engagement, social interaction. If a person lacks participation regarding at least one of four dimensions; he or she is considered socially excluded. Their analysis shows the degree of exclusion, from "the less the extent of participation" to "socially excluded to a greater degree", measured in terms of number of dimensions in which the person is excluded and of the duration of exclusion.

Today, several poverty studies emphasize the link between social exclusion and poverty, and the use of their combination in non-monetary poverty studies (see Lister 2004, Nolan and Whelan 2007, 2010, but also see Silver 1994, Silver and Wilkinson 1995 for the discussions on the conceptual difference between poverty and social exclusion). For instance, some studies

[184] Absolute poverty and overall poverty are defined based on the United Nations World Summit on Social Development in Copenhagen in 1995. Gordon (2000, 75) writes that "absolute poverty means not having the basic necessities of life to keep body and soul together. Overall poverty is not having those things that society thinks are basic necessities. Overall poverty also means not being able to do the things that most people take for granted (either because you can't afford to participate in usual activities or because you are discriminated against in other ways. What constitutes overall poverty will vary between different societies and at different points in time."

[185] The complete list of indicators is given by Gordon (2000, 68-70).

[186] Author mentions that this definition is based on a definition proposed by Julian Le Grand at a meeting of the Centre for Analysis of Social Exclusion.

[187] Burchardt points out a very important technical issue: household survey does not include the institutional or homeless populations (see Burchardt 2000 for more technical details on survey). Especially, for developing countries, it may be quite problematic using household survey for measuring social exclusion and poverty.

define poverty as a social exclusion (see Levitas 1996, 1998; Abrahamson and Hansen 1996), or others see social exclusion as a one dimension of poverty among others. Regarding Room (1995), social exclusion and poverty studies differ in three aspects. Room's first criterion is that poverty focuses more on income, expenditure, consumption whilst social exclusion rests on multidimensional deprivation and disadvantages. In line with this condition, Bergman (1995) perceives poverty as a lack of sufficient earnings but emphasizes the detachment from the mainstream society in defining social exclusion as a failure of social systems resulting in unfulfillment of citizenship rights to some group of people. Similarly, Burchardt et al. (2002, 6) mention that poverty focuses on deprivation, lack of resources; but social exclusion aims to identify "not only those who lack resources, but also those whose non-participation arises in different ways: through discrimination, chronic ill health, geographical location, or cultural identification, for example."[188] Room's second criterion is that social exclusion studies take more often a dynamic perspective. In this regard, Alcock (2006 [1993]) mentions that social exclusion is not a static phenomenon unlike poverty, rather it corresponds to a process focusing on what others do in society. Robbins et al. (1994) agree also that social exclusion has a more dynamic feature compared poverty, going beyond the concept of insufficient income, or monetary resources. It reflects the rupture of social link, the status of social participation in society. And Room's third criterion is that there is a shift from individual or household level to local community level in social exclusion studies.

In Europe, notably since the early 1990s, instead of underclass, the social exclusion approach is more and more used together with the concept of poverty. Marlier and Atkinson (2010, 285) define social exclusion from a multidimensional perspective as "the involuntary exclusion of individuals and groups from political, economic, and social processes, preventing their full participation in the society in which they live" by emphasizing that poverty is the main cause of social exclusion. Related to that, they suggest analyzing together poverty and social exclusion for promoting social inclusion. On the other hand, Evans (1998) argues that European Union social policy has targeted social exclusion, instead of poverty, since the early 1990s, due to increasing financial problems, emerging new social assistance demands. Social exclusion is mentionned in the European Union (EU) anti-poverty programme in 1988 and in the preamble of the European Social Charter in 1989 (see Bergman 1995). During the Lisbon Summit in 2010, the European Union member states have adopted for the use of the concepts of social cohesion and social inclusion in quality of life and poverty analysis; moreover, in the framework of the Europe 2020 project, the concept of social exclusion has been employed to measure poverty from a multidimensional perspective in addition to the classical indicator of material deprivation and relative monetary poverty. The 2000 Lisbon Agenda included social inclusion goals together with those of economic growth and employment, under the approach of 'Open Method of Coordination' (see Marlier and Atkinson 2010). At national level, the British government, under the Labour Government of Tony Blair in 1997, has established the Social Exclusion Unit (SEU onwards, see SEU 1998a, 1998b, 1999) in Britain. The Economic

[188] Brady (2009) argues that the capability approach to poverty (see Capabilities sub-section) and social exclusion have some similar parts, such as a socially excluded person has capability deficit in participating effectively in society. Atkinson (1998, 27) expresses that poverty involves social exclusion in the sense that the poor are "people being prevented from participation in the normal activities of the society in which they live or being incapable of functioning".

and Social Research Council's (ESCR onwards) Center for Analysis of Social Exclusion (CASE), established in 1997 at the London School of Economics, has a pioneer role in promoting the importance of social exclusion in social sciences (notably in the context of Britain)[189].

Most recently, in the context of Europe 2020 strategy (during the European Council in June 2010, 5 objectives were defined. One of these focuses on poverty. In this regard, some indicators are determined to measure the progress in eliminating poverty. This objective is also known as 'social inclusion' project, see Marlier et al. 2007), by employing the survey EU-SILC (European Union Statistics on Income and Living Conditions survey, see Atkinson and Marlier (eds.) 2010b) to monitor poverty and social exclusion in European Union. Eurostat has developed the AROPE indicator namely 'at risk of poverty or social exclusion' (i.e., to monitor the target of reducing poverty by moving at least 20 million people out of the risk of poverty or social exclusion by 2020 in European Union region). The AROPE indicator includes three groups: (1) people who are at risk of living in income poverty after social transfers (i.e., their equivalized disposable income is below 60% of the national median equivalised disposable income); (2) people who are severely materially deprived (i.e., they cannot afford at least four of nine material items. The nine items are: unable i) to pay rent or utility bills; ii) unable to keep home adequately warm; iii) unable to cover unexpected expenses; iv) unable to eat meat, fish or a protein equivalent every second day; v) does not take a week holiday away from home, or could not afford such (even if wanted to), vi) unable to afford a car; vii) unable to afford a washing machine; viii) unable to afford a colour TV: or ix) unable to afford a telephone.); (3) people who live in households with very low work intensity (i.e., people from 0–59 years of age living in households where adults[190] worked less than 20% of their total potential during the previous 12 months). In this light, in European Union, a person who experience at least one of these three situations is considered to be at-risk-of poverty or socially excluded.

Nevertheless, Maître et al. (2014) criticized the EU2020 indicator in the sense that neither union (i.e., being deprived at least regarding one indicator) nor intersection (i.e., being deprived regarding all indicators) of the three indicators is an efficient indicator for combatting poverty. They propose another concept called 'consistent poverty' focusing on individuals below the monetary poverty line but at the same time deprived in at least 3 material items. By excluding the indicator of low intensity of work and combining the indicators of material deprivation and relative monetary, Maître et al. (2014) argues that the 'consistent poverty' approach reveals substantial variations between countries compared to the AROPE approach (see also Nolan and Whelan 2011 that has precedingly shed in light some inconvenient of the European Union's approach on the elimination of poverty in the Euro zone). If the conceptualisation of poverty via social exclusion seems to be an important perspective to analyse poverty's multiple dimensions, however it seems very complex to define poverty like that because the social exclusion indicators itself are indirect, abstract and multidimensional in many cases (see Lister 2004, and Levitas 2000 [1998] for multidimensional character of social exclusion). Since there is no agreement among scholars on how to define and measure social exclusion, an alternative may be to consider social exclusion as one dimension of poverty.

[189] The ESRC (1997; cited in Pantazis et al. 2006, 126) mention social exclusion as "understanding the processess by which individuals and their communities become polarised, socially differentiated and unequal."

[190] Those aged 18–59, but excluding students aged 18–24 yrs. old.

As we have seen, in the European literature, there is a quite considerable amount of works on social exclusion approach to poverty for especially explaining the vulnerable situation of the poor regarding how they feel rejected or how they are in fact rejected by society as a whole. Now, we will see another approach called *social disqualification* developed by French sociologist Serge Paugam in the early 1990s. Within this framework, he focused on three disadvantaged groups, namely: *fragile*, *assisted*, and *marginal* based on his investigation realised in Saint-Brieuc in France (Paugam 1991a). His works were influenced by Durkheim's (2007a [1893], 2009 [1894], 2007b [1897], 2003 [1912]) sociological ideas and methodology, Tocqueville (1835) and Marx's (1847) ideas on pauperism and the role of charity, Simmel's (1998 [1907]) definition of being poor and assisted, and Castel's (1990, 1994, 1995a, 1995b, 2003, 2009) concepts of disaffiliation[191], exclusion, and social insecurity. Further, he also pointed out the question of social exclusion in Europe, notably in France, and he formulated a typology of poverty referred as *different elementary forms of poverty* (Paugam 2005). But, first of all we will focus on his social disqualification approach.

Paugam's work (1991a) on social disqualification has brought a new perspective to social exclusion and poverty studies in the sense that his theoretical framework is sociologically well established inspired by the question of social assistance, and social bonds in modern societies. Traditional social exclusion studies were more focusing on (i) description of social life conditions of the (monetary) poor compared to rest of society without referring to sociological theory; (ii) socio-political aspects in terms of implementation and evaluation of social policies, and of welfare regimes against poverty; but they were less focusing on typology of socially excluded groups. At that time, Paugam (1991a) has significantly filled this gap in the literature. He is interested in social relationships encountering a person in her/his life, and how people are connected to each other in society (e.g., to social groups, and to society as a whole). He recalls Durkheim's main approach in *Division of labour* (2007a [1893]), so that the (modern) societies are becoming more independent but at the same time individuals are becoming dependent to society. Durkheim mentions that understanding the mutations of social solidarity will be an efficient solution to resolve this contradiction. He adds that division of labour observed in modern societies is not an obstacle for social solidarity, in opposite he argues that division of labour is essential for social solidarity because it makes people feel useful regarding themselves and society (see Paugam 1991a). At that point, inspired by Durkheim's ideas, Paugam has considered labour market conditions and social ties as main application fields for understanding and examining social solidarity and therefore for explaining poverty. In following paragraphs,

[191] Original name in French is 'désaffiliation' to express 'social disengagement' emerged as a social product due to labour division and social values. In French, Castel's definition for 'désaffiliation' is "le décrochage par apport aux régulations à travers les quelles la vie sociale se reproduit et se reconduit (Castel 1990, 157). Through this expression, Castel means someone who stays outside of the main social mechanisms and regulations of her/his society, which are the essentials for reproduction and continuum of social life in society. He notably emphasizes that persons who are disaffiliated, are nor deviant neither out of social order; they are not in failure, but they are set back from their society (see also Castel 1991). He mentions that 'a marginal' (e.g., vagrants) is a socially disengaged person who has broken her/his ties with her/his original community (Castel 1996, 34- "Le marginal a rompu ses attaches avec sa communauté d'origine. C'est un désaffilié »). A marginal person is not necessarily a monetarily poor person, or a monetarily poor person is not necessarily a marginal. Moreover, Castel emphasizes that exclusion does not mean marginalization but, in some cases, exclusion may lead to marginalization.

we will shortly see Paugam's typology of social disqualification, social ties, social solidarity, and poverty respectively.

Paugam (1991a, 1993, 2005, 2015 [2007]) argues that society changes as social relationships and the look of people to life, to individuals, to society take different shapes over time. Herewith, in his research, poverty manifests as a social phenomenon that got different forms within society, and among countries. His early studies focus on *social disqualification*, which emerges in the life of people who pursue their life with social allowance (see Paugam 1991a). He describes the situation of being socially disqualified as not being able to participate effectively into social and economic life. Regarding him, this type of situation can be explored thanks to personal feelings and perceptions towards social contacts and different situations encountered in life. He showed different aspects of social disqualification as a result of his investigation among assisted people (i.e., having recourse to social assistance) in Saint-Brieuc. His analysis includes in-depth interviews as well as empirical studies such as descriptive statistics (e.g., employment and social protection situation, types of households in Saint-Brieuc[192]), and correspondence analysis among beneficiary of social assistance.

Paugam (1991a, 31) determines three groups of social action beneficiaries as socially disqualified people, respectively: "fragile people, assisted people, and marginal people". Specifically, the fragile ones get occasionally benefits from social action; the assisted ones are permanently dependent on social action; and the marginal ones get infra-intervention type social action. Regarding him, the main characteristic of the fragile group is that these people do not have a regular income or there is an incertitude regarding their employment status or contracts, for instance, unemployed, people having a working contract for a certain period of time (e.g., interim, seasonal job), interns, trainees, and people working part time in non-qualified jobs. In opposite, the assisted group receives regularly income from social protection unit of the state. Physically or morally disabled people, or people who do not have access to education, or people who do not have enough opportunities, means, or resources for taking care of their children are generally part of this group. On the other hand, the marginal group do not have nor a stable job neither a regular income; but also, they do not benefit regularly social action.

Based on this typology, Paugam (1991a, 50) put in evidence how socially disqualified people cope with their lower status in their society given that social disqualification may have a very negative consequence on psychological state of the person. He explains that: (i) the fragile group faces two types of fragility: internalized fragility (i.e., feeling humiliated, do not want to communicate with others, emotional instability, be angry to others) or negotiated fragility (i.e., young adults less than 25 yrs. without a regular job or a regular income. For them, this assistance situation to social services is just temporary until they find a regular job); (ii) the assisted group faces three types of assistance: deferred assistance, installed assistance or claimed assistance. Among these three types of assistance, the worst one is the installed assistance in the sense that even if all of assisted people do regularly contact social services, the dependency becomes stronger, and motivation to find a job or a better job becomes less for the group of installed assistance; (iii) the marginal group faces two types of marginality: conjugal marginality, or organized marginality. The conjugal marginals distinguish from the organized marginals in

[192] Based on 1985 Labour Force Survey, and 1982 population census carried out by the INSEE (Institut National de la statistique et des études économiques).

terms of their general will for having a better social and professional integration. However, regarding Paugam, it seems very difficult for marginals to get out of their vicious circle in a short term. Concerning the organized marginals, they have tendency of developing their own sub-culture in a good way, which helps them to resist against stigmatisation and diverse hardships. They are not against social contacts, but they are aware of their very disadvantaged situation, which often puts them apart of the rest of society.

After having introduced the concept of social disqualification, two years later in 1993, by still remaining close to the link between poverty and social ties, Paugam (1993, 51) has described poverty in French society. His main goal was to unveil new approaches and new realities to explaining multi-aspects of poverty. He has argued that insufficient amount of adequate housing was not any more a main indicator of poverty. People living in miserable conditions such as in slums or shantytowns were not the only poor asking social assistance. Regarding him, the profiles of applicants to social assistance were quite various. He has pointed out two main drivers of *new* poverty in France: (i) malfunction of labour market such as highly increasing precarious jobs, and long-term unemployment touching especially young adults without sufficient professional experience; (ii) increasingly weaker social ties within family members, and within society in terms of solidarity.

Paugam (1993, 167) suggests using indicators of social ties (i.e., the intensity of relationship with different social groups such as family, friends, institutions) and labour market (i.e., the position of social allowance recipients in labour market, how these people are seen in labour market in terms of the values given to their skills and professional experiences) to find out different forms of social and economic integration of socially disqualified poor people. According to results of his analysis covering people receiving a minimum income of insertion, called *Revenu minimum d'insertion*[193] (RMI), he has unveiled three types of groups experiencing: (i) positive value in labour market but weak social ties (type 1); (ii) lack of value in labour market but good intensity of social ties (type 2); (iii) greatly lack of value in labour market and greatly lack of social ties with different social groups (type 3). The type 1 includes mainly young people less than 35 yrs. with no stable job (as a result, low and uncertain income with several periods of unemployment), no stable housing, and very weak social ties with their family. Often, they do not have an adequate or sufficient professional skill, or their skills are not generally demanded in labour market. The type 2 includes people having difficulties of accessing to stable jobs because of their age, fragile health situation, lack of professional experience and skills. Generally, these people live in decent housing conditions and they do have a very good regular relationship with their family. And the type 3 includes mainly man with no partner and no child. They do not have or have a very few intensities of relationship with their family. Often, they become alcoholic and homeless or with no fix housing.

Paugam (2018 [2008]) has proposed a typology of social ties, for explaining social solidarity, and showing a typology of poverty. Regarding him, there are four types of social ties: *organic participation bond*, *parentage bond*, *elective participation bond*, and *citizenship bond*. Paugam (2015 [2007], 37) has built his social bonds typology based on two concepts: protection, and recognition. Protection includes family, community, professional, and social support

[193] It is a kind of social welfare, supplementary income benefits for people having no income or for people having an income below the standard. It is firstly put into effect in 1988 by social government in France, later it is removed in 2007, but in 2009 it is put again into effect.

encountering a person who may need them during her/his life against negative events. Recognition reflects how a person feels valued in the eyes of others. Thanks to protection, a person may feel stronger and ready against hazards, which in turn diminishes anxiety and social insecurity; and thanks to recognition a person may feel useful, precious by her/his existence in society, which makes sense to her/his life. Paugam showed empirically that the assisted poor people feel socially disqualified when they do not have a stable, good quality, good paid job. He calls this relationship between the individual and her/his work as the *organic participation bond*. Certainly, it is very important to be active in labour market; however, it is not sufficient. Individuals need to be protected, making ensure that their future is guaranteed against poverty, unemployment, precarity. Otherwise, social insecurity emerges (see Castel 2003, 2009; Bourdieu (ed.) 1993, chapter IV). Paugam (2015 [2007], 52) distinguishes two types of precarity: (i) job insecurity (*précarité de l'emploi*) meaning jobs with fixed contracts or with high probability of contract termination; in that case, the person has an inferior social status; (ii) labour insecurity (*précarité du travail*) when a person feels worthless and his or her job seems irrelevant and his or her job is not recognized by colleagues at workplace. On the other hand, *parentage bond* reflects the relationship of a person with her/his family. The *elective participation bond* covers the relationship outside the family, with other people such as within institutions, religious community, diverse leisure groups, neighborhood, friendship, marriage, so on, and its intensity and variety depend significantly on social backgrounds. The *citizenship bond* focuses on patriotic feelings. In democratic countries, all citizens should be recognized equal in front of the law. In that sense, accessibility to law, to rights, trust in people and trust in legal system become very important for providing social integration in society.

Social ties play substantial role, as each type of them has its own features and importance, in social and economic integration of society as a whole. Paugam (2015 [2007], 62) defines four types of integration: assured integration, fragilized integration, compensated integration, marginalized integration. *Assured integration* is the best way of integration, which seems quite utopian to be realized nowadays in societies. People of highest class in social hierarchy are more likely experiencing an assured integration. Furthermore, protection and recognition are greatly achieved in assured integration. In *fragilized integration*, the consistency of social ties over time is uncertain, however they are not totally broken. But individuals experiencing this type of integration feel frustrated in somewhere in their life. In *compensated integration,* individuals are more in the mood of resistance toward life in general. Their social ties are partially broken. They are integrated regarding some aspects, but they are not integrated regarding some other aspects of life (i.e., positive sides compensate negative sides in coping with social disqualification). In *marginalized integration,* social ties are systematically broken, individuals are in search of how to survive.

In the framework of Paugam, the analysis of social bonds is crucial for understanding how the individuals are attached to their society. After that, understanding types of integration existing in society allow us to be aware whether the society itself is integrated or not as a whole. All of these integration forms give shape to social regulation in society. In this perspective, in societies where social regulation reaches its optimum level, poverty in its different forms should be minimized. Each society can have its own configuration of social ties, therefore, its own integration and regulation. For instance, French society is dominated by the organic participation bond, which plays a regulator role in the integration of society as a whole. Fragilization of this tie, stimulated by mass unemployment and increasing precarious working

conditions, make emerge different social problems including a high level of social disqualification.

Based on these arguments mentioned in previous paragraph, Paugam (2005, see also Paugam 2018 [2008], Paugam and Duvoux 2008) has extended his theoretical framework and has proposed a typology of the elementary forms of poverty, namely: *integrated poverty*, *marginal poverty* and *disqualifying poverty*. In *integrated* poverty, labour market conditions are defined by a low level of economic development leading to informal employment and hidden unemployment. People do not receive a minimum guarantee income and they benefit a few from public social protection. But relationship with family members is characterized by strong solidarity sense, and people take generally a protective role with their close entourage. They are not very much stigmatized of their situation; they do not feel socially excluded notably thanks to solidarity among family members. In *marginal* poverty, there is quasi-full employment, which in turn manifests in less appeal on family help or at least keep at same level. Social protection system is widespread in this group, and minimum guarantee income is received by persons in destitution (i.e., extreme poor, absolute poor, severely poor, ultra-poor[194]). Often, the marginal poverty is less frequent compared to integrated poor, but the marginal poor feel strongly stigmatized, and they are generally considered as people unable to adapt themselves to modern society. In *disqualifying poverty*, there is a high number of unemployment, unstable jobs (i.e., integration in labour market is limited). In addition, social bonds are also weak, notably concerning unemployed and the most disadvantaged people in this group. As a result, they do appeal on minimum guarantee income and other social allowances. Heterogeneity in this group appears significantly in terms of social status, and living conditions (e.g., dwelling, health, social participation); but one main common point is almost all members of community do feel at risk of social exclusion. To Paugam, (i) *disqualifying* poverty is likely to be happen in post-industrial societies (notably in societies with high increase of unemployment and precarious status in labour market); (ii) *marginal* poverty is likely to be happen in advanced industrial societies, notably in countries where rate of unemployment is few and social protection coverage is high such as Scandinavian countries; (iii) *integrated* poverty is likely to be happen in traditional societies where economic and human development is limited.

Most recently, Paugam (2016) has linked three types of poverty perception[195] to four configurations of societal attachment based on social bonds. There are four types of configurations: *familialist* configuration (*parentage bond* is prominent), *voluntarist* configuration (*elective participation bond* is prominent), *organiciste* configuration (*organic participation bond* is prominent), *universalist* configuration (*citizenship bond* is prominent)[196], which are defined based on economic development, inequality level, social protection system, and citizenship responsibility in the country. In each configuration, he explained whether poverty is perceived as a normal process (*naturalization*) or as a result of incompetency of the poor in the sense that the poor are responsible for their poor situation (*blaming*) or as a result

[194] Lipton (1983, 1988) introduced the term 'ultra-poor' in identifying the poorest group in a monetary perspective. We also note that the World Bank's absolute poverty lines (e.g., $1.25 per day) are settled for detecting the extreme poor in monetary perspective.

[195] Van Oorschot and Halman (2000), Paugam and Selz (2005) provide also empirical studies on poverty perception in Europe.

[196] The term 'prominent' means that the bond in question is both integrator and regulator of society (see Paugam 2016, 129).

of dysfunction in society where the poor are the victim (*victimization*). *Familialist* configuration is characterized by a limited level of decommodification, low level of citizenship responsibility and high level of social inequalities. In this type of societies, *integrated poverty* emerges and poverty, therefore, is seen as a natural event. In *voluntarist* configuration, poverty is often perceived as an inevitable social phenomenon and as a kind of punishment for people who are not capable of achieving in their life or they are not enough brave for achieving to live a good life. The poor are generally judged as unskilful, dilly-dallier, irresponsible by public and when the poor ask for assistance to survive, they are in general *blamed*. In *organiciste* configuration, the poor constitute the lowest strata in society. They are, in fact, not out of social order but they are seen as a *victim* who do need social welfare. The intervention of state is more frequent to reduce inequalities, to provide social welfare. Egalitarianism in sense of citizenship is quite guaranteed but equality of living conditions is not fully fulfilled. In *universalist* configuration, *victimisation* of the poor appears more strikingly compared to organiciste configuration. The state primarily aims to provide equal rights to every citizen in a democratic angle, regarding both economic and social life, and to socialize the poor. In this perspective, poverty is rather perceived as a reflection of incapacity of society to achieve equality for every citizen. There is a political will of preventing the poor to stay permanently outside of society because poverty is recognized as an important threat against social cohesion.

Conceptually not very far from social exclusion and related to culture of poverty in some respects, the studies of *marginalidad* especially through the works of Gino Germani (1973) explore a certain way of living called *marginalidad* (*marginality* onwards) in Latin America[197]. Following paragraphs reflect the ideas of Germani in his book *El concepto de marginalidad* (1973), and its English version *Marginality* (1980). Germani (1973, 12-14) mentions that the notion of marginality[198] has been initially expressed in urban ecological studies, more precisely, in description of living conditions of segregated population sectors in areas deprived of urban service system with improvised dwellings and illegally occupied lands (Germani suggests see Mangin 1967, Morse 1965, Turner 1966, CEPAL1969). In particular, some descriptive evidence show that an important part of urban marginal class is composed of internal migrants moved from rural areas to urban areas. This population appears significantly different than the rest of urban population in terms of cultural divergence (i.e., culture of city) such as the marginal population adopts different configurations regarding behaviors, norms, values in opposition to the cultural features of modernity. For instance, Juppenlatz (1970) proposed the concepts of 'ruralization of the city' and 'cancer on the carapace of the city' to describe the fact that migrant peasants bring their rural and traditional culture with them to their city life, as they are later perceived as 'marginals' to the mainstream society (or 'on the margins of society', see Fassin 1996, 54). The 'marginals' are generally settled at the periphery of main cities in Latin America (see Nun 1969, 1999; Murmis 1969).

[197] Germani's work is in the same line with Chicago School's 'urban ecology' approach. Besides, in the literature, there is also Marxian approach focusing on *Marginalidad*, which considers some people as marginals as they stay at the outside of the capitalist production system (for instance, see Quijano and Westwell 1983). But we don't enter into details of this second approach in this review.

[198] Marginality is related to economic system such as regulation of production and consumption, to social aspects such as political participation, participation to unions, especially absence or exclusion from decision making process for any societal issues in community, working place, public or private institutions, to national culture, and to multiple aspects of life such as family organizations, norms, values, attitudes.

The marginalidad approach has been widely influenced by dependency theory- 'periphery' vs. 'center', which has notably dominated the Latin American debate in the 1960s and 1970s. Previously, in the 1920s, Park's work titled 'Human Migration and the Marginal Man' (1928) has introduced, in American context, the concept of 'Marginal Man' referring to persons who "living and sharing intimately in the cultural life and traditions of two distinct peoples, never quite willing to break, even if he were permitted to do so, with this past and his traditions, and not quite accepted, because of racial prejudice, in the new society in which he now seeks to find a place" (1928, 892). The 'Marginal Man' emphasizes the role of racial discrimination. which is not a primary interest in Latin American marginalidad studies, but we can see that Park points out well the cultural differences that some migrant faces (see also Stonequist 1961).

Germani (1980, 49-50) defines marginality as the lack of sufficient participation of individuals and groups in various spheres of life in which, according to determined criteria, they might be expected to effectively participate in. For an effective participation, according to him, a person needs (i) "material and non-material" items (e.g., "schools to educate oneself, employment to exercise nonmarginal occupations, purchasing power to have access to the goods and services"); (ii) no obstacles for reaching opportunities and to these items; and (iii) "personal features and capabilities[199] which individuals ought to have in order to exercise the roles". As a function of these three criteria, whether a role is marginal or not is determined. According to Germani, the interactions between deprivation of material and non-material items, obstacles, and personal features and capabilities lead to different configurations of marginality. The participation, in a broader sense, means the different roles taken by individuals/groups, such as "acting vs. not to acting, giving vs. receiving, carrying out obligations, duties, rights", in different spheres of life (e.g., "institutions, spheres of individual, collective life, family life, productive subsystem, political subsystem") given "sex, age, civil status, occupation, social status, norms, values, expectations". Germani (1980, 55-64) summarizes into five *categories* the spheres of life where a person/group can participate in. And he also mentions the marginal degree and marginal form of insertion to these categories. The categories are: "(i) productive subsystem; (ii) consumption subsystem; (iii) cultural subsystem; (iv) the subsystem dealing with education and the access to the scientific literary, and artistic patrimony; and (v) political subsystem". These categories are not mutually exclusive; they distinguish aspects and not the roles, actions or concrete behaviors[200].

First, the "*degree and form of insertion in the productive subsystem"*: the marginality in the productive subsystem covers "total unemployment, partial unemployment (visible underemployment), invisible underemployment, activities on one's own account in sub-economic conditions"; second, the "*degree and form of insertion in the consumption subsystem"*: the consumption subsystem includes "economic aspect of the participation of individuals and groups in the market as consumers", and various components of standards of living such as housing, food, clothing, education, health, and recreation; third, the "*degree and form of insertion in the cultural subsystem"*: the cultural subsystem focuses on the cultural

[199] This corresponds to "psychological features", and "the cognitive patrimony" (1980, 51). Sen (1981) also uses the term *capabilities*, but in a broader sense including resources, means, opportunities, personal abilities. However, Germani only uses the term *sources* and he links capabilities to personal conditions.
[200] As the reasons behind the emergence of marginality are plural and there are more than one forms of marginality, Germani qualifies the phenomenon of marginality as multidimensional and multi-forms (1980, 55)

features of modern urban society (i.e., "those individuals who maintain behavior patterns which prevent them from assuming and accomplishing adequately the roles corresponding to the structures of the modern society would be marginal persons") , and it focuses on the cultural features of society as a whole (i.e., domination of classes, some groups, or various sub-cultural differences among classes, regions); fourth, the "*degree and form of insertion in the subsystem dealing with education and the access to the scientific literary, and artistic patrimony"*: the education related subsystem encompasses knowledge and educational attainment, and cultural formation (i.e., all people in a society should attain these items; otherwise, they are considered as marginal) ; fifth, the "*degree and form of insertion in the political subsystem"*: the political subsystem reflects the degree of participation into decision making process as a citizen, or as a resident in a country or in a territorial unit. Besides, other types of marginality are emerged as well as a result of unequal or discriminatory distribution of public services such as public health facilities, social protection scheme, security system, other public services in a country or in a territorial unit among areas, or regions (1980, 63).

Germani suggests five factors to describe and analyze the phenomenon of marginality and its plural forms. The five factors of marginality are as follows: First, *economic-social factors*: marginality is generated by economic-social order (i.e., structural conditions, and its related policies, or conjunctural situations within such order). For instance, unemployment, underemployment, disguised unemployment can cause marginality, which are in turn induce a subculture characterized by a different level of living standards, of participation in-decision making, of the access to goods and services, compared to the rest of population. According to him, marginality can slow down economic development and as a result marginality can perpetuate (e.g., marginal situation of some regions in a country)[201] (1980, 10-13). Second, *political factors*: political regime, or political orientations of the government may contribute to the emergence of marginality, and its perpetuation, especially limitation of political participation as one of the features of marginality. Germani (1980, 14) emphasizes also stratification system and discrimination such as the 'racial' US discrimination (see underclass sub-section), but in many Latin American countries it is a cultural matter. Third, *demographic factors*: Germani (1980, 15) recognizes that at the time he was writing his book, Latin America's population growth was the highest one in history compared to developed and the third world countries. However, he mentions that population growth was not the primary cause of low economic growth but in fact it was inhibiting reduction of marginality (because in a such situation never experienced before, any economy and any kind of social order will have difficulties to handle mass population growth). Fourth, *cultural (in the anthropological sense) factors*: Germani (1980, 15-16) sees marginality as a result of "the social order imposed by the colonial society", more precisely, "the domination of the European or Europeanized minority over the great majority, mainly pre-Hispanic populations and/or former black slaves." This factor would be the main driver of "the superposition of cultures and civilizations" such as "duality of values, social structures, and political and administrative regimes." (DESAL, Chapter 1). According to him, this situation stimulated the gap between urban and rural areas, and differences among regions in terms of development in Latin-American countries[202]. Fifth,

[201] Germani has attempted to explain the relationship between underdevelopment and marginality for capitalist economies. He suggests for further research a similar analysis for non-capitalist economies (1980, 12).

[202] According to Germani (1980, 16), these power connections such as domination of one cultural group as a result of colonial system is the peculiarity of Latin-America, which is clearly different than other

psychosocial factors: Germani (1980, 16-18) say that inadaptation to modern society behaviors generate marginality, which in turn may reinforce difficulties in fully participating in industrial society. As a result of marginality, "status inferiority and systematic deprivation of rights can produce an underestimation of the self by the individual, loss of identity and other psychological effects, which may contribute, despite structural changes, to the perpetuation of marginality." Germani emphasizes the substantial role of "early socialization", and "cognitive and intellectual development[203]" on the emergence and perpetuation of the nature of marginality. At this point, he also refers to Lewis's culture of poverty concept which, according to Germani, has already contributed to explain the marginality from early socialization, psychosocial and structural aspects.

To summarize, according to Germani's marginality framework, individuals/groups are participating in the five main categories of living and they take roles in each category. For instance, an individual will have her set of marginalities (i.e., simultaneously defined), which in turn construct her own configuration of marginality in a given category. As individuals/groups may participate in more than one category, at the end, each individual/group will have her own combination of marginalities across five categories. Germani (1980, 63) distinguishes generalized marginality vs. partial form of marginality. Totally non-participating individuals/groups will be suffering of generalized marginality. Thus, there will be mainly three groups in the population: participating sector, generalized marginal sector, and partially marginalized sector. Finally, Germani (1980, 64) mentions that the nature of the socio-cultural context and the causes of marginality define dimensions and forms of marginality, and interrelations among dimensions. Besides Germani's paramount work on marginality, there are also other studies emphasizing the marginality, migrant poor in urban areas of Latin America (see Quijano 1974, Cornelius 1975, Lomnitz 1975, Perlman 1976, Eckstein 1977, Roberts 1978, Wacquant 1996).

How social relationships define capabilities of the person

In section 2.1 we proposed and showed different types of functionings including macro-level ones, which reflect the societal situation in terms of actions and activities realized by a group of people (through the interactions) having a direct or an indirect effect on the life of the person. Sen, in his framework, takes in consideration micro level functionings that are activities and actions realized by the person and points out that social interactions are also important, but they are not considered at the heart of the capability analysis. However, macro level functionings are very important because first of all people do not live alone and have different types of relationships with her entourage including close family ties, or friendships, or colleagues, or other people they encounter on their daily basis. Nowadays, even different sorts of media as one of the tools of transmitting information are the substantial elements of social interactions. Unless we don't follow closely the daily life of the person it becomes difficult to capture her interactions with society (including her behaviors in public, or with her close entourage and the behaviors of others the person met on her daily basis-which designates both the relationship of

typical inequalities driven by class structure. But of course, similar situations occurred in other countries such as in the US, South Africa.

[203] Germani (1980, 18) writes "it is also known that cognitive and intellectual development is influenced not only by social, interpersonal, family and environmental conditions but also by physical deprivation, especially nutritional, which destroys from a very early age all possibility of intellectual development in the child and annihilates his possible recuperation for a full participation in adult life."

the person with society in a broad sense and the relationship of society with the person. The person does not only act in her relationships but also give reactions to the acts of others). Social bonds inevitably constitute one of the fundamental elements of capabilities of the person. We cannot really know in which direction or how exactly the person's social bonds define her capabilities set unless we don't have exact information on her daily activities but via proxy indicators, we can obtain information on her family life, community life, labour life, social life. For instance, a single mother may need the help of her family to take care of her children while she is at work. Or a cleaning worker from poor community going to affluent neighborhood will encounter behaviors of upper class and such encounter will definitely create some ideas and observations in her mind.

For defining capabilities, we used the term of functionings to mention activities and actions that the person process and different doings and beings that the person got or achieved in her life at a given time. When the person values something she would like to achieve then she will take some decisions in her mind to define the way she would process to achieve what she values. In this process: (i) she will definitely encounter acts of others; (ii) she will have to deal with her actual situation including the state of her household, family, workplace, community, region, country, and world in a broad sense. Thus, at a given time before she processes what she values there is already her environment (i.e., her life situation or living conditions at the time before proceeding), which is a result of perpetual social interactions including external factors she cannot control, or she cannot decide herself alone. But also, during the processing, she will face the future conditions of her environment including social factors she cannot control, determine or predict. Assuming that life is a dynamic process and acts of every single individual play a role in determining the situation of society, social relationships or in other words how the person interacts with her environment becomes the essential element of capabilities. In the light of these explanations, we point out that the macro level functionings we determined in our typology of functionings correspond to different types of social relationships of the person with her entourage on her daily basis and to the situation of her environment where she lives including for instance her community, neighborhood, household, family. The mixed macro-micro level functionings we defined in our typology (see section 2.1) contain also social bonds given that these functionings reflect actions and activities the person has faced due to her social environment and the actions and activities the person does herself for others in her social environment.

In practice, depending on the structure of questionnaire, we can detect macro and mixed macro-micro level functionings. To concretize these explanations, we will give two examples. First example is about the daily life a housekeeper named Sharon. Suppose that she has one-year old baby and in weekdays she has to go to a rich household for cleaning the house, for cooking and taking care of children of her employer. She wakes up early around 5 o'clock in the morning in the weekdays, she prepares breakfast for her family and she does some household chores before leaving her place. Then, she takes every morning a very crowded subway for about 1 hour and in-between she has to drop in her child to her mother and she finishes her work at around 6 o'clock in the evening and when she comes at home, she has to prepare dinner and to do some household chores. If she can find a place to sit down in the subway, she feels less tired when she gets home but there is no guarantee for a free seat. But also, depending on the tasks given to her in every morning from her employer she organizes her job. She earns the equivalent of minimum income settled in the country and for increasing her household's income (given that her partner is unemployed for 6 months and if her partner cannot find a job in coming 3 months

then they won't receive any unemployment insurance and they will be asking for food help from her community) she would like to become an administrative worker in municipality. So, she has been preparing for the exam during the evenings after family dinner. This is a very short portray of Sharon, but this will help us to understand macro level functionings. First of all, her family ties define the way her baby is taken care in the day. Second, the situation of subway in terms of number of persons taking the subway, the number of available seats in the subway and the frequency of subways play crucial roles in terms of time she takes for going to work and for coming back to home. Third, the size of her household and dynamics (the roles each person takes) within the household define her household tasks. Fourth, Sharon fulfills her job as a function of her employer's schedule and requests. Fifth, her household income depends on unemployment situation of her partner and this depends also on the labour market conditions. Sixth, for additional income, she will be probably asking for some food help. The food help she will be receiving will be dependent on available food help types in her community administration. In each of these six conditions, we see a different sort of relationship that Sharon has with her entourage. The nature of these relationships defines the daily activities and actions of Sharon. The first condition corresponds to family relationship, the second reflects public transport situation, the third emphasizes relationships within the household, the fourth designates work environment and working conditions, the fifth reflects the impact of labour market conditions on her household, the sixth concerns one type of social welfare programme in her community. In none of these six cases, Sharon decides herself alone and as we can see her daily capabilities depend on different sorts of social relationships she has with her family, household, community, workplace, society where she lives. Notably, labour market, public transport situations are the result of complex social interactions, which are not direct social ties of Sharon as in the case of her relationship with her family, her partner or her employer. Through these six conditions, we can see that on the one hand Sharon realizes some actions and activities for other(s) (i.e., mixed macro-micro level functionings), but also others realize some actions and activities for Sharon (i.e., mixed macro-micro level functionings) and there are some macro level functionings such as labour market, public transport conditions as the result of public's behaviors (i.e., people do not specifically act for Sharon). As a micro level functioning, Sharon would like to achieve her exam. But the way she will be preparing for the exam depends on these macro level, mixed macro-level functionings. In other words, Sharon's capabilities set for achieving her exam includes her direct or indirect relationships with different individuals in her society. The interest in whether Sharon achieved or not her exam is important but to ameliorate living conditions of Sharon we need to know her profile and to understand how she lives. As we can see from this example, economic resources do not define alone Sharon's capabilities but in fact different social bonds Sharon shares with society shape significantly her capabilities. Thus, social interactions or social relationships should not be seen as the indirect tool of capabilities (nevertheless, Sen's framework considers social context as indirect factor on capabilities) and in fact these social factors should be directly considered in discovering the capabilities of people. In terms of poverty, Sharon witness or risks at experiencing some disadvantages regarding crowded public transport, household chores early morning and late evening, long hours of commuting between work and home, household income, unemployed partner, dependence on social assistance. These doings and beings Sharon has in her life at a given time are not (necessarily) the result of personal failure or they are not personal non-achievements; in fact, these functionings are the results of her interaction with her society in her daily life. These interactions including the activities, actions the others do, and the states of the others bring her some disadvantages that decrease Sharon's capabilities. Sharon

is deprived of decent public transports, decent commuting time, adequate household income, liberty of consuming without being dependent on social assistance. We don't have here a full portray of Sharon but of course she experiences also non-deprivations in terms of personal achievements (i.e., micro level functionings) and advantages (i.e., macro and mixed macro-micro level functionings such as receiving help from her family for taking care of her child). The combinations of deprivations and non-deprivations construct her profile and therefore inform about her capability set at a given time. As a summary we can say that Sharon interacts with people (i.e., with society in a broad sense) on her daily basis; as a function of relationships, she ties with others she realizes various actions and activities; she finds herself in different conditions/states (e.g., she thinks, observes, gives reactions, acts, takes decision); and she realizes various actions and activities. In the end, the combinations of the outcomes of all these functionings tell about the nature and extent of real liberty she has for leading the life she values, and she has right to value. Understanding types and dynamics of her relationships with society where she lives is the primary step for discovering her capabilities. Capabilities are not about only the deprivations and non-deprivations but most importantly capabilities take shape on the basis of social bonds. Economic results such as financial deprivation, household income are also configurated by social bonds. With this example, we showed that taking only economic perspective such as household income, or consumption expenditure pattern, or material deprivation and ignoring social interactions (or considering social interactions as indirect factors of capabilities) is the monumental error in evaluating quality-of-life of people.

Second example is about a single mother witnessing hardship during the pandemic. Kylie is a single mother with 2 children living dependent on government's social welfare programme receiving some supplementary income each month. Her children are aged 1 and 3. She is working as a costumer adviser at a travel agency and her monthly income is just 1.5 times higher than the minimum income. She is not able to afford nanny for her two children and for that reason she is only working half day and her elderly neighbor takes care of her children when she is at work. Suddenly, a pandemic arises, and her employer has to close the travel agency for about 6 months. So, for 6 months, Kylie becomes stringently dependent on government's help that is equivalent of minimum income. As her monthly income diminishes so she has to cut off some of her expenses and has to renounce some of her habits such as buying make up products, paying private tv channel for movies, and cooking a lemon cake every single day. And, when the pandemic surfaced and restrictions announced by the government, she could not afford to buy some additional food and cleaning products as stock. Before the pandemic, she planned to get credit from the bank to buy her own house. However, as she suddenly found herself not economically active for 6 months, then the bank refused to give her the credit for 1 year. Based on this short portray of Kylie, we can capture different sorts of social bonds she ties with her social environment including the country situation because of the pandemic. We can see that she has close relationship with her neighbor who takes care of Kylie's children without receiving any money. She has relationship with social welfare assistant as she receives some monetary help each month. This relationship indicates the dependency of Kylie on government's help for affording her monthly expenses. In weekdays, she encounters costumers at her work, and she has contact with her colleagues during the half day. The first two relationships correspond to mixed macro-micro level functionings in the sense that others do something for Kylie. The last one includes intrinsically interactions Kylie cannot exactly predict; for instance, she does not know in advance the demands of her clients and the relationship at office has different dynamics depending on acts of her colleagues. So that,

Kylie's work environment in an ensemble corresponds to a macro level functioning. The pandemic as a macro level functioning affects clearly the life of Kylie as she suddenly becomes not able to work for 6 months. The fact she could not do stocks for pandemic because of her financial situation makes her feel lonely and insecure. She also finds herself not able to get loan from the bank, which she considers a failure. If we evaluate why Kylie could not have achieved to get loan from the bank, we have to look at her capabilities set. Her capabilities set includes a diminishment in her monthly income but most importantly contains different sorts of social relationships she has with society. Her relationship with the pandemic is indirect, her relationship with her neighbor is more direct but they all contribute to shape Kylie's capabilities. She is being refused by the bank because of her financial situation but most importantly she is being refused due to unpredictable consequences of the social phenomenon 'the pandemic'. If we look only at the fact Kylie could not achieve to get loan from the bank and she does not have enough money for buying her a house, we will significantly miss the entire profile of Kylie. For improving Kylie's situation, it is important to look at her social relationships; for instance, the vulnerable situation of Kylie when such pandemic appears reveals her precarious working conditions. Kylie needs first of all social protection and recognition from the society. She does not need only a higher salary for each month to ensure her expenses, but on top of that she needs to be recognized as a citizen. We can see that Kylie is not only poor regarding income, but she is also poor in terms of several disadvantages she gets because of pandemic effect (e.g., no work, reduction of income, no credit from the bank). On the one hand, dependence on social assistance might be an advantage for Kylie during the pandemic in the sense that short period social assistance helps Kylie to feel socially recognized and makes her feel the solidarity in society. On the other hand, such dependence will decrease capabilities of Kylie in the long-term if the assistance becomes the only way of affording her needs. Social protection is important but when assistance becomes the social status of the person (i.e., socially recognized as the 'assisted') then Kylie's real liberty will be in peril. Thus, information on the type of relationship Kylie has with social assistance becomes crucial to detect and predict Kylie's future capabilities. As we mentioned in previous example, social bonds stand as the principal pillar of capabilities.

As we have seen Townsend's relative deprivation and Mack and Lansley's Poor Britain (based on socially perceived necessities) studies include the item of social life in analyzing poverty in the United Kingdom. However, it is important to point out that social life is not about only having dinner with friends, or with family or with colleagues, or having holidays, or inviting people for having drink or having dinner. Current European multiple deprivation and social exclusion studies include this type of indicators to reflect social life. Similarly, such indicators are also used in some multidimensional poverty analysis. However, social bonds as a concept is beyond such indicators. As we have seen from the examples of Sharon and Kylie above, social bonds emerge in many aspects of life on daily basis such as relationships within family, household, workplace, university, public places, public transports, institutions, etc. This is why we argue that social life should not be seen only by such indicators. Thus, we can say that the 'social bonds' is not a particular dimension of poverty itself alone but in fact social bonds is the key element of any dimensions of poverty. For instance, economic activity can be seen as a dimension of capabilities and social life in terms of having some activities with friends, colleagues, family members, community members can be seen also as a dimension of capabilities, but the most important point is that in both two dimensions there are different sorts of social bonds.

Social exclusion is generally defined as not being able to participate in economic, social, political activities in society; for instance, not being able to afford 'socially perceived necessities' defined by the majority in society (see Mack and Lansley 1985). As we have seen from the literature review, people may feel or find themselves outside of the mainstream because they are not able to achieve or realize socially perceived necessities. Many reasons are possible for such exclusion, but one thing is clear that when they do not have opportunities, means and sources for achieving what they value then this situation corresponds to the lack of capabilities, the lack of real liberty. From this perspective, social exclusion can be seen as a sort of multidimensional poverty. To understand from which aspect(s) the person feels socially excluded, we have to directly ask the person and get her profile via subjective indicators on social exclusion. However, it is important to note that not everyone should accomplish the socially perceived necessities or in other words not everyone should do the actions, activities that the majority of society does. People cannot be considered marginalized because they refuse to apply customs, traditions, or to obtain the items that the majority qualifies as the must have ones. The key element to differentiate the real socially excluded people is that we have to get information about whether the person values and would like to have or obtain the functionings that the majority of society values. Another important issue is the link between human rights and the socially perceived necessities. In developing countries where democracy and fulfillment of human rights are problematic and some traditions and customs are quite prevalent, it is extremely important to check whether the socially perceived necessities are not against the human rights. For instance, in some countries forced child marriage is perceived as a normal event. In such condition, we cannot say if a child is not married before 18 yrs. then she is socially excluded. In societies where people do not know about democracy, human rights, how they can search for their rights, how they can defend themselves, the measurement of poverty via socially perceived necessities becomes problematic. For increasing capabilities of every single human being, the fulfillment of human rights is the most relevant, absolute functioning.

Being not able to participate effectively in various activities happening in society reflects weak relationships of the person with society; however, types of social relationships are generally not described or explained in detail in traditional quantitative multidimensional poverty and social exclusion or in relative deprivation and social exclusion studies. At this point, Paugam's (1991a, 2005) social disqualification framework connects social bonds and social exclusion to explicate different types of poverty. As we have seen he has defined 3 main types of poverty (integrated poverty, marginal poverty, disqualified poverty) based on labour market conditions, social bonds and social protection system. For instance, in integrated poverty, (i) economic development is weak; (ii) there are informal economy and underemployment; (iii) family relationships are strong; (iv) there is strong protection via friendship, community, family relations; (v) social security programmes are not relevant or do not exist; (vi) no guarantee of minimum income (Paugam 2005, 89). Here we can say that a person living in such poverty experiences disadvantages due to economic conditions of society, and to weak relationship between the state and citizens in terms of social protection. On the other hand, the person has advantages from her strong family, friends, community relationships. According to our typology of functionings, we can say that these 3 drivers of integrated poverty contain macro-level and mixed macro-micro level functionings. The nature of these functionings show the capabilities of people in integrated poverty situation. In other words, Paugam shows the profiles of integrated poverty, marginal poverty, and disqualified poverty situations by combining functionings of 3 dimensions that are labour market, social bonds, and social protection system.

His view is very close to our capability analysis in terms of profiles and different configurations of poverty. In our multidimensional poverty analysis, we are interested in extraction of combinations of functionings; however, we will have a great number of indicators. He used the term 'social bonds' for mentioning mainly family, friend, community relationships but the idea of social bonds can be extended including all other sorts of relationships the person has with her society. For instance, unemployment situation has a direct or indirect effect on person's capabilities, or the structure of social assistance programme defines the possible relationship of the person with social assistance service. Thus, we can qualify that these 3 dimensions in Paugam's analysis correspond to economic relationships; family-friend-community relationships; social assistance relationship, respectively. When we combine these different relationships then we obtain *social attachment profile* of the person. As we have mentioned we suppose that the person does not live alone and encounters perpetual interactions with her environment on daily basis; so, the way she is connected with her society and the way the society is connected with her (i.e., which is the social attachment) reflects her existence/state, actions, activities, advantages and disadvantages. In other words, social attachment profile of the person shows the nature and extent of capabilities of the person.

2.4 Multidimensional poverty

As we have mentioned there are different approaches to conceptualize and measure poverty. Except monetary approach, other do include a wide range of social aspects of life[204]. In the literature, there is a clear distinction between monetary approach and non-monetary approach. In monetary aspect, poverty is simply defined by one element 'money' as an indication of not having a sufficient amount of income or of not reaching a minimum level of adequate subsistence due to lack of sufficient earnings. On the other hand, in multidimensional aspect (e.g., capabilities approach), poverty is perceived based on various components of life such as access to educational institutions, access to health services, having an adequate income, living in an adequate shelter, having clothes and other materials. In the literature, especially at international platform among the United Nations' agencies, it is widely accepted that money cannot be considered as the only one indicator of poverty. In many cases, it is not possible to improve the given condition by paying more or by buying a supplement. For instance, people living in big cities such as Ulaanbaatar in Mongolia suffer significantly the strong negative impact of pollution invading whole city due to recent climate change. Or, some minority groups in the world suffer the consequences of non-democratic behavior of governments such as Uygur people living at the north west of China. In recent years, due to globalization, the circulation of goods has relatively become easier; however, the free circulation of human beings (notably for the citizens of developing countries) is still significantly limited. Many other examples can also show the limit of the power of money for having freedom. We can see that actually real liberty of the majority of world population to achieve diverse functionings is very low. In this context, it becomes substantial to investigate in which life conditions people experience the lack of real liberty leading them to live in various forms of poverty. Compared to purely monetary vision, the multidimensional one is intrinsically associated with a sociological eye. The multidimensional aspect primarily focuses on human beings' living conditions without trying to explain diverse deprivations based on the lack of money; but rather it includes different types

[204] However, we know that in some cases such as basic needs, or absolute poverty, even if more than one aspect is included to define poverty, in the measurement process these aspects were reduced to one or they were converted to monetary measure.

of relationship that the person has with her/his entourage (e.g., social environment), and it sheds light on various social constructs (e.g., access to public health services, education, community habits, political structure). Notably, capabilities approach aims at giving insight into the society structure in terms of human rights, available opportunities, democracy and how social groups differ regarding these aspects.

In non-monetary approach, the tendency is to include various aspects of life to define poverty. For instance, Atkinson et al. (2002, 79) mention that "poverty is multidimensional because it does not consist merely of an insufficiency of resources, but also encompasses cumulative deprivation in relation to income, housing, education, and health care. In other words, it concerns non-participation in various important areas of life". Ellis (1984) points out that there are several dimensions of poverty such as economic poverty, social poverty, political poverty, legal poverty. In this light, if more than one dimension (each dimension can be seen as an aspect) is taken in consideration, this process is called multidimensional poverty measurement. Related to that, the first concise attempt in the literature has made by Sen (1985, 1992, 1999) based on his capabilities approach. Before Sen's capabilities, poverty was measured for instance, regarding basic needs, subsistence (including shelter, clothing, nutrition), underclass, marginality, or social exclusion where several social aspects notably social participation was emphasized; but these methods did not lead to a visible international movement to multidimensional measurement (i.e., the operationalization of multidimensional definition of poverty came first prominently out by Sen's capabilities approach). Especially, in the case of social exclusion, marginality, and underclass studies, more often they were carried out by qualitative methods (e.g., in-depth interviews, participant observations). However, current multidimensional poverty studies are almost quantitative, based on dashboards (e.g., Organisation for Economic Co-operation and Development (OECD onwards)'s Better Life indicators board, United Nation's Millennium Development Goals and Sustainable Development Goals indicators board, see also Alkire and Robles 2016 for dashboard approach), or on various statistical techniques such as Alkire and Foster's counting approach (2007, rev. 2009b, 2011a) (e.g., United Nation's Multidimensional Poverty Index (MPI)), composite indices[205] (e.g., Human Poverty Index (HPI)), or on multivariate statistical tools (see Kakwani and Silber (eds.) 2008, Alkire et al. 2015 for various applications).

Alkire and Santos (2009, 142) mention that multidimensional measures have become in demand especially by three factors: "national poverty measures, identify beneficiaries for targeted public programmes, evaluate progress." Regarding them, the idea of new multidimensional measures for better capturing multiple deprivations targeted to the poor has more and more attracted the interest of national governments, compared to income measures. The fact that poverty is beyond income deprivation is significantly recognized by international organizations such as the World Bank, OECD, United Nations. Most recently, the United Nations (UN onwards), in the context of Sustainable Development Goals (SDGs), has launched 'My World Survey' on the internet where people can select and evaluate the most immediate important 6 goals of SDGs to them and to their family, among the all goals (reduce inequalities; industry, innovation and infrastructure; decent work and economic growth; partnerships for the goals; life below water; zero hunger; responsible consumption and production; peace, justice and strong institutions; affordable and clean energy; gender equality; sustainable cities and

[205] See Mazziotta and Pareto (2017), OECD (2008b), Nardo et al. (2005) for more technical information on the construction of composite indices.

communities; quality education; no poverty; climate action; good health and well-being; clear water and sanitation; life on land)[206]. Besides the UN, for instance, the OECD 2012 report has emphasized the non-monetary side of poverty such as freedom, security, social inclusion, strong institutions, education, health.

Alkire and Sumner (2013) point out that the World Bank's famous monetary poverty indicator USD 1.25 per day is not sufficient to capture the extent of poverty because many people are not only deprived in financial aspect, but they do also in other dimensions including malnutrition, education, health and sanitary facilities, electricity, water. In their 2000 World Development Report (on 'Attacking Poverty'), the World Bank has mentioned three axes of reducing poverty in the world: promoting opportunities, facilitating empowerment, enhancing security (see World Bank 2001, vi). In following paragraphs, first we will shortly see some recommendations given by Alkire and Santos (2009), and other scholars for constructing a multidimensional measure of poverty (or of well-being), second, we will describe the two widely recognized international multidimensional poverty measures, namely the Human Poverty Index (HPI) and the Multidimensional Poverty Index (MPI), created by the United Nations Development Programme (UNDP onwards) based on Sen's capabilities approach. After that, we will see some poverty dashboards existing in the literature. Then, we will continue with several examples showing various types of application of multidimensional poverty measurement in the literature, carried out by using different quantitative tools.

Alkire and Santos (2009, 143-147) specified seven steps for constructing a multidimensional measure, which are respectively: determining a unit of analysis (e.g., household, individual, family, country, region, community, city)[207], order of aggregation (e.g., obtaining an overall poverty rate based on an aggregation function such as average, sum), dimensions, variables within the dimension, poverty lines, identification criterion, weight (e.g., relative importance assigned to an indicator, a dimension, or a specific group of people, or an individual such as the most disadvantaged, in order to increase/decrease her importance in the computation process). Moreover, Chiappero-Martinetti (2000, 5) defines 5 steps to be taken into account for an empirical application of capabilities, such as measuring multidimensional well-being or poverty. These steps are: "(i) the adequate evaluative space: capability vs. (achieved) functioning; (ii) a list of (essential, relevant) capabilities or functionings; (iii) a set of indicators related to the selected dimensions of well-being and adequate criteria to measure and represent them; (iv) how (and if) to aggregate the elementary indicators to obtain an overall evaluation for each single dimension (functioning/capability) of well-being; (v) how (and if) to add up all the dimensions and to reach an overall evaluation of well-being".

The selection of dimensions depends on the goal and context of study. For instance, hunger, malnutrition, sanitation, health facilities, education services are generally most important in the case of developing countries for measuring poverty; however, in developed countries other type of dimensions such as cannot got to holiday, or not having an appropriate social relationship with family, or not being able to participate to social activities are acknowledged as important poverty indicators. But also, cultural, environmental contexts may play a role in determining dimensions; for instance, in livestock and agriculture-based economies such as in Mongolia, vulnerability of people to natural disasters and climate change have become substantial risks to

[206] See the following link for 'My World Survey': https://myworld2030.org/. Accessed 30 June 2018.
[207] Most of time, household is chosen as a unit of analysis such as in the UNDP's poverty measures.

trapping in poverty. From technical side, data availability and data comparison are crucial elements for the selection of dimensions. Alkire (2008a, 97) mentions five drivers for selecting dimensions: "existing data or convention, theory, public consensus, ongoing deliberative participatory processes, empirical evidence regarding people's values".

After having chosen dimensions, then indicators/variables are defined within each dimension[208]. This step depends on theoretical framework, what the researcher would like to measure by dimension. For instance, if the theoretical framework is capabilities, indicators should address functionings. The relationship between indicators should be also taken in consideration for avoiding overlapping in a measurement process. Regarding poverty lines (or poverty thresholds), and identification criterion, they serve to define who is poor; specifically, people having a poverty score worse than the value of the poverty line are considered poor. Poverty lines are defined based on experts' opinions or national context, or international consensus. An objective process for setting poverty lines will be the best option, but mostly they are defined by value judgment.

Besides Alkire and Santos' suggestions mentioned above, there are also various recommendations drawn by other scholars such as Fitoussi et al. (2002), Alkire and Foster (2007, rev. 2009b, 2011a), Atkinson and Marlier (2010b), and the World Bank (2017). Atkinson and Marlier (2010a, 45) have recommended five criteria for an optimum international comparison of indicators[209]: "(i) an indicator should identify the essence of the problem and have an agreed normative interpretation; (ii) an indicator should be robust and statistically validated; (iii) an indicator should be interpretable in an international context; (iv) an indicator should reflect the direction of change and be susceptible to revision as improved methods become available; (v) the measurement of an indicator should not impose too large a burden on countries, on enterprises, or on citizens. Authors also mention that the set of (social) indicators should respect three following criteria: (i) an equal number of indicators in each dimension; (ii) indicators should be proportionally weighted; (iii) transparent and accessible data source information to citizens. Alkire and Foster (2009a, 78) have defined six criteria to be fulfilled for dealing with complex nature of poverty measure: "be understandable and easy to describe; conform to 'common-sense' notions of poverty, be able to target the poor, track changes, and guide policy; be technically solid; be operationally viable; be easily replicable"[210].

At international area, as we have seen previously, in 1990 the United Nation's Development Programme has collaborated with the economist and philosopher Amartya Sen to build a human

[208] But depending on available data, it is also possible to set the dimensions as a function of computed indicators.

[209] Authors developed this criterion for the context of social inclusion measurement in European Union.

[210] In the literature, there are on-going debates on which criteria a multidimensional poverty measure should be based on. For policy makers, it can be important to use both conceptually and statistically easy poverty measurement tools. But 'too simplicity' generates some problems as well. Notably, dividing the society into two distinct groups such as poor vs. non-poor is vastly applied by multidimensional poverty measures and it is generally considered as understandable and accurate operation; but, this 'common-sense' of poverty in fact ignores different forms of poverty which can affect different strata in a society. Moreover, this practice makes perceive 'being poor' as a specific stratum; however, being poor is a fluid, changeable process depending on political, social, economic context of where the person maintains her life (e.g., different contexts at different levels such as country, region, community, family).

development index (HDI). This index includes three dimensions: education, health and standard of living (see Anand and Sen 1994, Jahan 2003, Raworth and Stewart 2003). In 1997, this index was replaced by the Human Poverty Index (HPI)[211]. After that, in 2010, Alkire and Foster's counting methodology-AF onwards (see Alkire and Foster 2007, rev. 2009b, 2009, 2011a, 2011b, 2016; Alkire et al. 2011a, 2011b, 2010; Alkire and Santos 2010, 2013, 2014; Alkire and Robles 2016) has been used to create a new *global* Multidimensional Poverty Index (MPI) in line with the Millennium Development Goals, which replaced the HPI. We also note that independently from the UNDP, Alkire and Foster also try to expand the use of their methodology at international level through collaborations with scholars from different countries and national statistical agencies (e.g., Mexico, Colombia, Bhutan). In this light, we first start by exploring the HPI, then we will continue with the MPI.

The HPI is first reported in 1997 Human Development Report (UNDP 1997) with two versions, namely the HPI for developing countries (HPI-1), the HPI for developed countries (HPI-2). In these two cases, the HPI focuses on the same three dimensions as the HDI: education, health, and living standards, but the indicators differ. In the HPI-1, the 3 following indicators are used: the adult literacy rate (dimension of education), the probability at birth of not surviving to the age of 40 (dimension of health), and the average of the per centage of population without access to an improved water source and the per centage of children under weight for their age (dimension of living standards). Concerning the HPI-2, there is an additional fourth dimension that is *social exclusion*. The indicators for the HPI-2 are: the per centage of adults (15-65 yrs.) lacking functional literacy skills (dimension of education), the probability at birth of not living to the age of 60 (dimension of health), the per centage of population below the income poverty line that is defined as 50 per cent of the median household income (dimension of living standards), and the rate of long-term unemployment for at least 12 months (dimension of social exclusion).

Concerning the MPI, which is based on AF methodology (see above), is aimed at providing the overall level of poverty (for a settled poverty line) and is especially designed for measuring *acute* poverty[212] in cross-county comparison perspective. It includes the same three dimensions as the HDI, but with the following ten indicators*: the years of schooling*, and *the children enrolled* (in the dimension of education), *nutrition*, and *child mortality* (in the dimension of health), *cooking fuel*, *toilet*, *water*, *electricity*, *floor*, and *assets* (in the dimension of living standards). In fact, there are four types of MPI in practice: (i) the global MPI carried out by the UNDP and OPHI (Oxford Poverty and Human Development Initiative onwards) for annual

[211] The HDI was replaced by the HPI in poverty measurement context. However, for human development purpose, the HDI still continues to be used in the UNDP's annual human development report.

[212] As this *global* poverty index focuses on acute poverty, this is why her dimensions are limited to some basic items of education, health and living standards. In some other regional, national adaptations and various academic applications, there are more dimensions compared to the *global* MPI. But also, for making easier and enable international comparison of *acute* poverty among developing countries, a limited number of dimensions and indicators is purposely selected by the authors in collaboration with experts, in 2010. The ten indicators in the *global* MPI were comparable over 100 countries. All the necessary information for computing the *global* MPI for each county should come from the same survey, except a small number of country-specific surveys and a slight difference in survey dates (the latter depends on the latest available reliable data but the authors of the *global* MPI accept that the survey date differences limit cross-country comparisons). The Demographic and health Survey (DHS), the Multiple Indicators Cluster Survey (MICS) and the World Health Survey (WHS) are the three main surveys used in the *global* MPI (see Alkire and Santos 2010).

human development reports, which is comparable across countries (see Alkire et al. 2016, Dotter and Klasen 2014); (ii) the regional MPI carried out by regional institutions such as the Economic Community of Latin America and the Caribbean (ECLAC), for instance see ECLAC's Latin American MPI published in 2014 Social Panorama, or European Union MPI (Alkire and Apablaza 2016, Weziak-Bialowolska and Dijkstra 2014); (iii) the national MPI carried out by national statistical offices, based on the same methodology of the global MPI but the indicators in national MPI may vary as a function of country context (therefore, it is not very possible to compare across countries); and (iv) the academic MPI developed by scholars by using the AF methodology with some possible modifications in the selection of indicators for the country context. Regarding third group, Mexico (Foster 2010, CONEVAL 2010), Colombia (Angulo et al. 2016), and South Africa (Statistics South Africa 2014) can be considered as the pioneers in the adaptation of the global MPI in national contexts. For instance, in 2009, Mexico has adopted their MPI as the only national poverty statistics. Besides, there are also other countries such as Bhutan, Chile, Ecuador, Panama having their official MPIs as complementary statistics to their official monetary poverty statistics. Most recently, the United Nations Economic Commission for Europe-UNECE (2017) has recommended to use the MPI methodology for monitoring the SDG target 1.2. In following paragraphs, first of all we will focus on the *global* MPI, then we will shortly see the European Union regional adaptation, the German MPI developed by a German research institution. At the end, the Colombian, Mexican, South African multidimensional poverty indices will be examined.

In this paragraph, we will introduce some technical information on the construction of the *global* MPI-which we will refer MPI here, (see Alkire and Santos 2010, Alkire et al. 2011b, 2015 for the full list of mathematical formulas), as follows:

- Firstly, we need to decide the unit of analysis that is the household here. Indicators are constructed based on household information[213].
- Secondly, poverty dimensions and indicators are determined (i.e., a multidimensional measure should contain at least two dimensions. Within a dimension, there is one or more indicators). In the context of the MPI, the UNDP and OPHI have already settled the dimensions and indicators, which are suggested to be the same for every country to facilitate cross-country comparisons.
- Thirdly, poverty lines (or poverty thresholds) are determined to identify which households are poor[214]. In the MPI, this step is based on dual cut-off process[215] (i.e., intermediate threshold): first cut-off (i.e., *deprivation threshold*), second cut-off

[213] The survey questions used for building the indicators of the MPI come often from household questionnaire. Regarding this point, it is important to take in consideration intra-household disparities/inequalities, which is in practice neglected in computing the MPI. The authors of the MPI accept this gap and link it to the data constraints as the individual-level information for all indicators in the MPI is not fully available among internationally comparable surveys.

[214] People remaining under poverty line are considered poor by definition.

[215] The AF's dual cut-off system is considered as an *intermediate* approach to identify the poor, as opposite to the *union* and *intersection* approaches. In terms of the *union* identification, a person/household should be deprived in at least one indicator to be considered poor. On the other hand, in terms of the *intersection* identification, a person/household should concurrently be deprived in all indicators to be considered poor (see Atkinson 2003, Alkire et al. 2015). Generally, the number of poor people in a given population based on the union approach is likely to be high and those based on the intersection approach is likely to be low.

(*poverty threshold* denoted as k). The deprivation threshold serves for finding which household is deprived in which indicator. In a normative[216] framework, a deprivation threshold indicates a minimum necessary level of achievement for not being considered deprived regarding the indicator. If the household has a worser score compared to the deprivation threshold of the indicator, she is considered deprived in the indicator. For instance, if the household is deprived regarding the education indicator, the household takes the score one as deprivation score in this indicator. But, if the household is not deprived regarding the education indicator, then the household takes the score zero as deprivation score in this indicator (i.e., *household's deprivation score* in education is zero or one), see the table below for an example:

Table 8. An example of deprivation threshold and household score, source: Author

Indicator **(col. I)**	**Deprivation threshold** **(col. II)**	**Household score in the indicator** **(col. III)**	**Result** **(col. IV)**	**Deprivation score of the household in educational attainment** **(col. V)**
Educational attainment of the girls in the household	Household will be considered deprived in educational attainment indicator if at least one girl in the household does not have access to school	All girls in the household have access to school	Household's score in educational attainment (col. III) is better than the deprivation threshold (col. II), so the household is not deprived in the indicator	Household's deprivation score in educational attainment is equal to zero because her performance in educational attainment (col. III) is above/better than the deprivation threshold (col. II)

Once we got household deprivation information across all indicators, we can compute the *household total deprivation score* which is the sum of the *weighted* deprivation scores of the household. Once, the household performance regarding each indicator is obtained, then household's deprivation score in each indicator is multiplied by the related weight. In the MPI, equally importance process is applied over all indicators (i.e., nested weighting approach – all dimensions are weighted equally and the indicators within each dimension are weighted equally). This process is decided by the UNDP and OPHI).

[216] Arbitrarily designated or based on expert's opinions, personal judgement or general societal judgement or societal opinion.

- Fourthly, the household total deprivation score is compared with the poverty threshold which is set at 33.3% by the UNDP and OPHI[217]. If the household total deprivation score is equal or superior than $k = 33.3\%$, then the household is considered poor.
- Fifthly, individuals living in the poor households are considered *multidimensionally poor* and they are aggregated[218]/summed up to obtain the poverty headcount ratio (i.e., the proportion of multidimensionally poor people in the population).
- Sixthly, the intensity of poverty is computed (by considering only the poor). The intensity of poverty corresponds to the division of the sum of all poor households' total deprivation score by the total number of poor people.
- At the final step, the headcount ratio is adjusted by the intensity of poverty to obtain the multidimensional poverty index score (which is equal the multiplication of the headcount ratio by the intensity of poverty). Moreover, (i) if the household has a total deprivation score equal or superior than 20%, she is considered to be near multidimensional poverty (or 2011 Human Development Report qualifies the households having a total deprivation score between 20 and 33 percent as *vulnerable to poverty*; (ii) if the household has a total deprivation score equal or superior than 50%, she is considered severely multidimensionally poor (the thresholds are arbitrarily settled by the UNDP and OPHI. Similarly, Alkire and Seth (2016) identified people in destitution (i.e., the bottom level among the poor) by using different cut-offs than the usual ones, for measuring multidimensional poverty in Nepal in a dynamic perspective[219]. Such study, including evolution of multidimensional poverty over time, captures best chronic poverty).

To make more concrete the process of computation of the MPI, we will give two hypothetical examples. In first example, we will focus on 3 individuals: Ind1, Ind2, Ind3 (each individual is affiliated to a household H1, H2, H3 respectively). Deprivation matrix represented in the table below summarizes the information.

[217] The *poverty threshold* (or poverty cut-off) corresponds to a specific percentage of total deprivation score a household must have for being considered poor.

[218] *Aggregation* means here the passage of poverty information from household to societal/community level. After having identified which household is poor, we identified which individual is poor (i.e., individuals living in the poor household), at the end we sum all poor individuals to obtain the proportion of poor individuals in the society (i.e., headcount ratio of poverty).

[219] More specifically, Alkire and Seth (2016, 1) employ the term "stringent" for the multidimensionally *poorest* group among the multidimensionally poor. The rest of the multidimensionally poor, therefore, is called "moderately poor". Their method (2016, 13-14) consists of 'stringent' poverty cut-off and 'ultra' deprivation cut-offs, compared to the MPI's standard thresholds. They propose two possibilities: (i) increase poverty cut-off from 33% to 50% or (ii) increase deprivation cut-offs regarding indicators (e.g., for water indicator, the global deprivation cut-off is "if the household does not have access to safe drinking water or safe water is more than a 30 minute walk, round trip; 0 otherwise" but their 'ultra' deprivation cut-off is "if the household does not have access to safe drinking water or safe water is more than a 45-minute walk, round trip; 0 otherwise)".

Table 8. Example 1 of the computation of the MPI, source: Author

	Education (1/3)		**Health (1/3)**		**Living Standards (1/3)**					
	ED1 (0.166)	ED2 (0.166)	H1 (0.166)	H2 (0.166)	LS1 (0.055)	LS2 (0.055)	LS3 (0.055)	LS4 (0.055)	LS5 (0.055)	LS6 (0.055)
Ind1	1	0	0	0	1	1	0	0	1	1
Ind2	1	1	1	1	1	0	0	1	0	1
Ind3	1	0	0	1	1	0	0	0	0	1

Note: The weights are mentioned in parenthesis.

As we can see from the table above, there are 3 dimensions: education, health, living standards. There are 2 indicators in education (ED1, ED2), 2 indicators in health (H1, H2), 6 indicators in living standards (LS1, LS2, LS3, LS4, LS5, LS6). Equal weight process (i.e., nested weighting) is applied across indicators according to the MPI methodology: at first step, each dimension is equally weighted (i.e., as there are 3 dimensions, each dimension has a weight of $\frac{1}{3} = 0.33$ *or* 33%); at second step each indicator within the dimension is equally weighted (i.e., as there are 2 indicators in education, each indicator within this dimension ED1, ED2 has a weight of $\frac{1}{3} \div 2 = 0.166$ *or* 16.7%; as there are 3 indicators in health, each indicator within this dimension H1, H2 has a weight of $\frac{1}{3} \div 2 = 0.166$ *or* 16.7%; as there are 6 indicators in living standards, each indicator within this dimension LS1, LS2, LS3, LS4, LS5, LS6 has a weight of $\frac{1}{3} \div 6 = 0.055$ *or* 5.6%). The indicator weights are shown in parenthesis in the table above. If an individual is deprived in an indicator, then the person has a score of 1 otherwise 0. For instance, first individual (Ind1) is deprived regarding the indicators ED1, LS1, LS2, LS5 and LS6. Now, we will compute the weighted deprivation score (DS) for each individual as follows:

DS of Ind1 = $1 \times 0.166 + 0 \times 0.166 + 0 \times 0.166 + 0 \times 0.166 + 1 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 = 0.386 = \mathbf{38.6\%}$ **.**

DS of Ind2 = $1 \times 0.166 + 1 \times 0.166 + 1 \times 0.166 + 1 \times 0.166 + 1 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 = 0.829 = \mathbf{82.9\%}$ **.**

DS of Ind3 = $1 \times 0.166 + 0 \times 0.166 + 0 \times 0.166 + 1 \times 0.166 + 1 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 = 0.442 = \mathbf{44.2\%}$ **.**

The weighted deprivation scores of Ind1, Ind2, and Ind3 are all superior than 33% which is the standard poverty cut-off in the MPI. In this context, these individuals are considered multidimensionally poor. In other words, the headcount ratio of poverty (H) is equal to $100\% \left(= \left(\frac{3}{3}\right) \times 100\right)$. The intensity of poverty (A), which corresponds to the average deprivation among the poor, is equal to 0.552 $\left(= \frac{(0.386 \times 1 + 0.829 \times 1 + 0.442 \times 1)}{3} \right)$. The average poor person is deprived in 55.2 per cent of the weighted indicators. Herewith, the adjusted headcount ratio (M_0) (or MPI) is $H \times A = 1 \times 0.552 = 0.552 = 55.2\%$. The AF methodology is also known as *counting approach* (which is already used in some poverty studies in the literature,

for instance, Townsend's poverty study in Britain (1979), the Unsatisfied Basic Needs poverty framework developed by the ECLAC). In fact, identification process of the poor is based on counting process (counting and summing the weighted deprivations that the household experiences, to obtain household total weighted deprivation score) but the aggregation process (i.e., obtain an overall poverty rate such as the poverty headcount ratio) is based on the FGT family of measures (the AF methodology adapts the unidimensional FGT family of measures[220] (Foster al. 1984) to the multidimensional level).

In second example, we will focus on four households H1, H2, H3, H4, H5 composed of 3, 4, 2, 2,1 individuals respectively (i.e., household size). As in the first example above, we will use the same number of dimensions, and indicators, and the same equal weighting process. But, this time, we will mention the name of indicators and to what they correspond, as in the original MPI (see UNDP Technical notes of the Human Development Report 2016). The full information on indicators is described in the table below (UNDP 2016, 9-10)[221].

Table 9. MPI Indicators and their description, source: UNDP Technical notes of the Human Development Report 2016, 9-10

Short name of indicator	**Indicator**	**Deprivation threshold**
ED1	School attainment (in terms of years of schooling)	Household is deprived if no one in household has completed at least six years of schooling
ED2	School attendance (for children)	Household is deprived if a school-age child (up to grade 8) is not attending school
H1	Nutrition	Household is deprived if a household member (for whom there is nutrition information) is malnourished, as measured by the body mass index for adults (women ages 15–49 in most of the surveys) and by the height-for-age *z* score calculated using World Health Organization standards for children under age 5
H2	Child mortality	Household is deprived if a child has died in the household within the five years prior to the survey

[220] See monetary poverty section for the essential FGT measures to measure poverty in a unidimensional perspective (i.e., monetary perspective).

[221] With the same indicators listed below in Table 10, Alkire and Robles (2016) propose also a dashboard approach. In dashboard, there is no index construction process. The deprivation score (in percentage) in each of the ten indicators is produced (i.e., level of deprivation/the percentage of people deprived in each of the ten indicators); for instance, they found that among 5.2 billion people across 101 countries, 21.8% of them lack electricity, 25.1% of them is deprived in safe water or must walk 30 minutes or more to obtain it.

LS1	Electricity	Household is deprived if household do not have access to electricity
LS2	Safe drinking water	Household is deprived if household do not have access to clean drinking water or if the source of clean drinking water is located more than 30 minutes away by walking
LS3	Improved sanitation	Household is deprived if household do not have access to improved sanitation or if improved, it is shared
LS4	Cooking fuel	Household is deprived if household uses 'dirty' cooking fuel (dung, wood or charcoal)
LS5	Flooring	Household is deprived if household lives in a home with a dirt, sand or dung floor
LS6	Assets ownership	Household is deprived if household do not have at least one asset related to access to information (radio, TV, telephone) and do not have at least one asset related to mobility (bike, motorbike, car, truck, animal cart, motorboat) or at least one asset related to livelihood (refrigerator, arable land, livestock).

In the table below, we show the performance of five households in terms of indicators' values as a hypothetical example (inspired by the example given in the technical notes of the UNDP Human development reports).

Table 10. Example 2 of the computation of the MPI, source: Author

	Household				
Variables	**H1**	**H2**	**H3**	**H4**	**H5**
Household size	3	4	2	2	1
ED1	0	0	1	1	1
ED2	1	0	1	0	1
H1	1	1	1	0	0
H2	0	0	0	1	0
LS1	1	1	1	0	0

LS2	0	1	1	1	0
LS3	1	0	0	1	0
LS4	1	1	1	1	1
LS5	1	0	1	0	1
LS6	0	0	1	0	1
Weighted deprivation score (DS)	0.552	0.331	0.773	0.497	0.497

Note: Household size corresponds to the number of people living in the household.

According to the MPI methodology, the indicators ED1, ED2, H1, H2 has 0.166 as a weight and the indicators LS1, LS2, LS3, LS4, LS5, LS6 has 0.055 as a weight. Household is considered deprived regarding the indicator if she has the score 1 in the indicator (otherwise 0: non-deprived). The computation of the total weighted deprivation score (DS) of each household is as follows:

DS of H1 = $0 \times 0.166 + 1 \times 0.166 + 1 \times 0.166 + 0 \times 0.166 + 1 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 = 0.552 = \mathbf{55.2}\%$.

DS OF H2 = $0 \times 0.166 + 0 \times 0.166 + 1 \times 0.166 + 0 \times 0.166 + 1 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 = 0.331 = \mathbf{33.1}\%$.

DS of H3 = $1 \times 0.166 + 1 \times 0.166 + 1 \times 0.166 + 0 \times 0.166 + 1 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 = 0.773 = \mathbf{77.3}\%$.

DS of H4 = $1 \times 0.166 + 0 \times 0.166 + 0 \times 0.166 + 1 \times 0.166 + 0 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 = 0.497 = \mathbf{49.7}\%$.

DS of H5 = $1 \times 0.166 + 1 \times 0.166 + 0 \times 0.166 + 0 \times 0.166 + 0 \times 0.055 + 0 \times 0.055 + 0 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 + 1 \times 0.055 = 0.497 = \mathbf{49.7}\%$.

Based on the total (weighted) deprivation score of the households, we can say that first, third, fourth, and fifth households' DSs are higher than the poverty threshold (33.3%). The individuals, therefore, living in these poor households are considered poor (i.e., there are $3 + 2 + 2 + 1 = 8$ poor persons). The headcount ratio of poverty ($H = \frac{q}{n}$ where q is the number of (multidimensionally) poor[222] and n is the number of population) is, therefore, obtained as follows: $\left(\frac{8}{12}\right) \times 100 = 66.6\%$. This means that the proportion/incidence of the poor within this population is 66.6% (i.e., 66.6% of the population is poor). On the other hand, the intensity (or breath) of poverty (A), computed based on the poor households' scores, is obtained as follows: $\frac{0.552 \times 3 + 0.773 \times 2 + 0.497 \times 2 + 0.497 \times 1}{8} = 0.587$. This latter statistics means that the average

[222] As a reminder, here, *poor* is the person having a total deprivation score equal at least a third of the weighted indicators (i.e., the person living in a household having a deprivation score equal at least a third of the weighted indicators).

poor person is deprived in 58.7 per cent of the weighted indicators[223]. By adjusting the headcount ratio (H) with the intensity of poverty (A)[224], we obtain the multidimensional poverty index M_0 (or MPI), as the product of H and A , as follows: $H \times A = 0.666 \times 0.587 = 0.390 = 39\%$. As we can see the MPI yields a *single* statistic to measure poverty (Alkire and Robles qualifies the MPI as "a summary headline indicator", p.12). It is important to note that the (multidimensional) headcount ratio and intensity of poverty are obtained based on the poor's deprivation scores; in other words, the deprivation scores of non-poor households are *not* taken into account. The AF methodology calls this process *censoring* (i.e., the deprivation scores of the non-poor are *censored* from the MPI computation). Thus, (multidimensional) headcount ratio is called *censored* headcount ratio. The MPI methodology allows also sub-group analysis (e.g., gender, urban/rural, region, ethnic group) and computes the contribution of each group to overall poverty, the censored headcount ratio of indicator (i.e., the number of deprived people in the indicator divided by the number of population), the percentage contribution of each indicator to overall poverty based on censored headcount ratio. Additional robustness checks notably in terms of sensibility of country rankings to variations in poverty or deprivations thresholds, in weights, in sample variability are suggested in the MPI[225].

The MPI indicators are linked to the Millennium Development Goals (MDGs)[226], for monitoring the progress among developing countries: (i) the years of schooling and child school attendance are used for monitoring the second goal of the MDGs; (ii) the nutrition is used for monitoring the first goal of the MDGs; (iii) child mortality is used for monitoring the fourth goal of the MDGs; (iv) the four indicators of living standards, namely improved sanitation, safe drinking water, cooking fuel and assets ownership are used for monitoring the seventh goal of the MDGs. The two other indicators of the living standards, namely flooring and electricity are used for evaluating the quality of housing conditions.

As we have mentioned above, the MPI is also adapted at national, or regional level, or by researchers at scientific platform for country analysis. Alkire and Apablaza (2016) developed the MPI for the European Union region (except Belgium, Bulgaria, Ireland, Croatia, Malta, Romania, Iceland, Norway and Switzerland) by employing the EU-SILC data from 2006 to 2012. Their computation is based on 6 dimensions: income, employment, material deprivation, education, environment, health. The table below shows indicators per dimension, which are used in the analysis (see Alkire and Apablaza 2016, 13).

Table 11. Dimensions and Indicators used in the computation of the MPI across European Union countries by Alkire and Apablaza (2016, 13).

Dimension	**Indicator**	**First cut-off**
Income	At-risk-of poverty (1/6=0.166)	Individual is deprived if he or she lives in a household whose equivalized disposable income

[223] The intensity of poverty corresponds to the average share of weighted deprivations that the multidimensionally poor person experiences (i.e., the average deprivation score of multidimensionally poor people).
[224] The headcount ratio (H) is also called as the *multidimensional* headcount ratio.
[225] The MPI is already tested regarding these three types of variations (see Alkire and Santos 2013, Alkire et al. 2010).
[226] Please refer to the link below for the MDGs (accessed 30 June 2018):
http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm

		is above 60% of the national median
Employment	Quasi-joblessness (1/6=0.166)	Individual is deprived if he or she lives in a household where the ratio of the total number of months that all household members aged 16-59 have worked during the income reference year and the total number of months the same household members theoretically could have worked in the same period is higher than 0.2
Material deprivation	Severe material deprivation (1/6=0.166)	Individual is *not* deprived if he or she has at least six of the following achievements: the ability to avoid arrears; to afford one week of holidays; a meal with meat, chicken, fish or vegetarian equivalent; to face unexpected expenses; and to keep home adequately warm; owns a car; a color TV; a washing machine; a telephone.
Education	Education (1/6=0.166)	Individual is not deprived if he or she has completed primary education
Environment	Noise (1/6/4=0.0416)	Individual is deprived if he or she lives in a household that experiences low noise from neighborhood or from the street
	Pollution (1/6/4=0.0416)	Individual is deprived if he or she lives in a household that experiences low pollution, grime or other environmental problems
	Crime (1/6/4=0.0416)	Individual is deprived if he or she lives in a household that experiences low crime, violence or vandalism in the area
	Housing (1/6/4=0.0416)	Individual is deprived if he or she lives in a dwelling with no leaking roof, damp walls, rot in window frames or floor

Health	Health (1/6/4=0.0416)	Individual is *not* deprived of he or she evaluates her own health as fair or above
	Chronic illness (1/6/4=0.0416)	Individual is *not* deprived if he or she has no chronic illness or long-term health condition
	Morbidity (1/6/4=0.0416)	Individual is *not* deprived if he or she reports no limitations due to health problems
	Unmet medical needs (1/6/4=0.0416)	Individual is *not* deprived if he or she does not report unmet medical needs

Compared to the global MPI: (i) the three indicators of Europe 2020 (income, employment and material deprivation); (ii) the dimension environment; (iii) more indicators are included in health dimension, in Alkire and Apablaza's work. The weights are assigned equally across dimensions and across indicators within dimensions (they are mentioned in parenthesis in second column of the table above).

Similar to Alkire and Apablaza (2016), Weziak-Bialowolska and Dijkstra (2014) have also proposed the regional adapted MPI (MPI-Reg onwards) in the European Union region (as a JRC science and policy report by the joint research centre and the European Commission). They used three dimensions (health, education, living standards), with in total 7 indicators. Concerning health, they used the indicators of general health, unmet medical need due to lack of affordability and accessibility, unmet dental need due to lack of affordability and accessibility; concerning education, they used the indicators of educational attainment; and concerning living standards, they used the indicators of material deprivation, housing problems, environment. The indicators are not exactly the same in the two studies; for instance, Alkire and Apablaza (2016) included housing problems in environment dimensions, they did not employ the aspects of unmet dental needs and household crowding, but they used income and employment as dimensions compared to Weziak-Bialowolska and Dijkstra (2014). Furthermore, as a contribution to the MPI calculations, Weziak-Bialowolska and Dijkstra (2014) proposed three sub-indices: the multidimensional poverty in health index (MPI-H), the multidimensional poverty in education index (MPI-E), the multidimensional poverty in living standards index (MPI-L). Each person has got a (weighted) deprivation score in each dimension. For instance, if a person has a weighted deprivation score equal or higher than 2/3 in health dimension, then this person is considered multidimensionally poor with respect to health.

Besides the examples we saw above, Suppa (2015) has proposed the MPI for Germany by employing the German socio-economic panel (SOEP) data. Regarding dimensions, he used education, housing, health, material deprivation, social participation, and employment (we note that the indicators of social participation are quite similar to some indicators used in material deprivation dimension in traditional EU-SILC analysis). As the questionnaires of SOEP and EU-SILC are different, the indicators developed by Suppa bring a new perspective to multidimensional poverty studies in European context. Some of them are: (i) in education, he combined quantity and quality of education; (ii) in employment, he included the indicator of precarious employment; (iii) in health, he took in consideration disability, and body mass index.

In total, there are 15 indicators which are weighted according to equal weighting methodology of the MPI. The indicators are listed in the table below (Suppa 2015, 30).

Table 12. Dimensions, indicators, first cut-off developed by Suppa (2017) for the German MPI, source: Suppa (2015, 30)

Dimension	**First cut-off**
Education	Individual is deprived if elementary schooling not completed or elementary schooling completed but no vocational qualification
	Individual is deprived if there are less than 10 books in her/his household
Housing	Individual is deprived if the dwelling where he or she lives requires major renovation or is ready for demolition
	Individual is deprived if any of the following are absent in the household: bath, shower, kitchen, warm water, toilet
	Individual is deprived if the dwelling where he or she lives is overcrowded (i.e., less than one room per person)
Health	Individual is deprived if he or she is partially or severely disabled
	Individual is deprived if he or she reports 2/4 health issues (the four health issues are (i) a strong limitation when climbing stairs: (ii) a strong limitation for tiring activities; (iii) physical pain occurred always or often during the last 4 weeks, and (iv) the health condition limited always or often socially)
	Individual is deprived if her/his body mass index is larger than 30
Material deprivation	Individual is deprived if he or she reports 2/4 goods missing for financial reasons (the four goods asked for are (i) a warm meal; (ii) whether friends are invited for dinner; (iii) whether money is put aside for emergencies; and (iv) whether worn out furniture is replaced)
	Individual is deprived if any of the following are absent in her/his life: life insurance, pension, owning the house or apartment, other house, financial assets, commercial enterprise, tangible assets
Social participation	Individual is deprived if 5/7 activities are never performed; remaining at most less than monthly (activities included are (i) going to the movies, pop music concerts, dancing, disco, etc.; (ii) going to cultural events (such as concerts, theatre, lectures); (iii) doing sports yourself: (iv)

	volunteer work; (v) attending religious events; (vi) helping out friends, relatives or neighbors; (vii) involvement in a citizens' group, political party, local government.
	Individual is deprived if he or she never meets friends
Employment	Individual is deprived if he or she is unemployed
	Individual is deprived if he or she works involuntarily work less than 30 hours per week
	Individual is deprived if he or she is engaged in precarious employment

Through the three examples of adaptation of the MPI shown above, we can say that some dimensions are often used such as education, health, dwelling; but the selected indicators are mainly varying as a function of questionnaires. Concerning the relative importance of dimensions and indicators, the equal weight methodology of the global MPI is often applied.

The Colombian Multidimensional Poverty Index (CMPI onwards) is launched by the Colombian National Planning Department for better capturing the actual living conditions of the poor in a multidimensional perspective and for better evaluating the impact of public policies on the elimination of poverty (e.g., The Division of Urban Development of the National Planning Department's 2014 goals for providing better access to public utilities and for increasing housing conditions). Before the CMPI, Colombia was already familiar with multidimensional measures such as the Unmet Basic Needs Index (UBN), the SISBEN (Sistema de Identificación de Potenciales Beneficiaros de Programas Sociales), the Human Development Index (HDI), the Living Conditions Index (LCI), the Human Opportunities Index (HOI) but for some technical aspects these measures were considered insufficient[227] (see Angulo et al. 2016). The CMPI (see Angulo et al. 2016 for all technical details) is based on the AF methodology as a national adaptation of the global MPI (with the same nested weighting process, the household as the unit of analysis, the poverty cut-off settled at 33%), is constructed by using the Colombian Living Standards Measurement Survey (LSMS) 1997, 2003, 2008 and 2010[228]. As dimensions, CMPI employed education, childhood and youth conditions, employment, health, access to public utilities and housing conditions. Concerning indicators[229] (i) in *education*: educational

[227] Angulo et al. (2016) mention three major problems in these existing indices before the CMPI: (i) non-fulfillment of some axiomatic rules defined by Sen (1976), such as the Unmet Basic Needs Index is not sensitive to greater deprivations (e.g., if household becomes more deprived); (ii) they do not necessarily focus only on the poor, they include both variations at poor and non-poor levels; (iii) in terms of content (e.g., indicator, dimension), they do not have a good coverage. In general, these existing indices were considered poor instruments in the context of Colombian for measuring multidimensional poverty.

[228] Angulo et al. (2016) mention in their article that their computations are based on LSMS 1997, 2003, 2008, and 2010 data whilst the official statistical office of Colombia (DANE) mention on their website that they obtained CMPI estimations for 1997, 2003, 2008 and 2010 by employing Encuestas de Calidad de Vida (ECV) del DANE (Quality of Life Surveys carried out by DANE). For further information, please refer to the link below (i.e., the publication on the methodology of poverty measurement in Colombia published on DANE's website (CONPES 2012, 11- accessed 9 January 2020)):
https://www.dane.gov.co/files/acerca/Normatividad/CONPES-150.pdf

[229] For description of indicators, see Angulo et al. (2015).

achievement, literacy; (ii) in *childhood and youth conditions*: school attendance, no school lag, access to childcare services, children not working; (iii) in *employment*: no one in long-term unemployment, formal employment; (iv) in *health*: health insurance, access to health services; (v) in *access to public utilities and housing conditions*: access to water source, adequate elimination of sewer waste, adequate floors, adequate external walls, no critical overcrowding are used.

In general, selection of indicators can be considered as a delicate process due to correlations between indicators and target group of indicators. For instance, in the dimension of education, the following indicators are used: average education level for people 15 and older living in a household, percentage of people living in a household 15 and older who know how to read and write. At first glance, these two indicators seem likely correlated due to their target group and content. Notably, the employment related indicators such as child labour, long-term unemployment, formal vs. informal employment represent a new wave in the MPI. Notably, the CMPI aims at detecting more detailed information in health and childhood domains such as percentage of household members over the age of 5 that are insured by the Social Security Health System (in addition to more usual question having access or not to health services in case of need, the school lag and access to childcare services (in health, nutrition and education). Compared to the global MPI, we can say that the indicators of nutrition, child mortality, electricity, cooking fuel, assets ownership are not considered in the CMPI, despite these indicators can be considered as some of primarily important needs as a human being. The CMPI is computed, in addition, at urban, rural, region levels. It is reported that in 2010, 20% of population was poor both in income and multidimensional perspectives, 21% was only income poor, 10% was only multidimensionally poor and 49% was nor income neither multidimensionally poor. As one of the main weakness of the CMPI, this measure does not show the distribution of indicators among the non-poor population. The 49% of population is certainly deprived in some respects in their life, but also some of them may be in risk boundary (i.e., fall below the poverty line over time).

The Mexican multidimensional poverty measurement (see Foster 2010, CONEVAL 2010 for technical details) applied by the CONEVAL is based on the AF methodology and is first attempted in 2008 by employing the 2008 Socioeconomics Conditions Module in the National Survey of Household Income and Expenditures (MCS-ENIGH). The multidimensional poverty measure is aimed at evaluating three spaces: social rights, economic wellbeing and the territorial context (CONEVAL 2010, 57). In this context, economic wellbeing is measured by traditional income measure, social rights are measured by social deprivation index. In other words, the Mexican measure combines both income and social aspects. The social deprivation index includes six dimensions which are: education, access to health services, access to social security, shelter characteristics, access to basic services, access to food[230]. The social deprivation index is built over two steps. First, *deprivation identification* process determines whether the person is deprived or not in the indicator. The unit of analysis is settled individual

[230] Level of social cohesion is also recommended but it is not included in 2008 calculus, it is not taken in consideration. In fact, the social cohesion is the only dimension associated with the territoral context. The four indicators are defined for this dimension: "the Gini index, the degree of social polarization of every state and municipality, the income ratio of the population living in extreme multidimensional poverty relative to the popuation that is not living in multidimensional poverty and that is not vulnerable, the social netwroks perception index" (CONEVAL 2010, 36).

but in case individual level measures are not available the household level measures are applied (i.e., household condition is valid for every member of the household). In total, there are six social indicators and all of them are binary variable taking the value 1 if the person is deprived in the indicator otherwise 0. As there are six variables in total for six dimensions, we can see that each dimension is represented by one indicator[231]. Second, *social deprivation index* is constructed by simply summing the scores of the six indicators. Once, the score of social deprivation index is obtained, the person is considered socially deprived if and only if her/his deprivation index score is equal or superior than one (i.e., called *deprivation threshold*). Each indicator is counted equally important (i.e., each indicator is associated with the same weight). As a justification of the equal weighting, the CONEVAL argues that the Mexican multidimensional poverty measure is based on human rights approach and "the principles of indivisibility and interdependence of human rights, which recognize that the non-fulfilment of any of the human rights infringes upon the fulfilment of the others, and that there is no human right that can be considered more important than any other" (2010, 30).

The methodology brings three main statistics (headcount ratio, depth of poverty, intensity of poverty) for four groups: *multidimensional poor* (people with an income level below the income poverty line[232] and deprived in at least one social right), *vulnerable due to social deprivation* (people with an income level above the income poverty line but deprived in at least one social right), *vulnerable due to income* (people with an income level below the income poverty line but not deprived regarding social rights), *not multidimensional poor and not vulnerable* (people with no social deprivation and with an income level above the income poverty line). Moreover, the multidimensional poor population is distinguished into two groups: *extreme* multidimensionally poor vs. *moderate* multidimensionally poor. The first group includes individuals with a very low-income level, "even if spent entirely on food, they could not buy the necessary nutrients for a healthy life" and deprived at least in three social rights (2010, 33). The second group, therefore, covers the rest of multidimensionally poor.

The South African Multidimensional Poverty Index (SAMPI onwards, see Statistics South Africa 2014 for technical details[233]) is constructed by the Statistics South Africa by employing the AF method and the population censuses of 2001 and of 2011. Especially, this new index was built to monitor the evolution of poverty at province and municipal levels. Previously, the Provincial Indices of Multiple Deprivation were applied by using the population census of 2001 (PIMD onwards, see Noble et al. 2006 for technical details[234]). However, Statistics South Africa (2014) mention that the PIMD was not fully decomposable by demographic and geographical indicators. The SAMPI is composed of four dimensions and eleven indicators as follows: in the dimension of *health*: child mortality, in the dimension of *education*: years of schooling, school attendance, in the dimension of *standard of living*: fuel for lighting, fuel for heating, fuel for cooking, water access, sanitation type, dwelling type, asset ownership, in the dimension of *economic activity*: unemployment. Compared to the global MPI, (i) the SAMPI includes the

[231] Indicators are chosen regarding four criteria: "conceptual relevance", "empirical feasibility", "specificity", "usefulness for public policy" (CONEVAL 2010, 42).
[232] In the official document of CONEVAL (2010, 32), the income poverty line is called the *wellbeing threshold*.
[233] The document is available in the link below (accessed 7 January 2020):
http://www.statssa.gov.za/?page_id=1854&PPN=Report-03-10-08
[234] The document is available in the link below (accessed 7 January 2020):
http://www.statssa.gov.za/census/census_2001/PMID/PIMDTechnicalReport2006.pdf

dimension of economic activity, the indicators of lighting, heating; (ii) due to available data, there is no the indicator of flooring in the SAMPI, but instead dwelling type is included in; (iii) in the SAMPI, the health dimension does not include the information on nutrition, hunger, food security; (iv) the deprivation cut-offs in cooking, water access, sanitation type, and asset ownership differ slightly in the SAMPI. Equal weighting is applied over dimensions and indicators, and traditional AF statistics (headcount ratio of poverty, intensity of poverty, multidimensional poverty index score) are produced.

Moreover, Rogan (2016) has suggested computing multidimensional gender poverty gap based on the 2008 wave of the South African National Income Dynamic Study. Rogan has produced gender differences in both the global MPI (by employing 3 dimensions, education, health, and living standard; and 9 indicators, schooling years, child enrolment, child mortality, nutrition, water, electricity, sanitation, cooking fuel, and assets) and income poverty. Mushongera et al. (2017), has adapted the Global Multidimensional Poverty Index for Gauteng province, by using the 2011 and 2013 Quality of Life surveys. Four dimensions and nine indicators are used as follows: in the dimension of *standard of living*: housing dwelling, housing overcrowded, water access, sanitation, energy access, communication assets, in the dimension of *food*: access to food, in the dimension of *economic activity*: unemployment, in the dimension of *education*: years of school attendance. As results, both deprivation scores (in percentage) among the poor and the non-poor, distribution of multidimensional poverty index score by municipalities are provided.

In addition to composite indices such as the HPI or MPI, there are also various poverty dashboards which are relatively easy in terms of construction and interpretation. A dashboard includes a certain number of dimensions regarding the purpose of the study and generally, each dimension includes only one indicator (for instance, in European context, most recently, to monitor the Europe 2020 poverty and social exclusion target, 3 dimensions and 1 indicator for each dimension were defined[235]. The indicators are: (i) at risk-of-poverty rate; (ii) low-work intensity population (quasi-jobless households) rate; (iii) severe material deprivation rate. For more information, see social exclusion in Appendix 9.1). However, there are some dashboards including a set of indicators within a dimension (e.g., OECD Better Life indicators, United Nation's Millennium Development Goals (MDGs onwards) dashboard is composed of 49 indicators to monitoring 18 targets within 8 goals such as the proportion of underweight children under 5 years of age, child mortality rate, etc.), showing that one indicator per each dimension is not necessary the rule. The most important difference between a dashboard and a (composite) index is the former does not include an aggregation process; for instance, if in education dimension there are two indicators, then the scores of these two indicators are not aggregated together, or if there is one indicator per dimension there is no any aggregation process over dimensions. Each indicator keeps its singularity, and the focus is on the percentage of people deprived regarding each indicator. In dashboards, therefore, the aim is not to obtain an overall rate of poverty in opposition to composite indices. In this regard, the United Nations Economic Commission for Europe (UNECE onwards) recommends that "it is desirable for dashboards to have a clear hierarchy and set of priorities, and to name a small set of 5-10 indicators as tier 1 or key indicators" (see UNECE 2017, 117). The main advantage of

[235] See http://ec.europa.eu/social/ (accessed 30 June 2018): But also, Atkinson et al. (2002), Marlier et al. (2007) provide information on the advantage of using social indicators for combatting social exclusion and poverty in the European Union.

dashboards is there is no necessity to gather data from only one source. In opposite, due to aggregation process, composite indices should be drawn from one data set. In some cases, not all variables do exist in the same data set, for instance, long-term unemployment rate can be find based on micro data and information on the ratio of informal sector can be find by employing administrative data. Dashboards, therefore, allow to combine different data sources for gathering information on poverty from different aspects. However, there are also a number of drawbacks in their methodology[236]: first, when there are many indicators, which indicators should be given importance does not appear clearly; second, when the indicators may come from different data sources the information on people suffering multiple deprivations at the same time (i.e., joint deprivation) cannot be obtained. Moreover, non-weighting across indicators and the lack of an overall poverty rate are other weak points of dashboards. Often, there is no clear process on weighting and policy makers have to interpret the indicator scores separately. The last point makes problematic the comparison of macro level poverty rates across countries, regions, or communities. In composite indices, the most appreciated point is that they provide a 'headline figure' (i.e., overall poverty rate) of poverty whose tendency over years can be easily obtained.

Now, we will see in more details the dashboard of the MDGs, Sustainable Development Goals (SDGs onwards) and OECD Better Life Index. The MDGs including eight goals are introduced by the United Nations in September 2000. In this international meeting in New York at United Nations headquarters, the United Nations Millennium Declaration[237] was adopted by world leaders to commit in a global partnership by promoting a set of time-bound development targets (with a deadline of 2015) notably for reducing extreme poverty. In this declaration, in the section of Development and poverty eradication (there are also other points such as peace, security and disarmament), it is pointed out to resolve following problems by the end of 2015 (see United Nations 2000, 5):

"

• To halve, by the year 2015, the proportion of the world's people whose income is less than one dollar a day and the proportion of people who suffer from hunger and, by the same date, to halve the proportion of people who are unable to reach or to afford safe drinking water.

• To ensure that, by the same date, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling and that girls and boys will have equal access to all levels of education.

• By the same date, to have reduced maternal mortality by three quarters, and under-five child mortality by two thirds, of their current rates.

• To have, by then, halted, and begun to reverse, the spread of HIV/AIDS, the scourge of malaria and other major diseases that afflict humanity.

[236] UNECE (2017) provides comprehensive discussions on advantages and drawbacks of dashboard methodology.

[237] The official document is available in the link below:
https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_55_2.pdf
Accessed 8 May 2020.

• To provide special assistance to children orphaned by HIV/AIDS.

• By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers as proposed in the "Cities Without Slums" initiative.

• To promote gender equality and the empowerment of women as effective ways to combat poverty, hunger and disease and to stimulate development that is truly sustainable.

• To develop and implement strategies that give young people everywhere a real chance to find decent and productive work.

• To encourage the pharmaceutical industry to make essential drugs more widely available and affordable by all who need them in developing countries.

• To develop strong partnerships with the private sector and with civil society organizations in pursuit of development and poverty eradication

."

In this context, the eight goals are defined as follows: (i) eradicate extreme poverty and hunger; (ii) achieve universal primary education; (iii) promote gender equality and empowerment of women; (iv) reduce child mortality; (v) improve maternal health; (vi) combat HIV/AIDS, malaria and other diseases; (vii) ensure environmental sustainability; (viii) global partnership for development. Through 8 dimensions, in total 49 indicators are defined[238]. For instance, regarding first goal, there are 3 research axes with in total 9 indicators. First axe focuses on reducing by half, between 1990 and 2015, the proportion of population having an income level inferior than US$1 per day. To monitor this first axe, 3 indicators are defined: (i) the proportion of population living under US$1.25 per day in terms of PPA (or based on national monetary poverty line); (ii) poverty gap index; (iii) the share of the poorest quintile of the population in national consumption. Second axe focuses on guaranteeing full employment and possibility of finding a decent and productive job, including for women and youth population. To monitor this second axe, 4 indicators are defined: (i) PIB growth rate per person; (ii) the ratio of employment and population; (iii) the proportion of economically active population living under US$1.25 per day including PPA; (iv) the proportion of self-employers and family workers regarding the total economically active population. Third axe focuses on reducing by half, between 1990 and 2015, the proportion of population suffering hunger. To monitor this third axe, 2 indicators are defined: (i) the ratio of underweight children under 5 years old; (ii) the proportion of population that does not reach the minimum level of calorie intake. As we can see the progress monitoring is carried out via the defined indicators. There is no any aggregation process. Policy makers and researchers look directly at the indicator scores for country, regional, community comparisons. The 49 indicators construct the MDGs dashboard. In latest 2013 progress chart report[239], for instance, regarding first goal, Sub-Saharan Africa, Southern Asia, Oceania have experienced the lowest performance in reducing extreme poverty by half, in providing productive and decent work and in reducing hunger by half (except Oceania faces moderate hunger). Countries, regions, and communities are, therefore, compared in each

[238] The list of indicators is available in the link below:
http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm
Accessed 30 June 2018.

[239] The progress chart is available in the link below (accessed 30 June 2018):
http://mdgs.un.org/unsd/mdg/Resources/Static/Products/Progress2013/Progress_E.pdf

indicator (there is no intra-axes neither intra-goal compensation such as having a good performance in one indicator does not compensate for a bad performance in another indicator. This is the main feature of dashboards).

After the MDGs, the United Nations has pursued his initiative for improving development and reducing poverty in worldwide, with a new development programme called *Sustainable Development Goals* (SDGs)[240] launched in 2015. The 2030 Agenda for Sustainable Development adopted by countries includes 17 Sustainable Development Goals (in 2016, this agenda in reinforced by the Paris Agreement on climate change) to achieve by 2030. The 17 goals are as follows: (1) no poverty (end poverty in all its forms); (2) zero hunger; (3) good health and well-being; (4) quality education; (5) gender equality; (6) clean water and sanitation; (7) affordable and clean energy; (8) decent work and economic growth; (9) industry, innovation and infrastructure; (10) reduced inequalities; (11) sustainable cities and communities; (12) responsible consumption and production; (13) climate action; (14) life below water; (15) life on land; (16) peace, justice and strong institutions; (17) partnership for goals. The SDGs are very close to the MDGs, but the former include in addition new goals for promoting the importance of nature protection and climate change. The SDGs draw attention on non-monetary side of poverty: "Poverty is more than the lack of income and resources to ensure a sustainable livelihood. Its manifestations include hunger and malnutrition, limited access to education and other basic services, social discrimination and exclusion as well as the lack of participation in decision-making"[241]. To monitor the SDGs, the UN General Assembly adopted the global indicator framework on July 6, 2017. Later, in March 2018, during the 49th Session of the Statistical Commission, the official indicator list within this global indicator framework is defined. The official list includes 232 indicators in total[242].

The first goal of the SGDs related to the end of poverty, includes 5 research axes with in total 10 indicators to monitor the goal. The table below shows the indicators per research axe.

Table 13. SDG Indicators (dashboard of indicators) for Goal 1. End poverty in all its forms everywhere.

Axe	**Indicators**
1. By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day	I.1.1 Proportion of population below the international poverty line, by sex, age, employment status and geographical location (urban/rural)

[240] Please see the link below for further information on SDGs (accessed 30 June 2018):
https://www.un.org/sustainabledevelopment/development-agenda/
The information on Paris Agreement is available in the link below:
https://www.un.org/sustainabledevelopment/climatechange/

[241] Please see the link below for further information (accessed 30 May 2018):
https://www.un.org/sustainabledevelopment/poverty/

[242] It is mentioned that "the total number of indicators listed in the global indicator framework of SDG indicators is 244. However, since nine indicators repeat under two or three different targets, the actual total number of individual indicators in the list is 232" (see https://unstats.un.org/sdgs/indicators/indicators-list/, accessed 30 May 2018). The list of indicators is available in the link below:
https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%20refinement_Eng.pdf
Accessed 30 June 2018.

2. By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	I.2.1 Proportion of population living below the national poverty line, by sex and age
	I.2.2 Proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions
3. Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable	I.3.1 Proportion of population covered by social protection floors/systems, by sex, distinguishing children, unemployed persons, older persons, persons with disabilities, pregnant women, newborns, work-injury victims and the poor and the vulnerable
4. By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources, as well as access to basic services, ownership and control over land and other forms of property, inheritance, natural resources, appropriate new technology and financial services, including microfinance	I.4.1 Proportion of population living in households with access to basic services
	I.4.2 Proportion of total adult population with secure tenure rights to land, (a) with legally recognized documentation, and (b) who perceive their rights to land as secure, by sex and type of tenure
5. By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters	I.5.1 Number of deaths, missing persons and directly affected persons attributed to disasters per 100,000 population
	I.5.2 Direct economic loss attributed to disasters in relation to global gross domestic product (GDP)
	I.5.3 Number of countries that adopt and implement national disaster risk reduction strategies in line with the Sendai Framework for Disaster Risk Reduction 2015–2030
	I.5.4 Proportion of local governments that adopt and implement local disaster risk reduction strategies in line with national disaster risk reduction strategies

Source : https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%20refinement_Eng.pdf. Accessed 30 May 2018.

As we can see from the table above, several indicators as a dashboard were defined to monitor the goal of ending poverty in all its forms everywhere within the SDG context.

Besides United Nation's initiative, OECD's Better Life Index (BLI onwards) is another good example of dashboard[243]. In BLI dashboard, there is in total 24 indicators (see the table below). OECD countries are compared regarding each indicator. On the website of OECD, there is also a possibility of rating dimensions according to their importance to the person (see http://www.oecdbetterlifeindex.org/#/53555355551, accessed 31 May 2018). In other words, the person who visits the website of BLI can evaluate the performance of her/his country regarding each dimension (but not regarding each indicator). The BLI focuses more on well-

[243] There is no aggregation process in BLI nevertheless she is called index but in fact she is a dashboard.

being than poverty, but it includes a number of same dimensions of multidimensional poverty studies. There are 11 dimensions: housing, income, jobs, community, education, environment, civic engagement, health, life satisfaction, safety, work-life balance[244]. Several indicators are defined within each dimension, as follows (see the table below):

Table 14. OECD Better life index dashboard

Dimension[245]	**Indicator**
Housing	Housing expenditure, dwellings with basic facilities, rooms per person
Income (and wealth)	Household financial wealth, household net adjusted disposable income
Jobs (and earnings)	Job security, personal earnings, long-term unemployment rate, employment rate
Community (social connections)	Quality of support network
Education	Years in education, student skills, educational attainment
Environment	Water quality, air pollution
Civic engagement (and gouvernance)	Stakeholder engagement for developing regulations, voter turnout
Health	Self-reported health, life expectancy
Life satisfaction (subjective well-being)	Life satisfaction (i.e., the average self-evaluation of life satisfaction, on a scale from 0 to 10)
Safety	Homicide rate, feeling safe walking alone at night
Work-life balance	Time devoted to leisure and personal care, employees working very long hours

Source (accessed 30 May 2018): http://www.oecdbetterlifeindex.org/#/54535435353

Among BLI dimensions, especially housing, income, education and health are those in common compared to usual multidimensional poverty measurements such as UNDP's Multidimensional Poverty Index. Regarding the dimension of *housing*, the indicator of housing expenditure is the ratio to housing costs on household gross adjusted disposable income, latest available year; the indicator of dwellings with basic facilities is the percentage of people with indoor flushing toilets in their home, latest available year; the indicator of rooms per person is the average number of rooms shared per person in a dwelling, latest available year. For the dimension of *income*, the household financial wealth is measured via the average value of a household's financial assets (saving, stocks) minus their liabilities (loans), latest available year; the household net adjusted disposable income is measured by the average amount of money that a household earns per year, after taxes, latest available year. For the dimension of *education*, years in education is measured via the average duration of formal education in which a five-year old child can expect to enroll during her/his lifetime until the age of 39; student skills is

[244] Please refer to the link below for further information (accessed 30 May 2018): http://www.oecdbetterlifeindex.org/#/53555355551

[245] OECD uses the term *topic* instead of dimension.

measured by the average performance of students, according to PISA (Programme for International Student Assessment); educational attainment is measured by the percentage of people, aged 25 to 64, having at least an upper-secondary (high-school) degree, latest available year. For the dimension of *health*, self-reported health is measured by the percentage of people reporting their health to be 'good or very good', latest available year; the life expectancy corresponds to the average number of years a person can expect to live, latest available year. Among other dimensions, life satisfaction is generally more employed in happiness or quality of life studies or more recently in subjective poverty investigations; jobs, environment, and safety are used in some multidimensional studies, but their employment depends on the structure of data and their availability across regions and countries. Other dimensions such as community, civic engagement, work-life balance are relatively few used in multidimensional poverty studies, they are more often employed in well-being measures.

Most recently, in OECD's *How's life in 2017? Measuring Well-being* report, the ranking of OECD countries in each of the BLI's indicators is provided (OECD 2017, 28-29)[246]. This simple ranking reflects the position of the country in the sample. If the country is at top third of the OECD, she takes the category one; if the country is at middle third of the OECD, she takes the category two; if the country is at bottom third of the OECD, she takes the category three. Then, numerical values are attributed to each category: the value 0 for the third category, the value 5 for the second category, and the value 10 for the first category. For obtaining an aggregate number of well-being for each country, the function of average is used: first, in each dimension, an average score of indicators based on numerical values is computed for each country; second, based on average scores obtained at first step, a final average score across all dimensions is computed for each country. We note that each indicator has an equal weight. As we can see, in this report, an index of well-being is obtained based on the indicators of the BLI dashboard.

UNECE (2007) recommends using the dimensions of work, lived environment in addition to classical dimensions such as education, health, and living standards. As we have already mentioned, one of the flexible points of dashboards is the use of different data sources for measuring different topics[247]. For instance, within the dimension of housing in the BLI, the indicator of dwellings without basic facilities comes from European Union Statistics on Income and Living Conditions (EU-SILC), and national statistical offices; however, housing expenditure is computed based on OECD national accounts database. Another possibility of dashboards is to obtain a joint distribution of two deprivations (i.e., two by two distributions) by employing a simple cross-tabulation process (e.g., the percentage of people deprived both in years of schooling and nutrition, or percentage of people deprived in years of schooling but not deprived in nutrition at the same time). However, this pairwise simultaneous deprivation does not allow to see multiple deprivations (i.e., more than two deprivations) at the same time.

[246] Please refer to the link below for the report (accessed 30 May 2018):
https://www.oecd-ilibrary.org/economics/how-s-life-2017/how-s-life-in-2017_how_life-2017-5-en
On the one hand, the indicators income and wealth, jobs and earnings, housing are considered part of the *material conditions*. On the other hand, the indicators work-life balance, health, education, social connections and community, civic engagement and governance, environment, safety, subjective well-being (life satisfaction) are considered part of the *quality of life*.

[247] In the link below, Better Life Index- Edition 2017, the OECD provides detailed information on the construction of indicators in terms of unit, data source(s) used, other data characteristics, and statistical concept: http://stats.oecd.org/index.aspx?DataSetCode=BLI (accessed 31 May 2018).

Venn diagrams are generally produced to observe joint distribution of three or four indicators. But still big diagrams and tables for reflecting multiple joint distribution are likely not easy for interpreting (Alkire and Robles 2016).

As we have mentioned, in the literature, there are various quantitative techniques such as Alkire and Foster's counting approach, composite indices, multivariate statistical techniques, fuzzy set, partial order theory, which have been used for measuring poverty from a multidimensional perspective. To Alcock (2006 [1993], 116), "the great advantage of quantitative measures is their scale and their anonymity. A statistical survey, if it is large enough and if the sample of respondents providing the data is carefully chosen, can provide an objective, arguably scientific, picture of the broader group or society from which it has been selected, particularly if standard measures of the statistical significance of the data have been reached". As we have seen the AF method (Alkire and Foster 2007, rev. 2009b, 2011a) is particularly known for her employment in the construction of the UNDP's global MPI as well as of the various national MPIs[248]. Among its internationally wide application, we can mention some of them as follows: in U.S. context, White and Yamasaki (2017), Wagle (2014), Mitra and Brucker (2014); in European context, Alkire et al. (2014), Alkire and Apablaza (2016); concerning destitution and absolute poverty in the Global South, Alkire and Santos (2010), Alkire and Seth (2016); in dynamic perspective, Alkire et al. (2017); at global level, Alkire et al. (2016); in India, Alkire and Seth (2015); in Lao PDR, Bader et al. (2016); in Pakistan, Khan et al. (2014, 2016); in Germany, Nowak and Scheicher (2017); in Indonesia, Hanandita and Tampubolon (2016); in six Latin American countries (Argentina, Brazil, Chile, El Salvador, Mexico and Uruguay), Battiston et al. (2009); in Latin America, Santos et al. (2015); in Gauteng province of South Africa, Mushongera et al. (2017).

The AF methodology is particularly suitable for binary and categorical data which are often employed in large scale surveys, e.g., question involving binary response such as yes or no. Meanwhile, in the literature, several multidimensional poverty indices based on quantitative data[249] are proposed and/or employed by scholars to measure poverty, such as Kakwani (1980a, 1980b), Clark et al. (1981), Foster et al. (1984), Foster and Shorrocks (1991), Bourguignon and Chakravarty (2003), Chakravarty et al. (1998), Chakravarty and Silber (2008), Tsui (2002), De Muro et al. (2011), Lustig (2011), Ravallion (2011), Duclos and Tiberti (2016), Mazziotta and Pareto (2016), Sessa (2016). Concerning multivariate statistical techniques, Coromaldi and Drago (2017) realized multidimensional poverty analysis in Italy by applying the multiple correspondance analysis, Luzzi et al. (2008) did a cluster analysis of multidimensional poverty in Switzerland, Asselin and Anh (2008) applied multiple correspondance analysis for measuring multidimensional poverty in Viet nam, Coromaldi and Zoli (2012) employed the non-linear principal component analysis to measure multidimensional poverty in Italy, Gasparini et al. (2013) used factor analysis to measure multidimensional poverty in Latin America and the Carribean, Ntsalaze and Ikhide (2018) employed multiple correspondence analysis to measure multidimensional poverty in South Africa. Among fuzzy approach-based studies, we can mention for instance the studies of Annoni et al. (2011), Cerioli and Zani (1990), Lemmi and Betti (eds.) (2006), Cheli and Lemmi (1995), Chiappro-Martinetti (1996, 2000), Verma et al. (2017). Moreover, there is growing new body of multidimensional poverty studies

[248] The MPIs of some upper-middle-income countries are portrayed in poverty measurement practices in upper-middle-income countries section.
[249] Quantitative data includes metric scale figures such as per centage, ratio numbers.

by employing partial order techniques, such as Fattore et al. (2011a, 2011b, 2012), Fattore and Maggino (2014), Fattore (2016), Maggino (2017), Annoni et al. (2015, 2017), Beycan and Suter (2017), Beycan et al. (2019).

Multidimensional poverty is, therefore, measured based on social dimensions such as education, health, living standards, instead of traditional income, or consumption patterns. Concerning its disadvantages, we can mainly mention the two follows: first, in order to show different capabilities, sub-group analysis become very important. But, in practice, a limited number of decomposition analysis such as in terms of rural vs. urban, regional disparities, male vs. female is generally applied. Social theory should be more used for detecting various patterns of capabilities. Second, an overall index method (i.e., aggregation) risks to hide different combinations of functionings and to take in consideration achievements as a proxy of capabilities. When the focus is on achievements, the notion of capabilities becomes not different than deprivation. In typical deprivation studies, the centre is directly on whether a person has or not a specific item; but in capabilities the centre should be on how the combinations of achievements and failures may affect the living style or the way that the person wishes to live without damaging the life of others.

3. Poverty in upper-middle-income countries

This chapter focuses on upper-middle-income countries through 3 sections. First of all, we will start off by exploring the socio-economic situation of select countries in this category based on some macro level statistics provided by the United Nations (UN), and the World Bank (WB). We will then continue by examining actual national poverty measurement practices/techniques applied by national statistical offices. In third section, we will more specifically work on the four upper-middle-income countries, namely South Africa, Turkey, Mexico and Brazil, which are the target group in our empirical analysis, and we will show the national poor relief and poverty eradication programs applied by their governments.

The World Bank classifies the following 60 countries as being part of the upper-middle-income group: Albania, Algeria, American Samoa, Argentina, Armenia, Azerbaijan, Belarus, Belize, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, China, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Equatorial Guinea, Fiji, Gabon, Georgia, Grenada, Guatemala, Guyana, Iran, Islamic Rep., Iraq, Jamaica, Jordan, Kazakhstan, Kosovo, Lebanon, Libya, Malaysia, Maldives, Marshall Islands, Mauritius, Mexico, Montenegro, Namibia, Nauru, North Macedonia, Paraguay, Peru, Romania, Russian Federation, Samoa, Serbia, South Africa, Sri Lanka, St. Lucia, St. Vincent and the Grenadines, Suriname, Thailand, Tonga, Turkey, Turkmenistan, Tuvalu, and Venezuela, RB. We note that this list is annually open to modifications (entry of a country or exit of a country from the list) as a function of the countries' annual economic performance. We take in consideration the list updated by January 2020 on the website of the World Bank[250]. However, due to a considerably high number of countries in this category, we will rather focus on a number of select upper-middle-income countries which are Argentina, Belize, Brazil, Colombia, Costa Rica, Dominican Republic, Ecuador, Mexico, Paraguay, and Peru from Latin America; Azerbaijan, Georgia, and Turkey from Central Asia and Europe; Botswana, and Republic of South Africa from Africa; Jordan, and Lebanon from Middle-East; and Malaysia, and Thailand from East Asia.

3.1 Socio-economic background of select upper-middle-income countries

In this section, we will see a list of macro level indicators for the countries categorized as upper-middle-income according to the World Bank criteria. The macro level indicators will be presenting here are the aggregate scores at country level reflecting different aspects of life such as unemployment, crime rate. We will use them as the complementary information to our empirical work, which will be proceeded by using micro data of national surveys.

By employing micro data of national surveys, we have access to information on all individuals of the representative sample in the data. Based on the available variables, and the indicators we create in the data we are able to find out the profile of these individuals, in other words we can reveal their living conditions in a multidimensional aspect. However, the creation of micro level indicators depends on the structure of the survey in other words on the availability of the questions in the survey questionnaire. As the living conditions are multiple in a country and intrinsically poverty requires a multidimensional aspect on it, in order to reflect also the aspects which cannot be examined with the surveys, we suggest that it is important to collect additional information on the situation of the country by employing macro level indicators which are already computed and published in the reliable statistical databases of the national statistical

[250] Please refer to the link below for further information (accessed 5 January 2020):
https://data.worldbank.org/income-level/upper-middle-income

offices and/or those of the international organizations such as the World Bank, and United Nations. The macro-level indicators can be, therefore, seen as the supplementary tools for reflecting living conditions of the society where the person lives, as a part of the social environment profile of the person.

The three tables shown below present some socio-economic indicators found at the UN and World Bank data bases for select 19 upper-middle-income countries. Table 17 includes 10 indicators: life expectancy at birth, infant mortality rate, education: primary gross enrollment ratio, education: secondary gross enrollment ratio, education: tertiary gross enrollment ratio, Intentional homicide rate, seats held by women in national parliaments, CO2 emission estimates, unemployment rate, and consumer price index, published by the UN data. Based on these statistics we can observe that: (i) South Africa has the lowest and Costa Rica has the highest life expectancy for both male and female given that life expectancy of female is higher than male across whole series. Botswana has also lower life expectancy for both female and male compared to other countries in the list except South Africa. Except Botswana and South Africa, other countries have similar life expectancies varying between 75 and 80 for female except for Costa Rica it is 81.7 and for Lebanon it is 80.6, and varying between 67 and 73 for male except for Costa Rica it is 76.7 and for Lebanon it is 76.9; (ii) concerning infant mortality, Azerbaijan, Botswana, Dominican Republic, South Africa, Jordan, Paraguay have higher figures, but also Belize, Brazil, Colombia, Ecuador, Mexico, and Peru have values higher than the average (13.68); (iii) in terms of primary education gross enrollment, Lebanon, Thailand, Jordan, Turkey, and South Africa have lower achievement compared to the series; on the other hand, Argentina, Belize, Brazil, Colombia, and Costa Rica have very good achievement above the average for both female and male. Regarding discrepancy in terms of gender, difference is above 6 points in favor of male in Belize, Dominican Republic, and Lebanon; (iv) in terms of secondary education gross enrollment, Belize, Dominican Republic, Jordan, Lebanon, Malaysia, Paraguay, and Peru have clearly lower values; on the other hand, Costa Rica, has the remarkably best performance but also Ecuador, Georgia, and Thailand have good performance compared to other countries in the list. Regarding discrepancy in terms of gender, except Turkey, Peru, Lebanon, female achievement is higher than male; (v) in terms of tertiary education gross enrollment, Argentina and Turkey have the best performances, and achievement of female is higher than male in all countries except Turkey; (vi) concerning crime rate, Belize and South Africa have the worst performance, but also Brazil, Colombia, Mexico, Botswana, Costa Rica, Dominican Republic have scores higher than 10; (vii) regarding seats held by women in national parliaments, Argentina, Costa Rica, Dominican Republic, Ecuador, Mexico, Peru, and South Africa have values higher than the average (22.45). Notably, Belize, Botswana, Lebanon, and Thailand's performances are under 10 per cent; (viii) from environmental perspective, concerning CO2 emission estimates, Belize, Colombia, Costa Rica, Dominican Republic, Georgia, Paraguay, and Peru have best scores, a maximum score of 2.2 in terms of tons per capita; (ix) as one of the most important economic indicator, South Africa has a significantly worst score, around 28 per cent of unemployment, but also Botswana and Jordan have also worse performances compared to the series. And unemployment rate in Brazil, Georgia, and Turkey are slightly higher than 10 per cent; (x) Concerning consumer price index which reflects inflation rate, Turkey has significantly worst performance higher than 200; however, except Belize, performances of other countries are not positive as well-notably, Argentina, Azerbaijan, Brazil, and South Africa have values higher than 150.

Based on SDGs data shown in Table 16, we can observe that: (i) the proportion of population with severe disabilities receiving disability cash benefit are quite few in case of Turkey, Peru, and Colombia; (ii) in fact there Is not much data available on the proportion of unemployed persons receiving unemployment cash benefit; however, it is quite low in case of Azerbaijan, Brazil, Colombia, and Dominican Republic; (iii) in terms of proportion of population using basic drinking water services in urban areas, per centages are quite high with at least 95 per cent; on the other hand, concerning rural areas, the performances of Azerbaijan, Botswana, Brazil, Colombia, Dominican Republic, Ecuador, Lebanon, Malaysia, Peru, South Africa need to be ameliorated; (iv) in terms of proportion of population using basic sanitation services in urban areas, notably Peru, South Africa, Dominican Republic, and Botswana need to ameliorate their situation; on the other hand, in case of rural areas, performances of Botswana, Brazil, and Peru are quite bad and except Costa Rica, Jordan, Lebanon, Malaysia, and Thailand, the conditions need to be ameliorated in other countries; (v) in terms of proportion of population with access to electricity in urban areas, Botswana remains clearly as the worst country and South Africa also needs to make more effort to increase her situation; on the other hand, concerning rural areas, the situation of Botswana is significantly deteriorated and Peru, and South Africa need to do more efforts to fill the gap between urban and rural areas; (vi) and concerning number of undernourished people, in Thailand 6.2 millions of people were found in this situation, in Mexico it is 4.8 millions of people, in Brazil it is less than 5.2 million, in South Africa it is 3.4 million, in Colombia it is 3.2 million, and in Peru it is 2.8 million.

Moreover, 11 indicators selected from the World Bank's world development indicators data base (see Table 18 below) show that: (i) in terms of access to clean fuels and technologies for cooking, in particular Botswana, Paraguay, Thailand, Peru, Georgia need to ameliorate their situation but also Belize, Mexico, and South Africa; (ii) in terms of adolescents out of school, there is no information for some countries; however, we can see that the rate is relatively considerable in case of Belize, Malaysia, Thailand, and Turkey for both female and male; (iii) in terms of adults (ages 15+) and children (ages 0-14) newly infected with HIV, the situation is significantly alarming in case of South Africa, but also Brazil, Botswana, and Mexico need to pay attention on that issue; (iv) in terms of coverage of social insurance programs, there is not much available information but Turkey has reported that 36 per cent of population have only coverage and in Ecuador it is 10.6 per cent; (v) Gini index aiming at reflecting inequality level in the country shows that scores are at middle level or slightly lower than it; (vi) inflation rates is quite striking in case of Turkey, South Africa, Brazil, and Colombia; (vii) the rate of informal employment are also considerable, notably in Ecuador, Paraguay, Peru, Thailand, Colombia, Dominican Republic; (viii) and given the available information on literacy rate among people aged 15 or over, performances of countries are quite positive.

Table 15. Some of Sustainable Development Goals (SDGs) indicators for select 19 upper-middle-income countries based on UN (United Nations) data base

UN data base, https://unstats.un.org/sdgs/indicators/database , 2016 numbers	Proportion of population with severe disabilities receiving disability cash benefit (%)	Proportion of unemployed persons receiving unemployment cash benefit (%)	Proportion of population using basic drinking water services URBAN areas (%)	Proportion of population using basic drinking water services RURAL areas (%)	Proportion of population using basic sanitation services URBAN areas (%)	Proportion of population using basic sanitation services RURAL areas (%)	Proportion of population with access to electricity URBAN areas (%)	Proportion of population with access to electricity RURAL areas (%)	Number of undernourish people (millions)
Argentina	8,2	7,2	99,6	93	95,85	76,77	99,98	99,7	1,7
Azerbaijan	100	1,6	99,1	80,4	96,44	87,6	100	100	<0,2
Belize	NA	NA	98,96	96,4	92,55	83,34	97,95	96,2	<0,1
Botswana	NA	31,5	96,8	73,7	88,44	49,68	77,86	23,98	0,6
Brazil	100	7,8	99,46	87,86	92	58,72	99,9	98,46	<5,2
Colombia	6	4,6	99,8	85,54	93	74,5	99,8	92,76	3,2
Costa Rica	NA	NA	99,74	99,5	98,2	95,2	99,8	98,45	0,2
Dominican Republic	NA	4,2	98	90	86,15	73,44	99,9	95	1,1
Ecuador	34.5	NA	100	82,67	90,56	81,3	99,8	96,78	1,3
Georgia	100	NA	99,96	95,7	95,15	83,52	100	100	0,3
Jordan	NA	NA	99,96	97,7	97,5	96,4	100	100	1,3
Lebanon	NA	NA	92,6 (all area)		96,93 (all area)		100	100	0,7
Malaysia	NA	NA	99,1	89,6	99,86	98,67	100	100	0,9
Mexico	NA	NA	99,76	95,4	92,74	80,77	99,8	98,33	4,8
Paraguay	21,6	NA	99,5	97	93,65	81	99,9	96	0,8
Peru	3,9	NA	95,34	74,15	79,43	54,54	98,9	77,97	2,8
South Africa	64,3	NA	98,87	79,8	76	72,6	92,97	67,66	3,4
Thailand	35,7	NA	99,7	99	98,6	98,17	99,92	100	6,2
Turkey	5	NA	98,6	99,6	99,8	89,86	100	100	<2

Note: Due to keyboard, computer configurations, point and comma are used interchangeably for mentioning decimals. Methodological notes concerning meta data information, definition of indicator, source of indicator can be found at the UN data base via the link mentioned in the table.

Table 16. Some socio-economic indicators for select 19 upper-middle-income countries based on UN (United Nations) data base

	Life expectancy at birth (females/males, years)	Infant mortality rate (per 1 000 live births)	Education: Primary gross enrol. ratio (f/m per 100 pop.)	Education: Secondary gross enrol. ratio (f/m per 100 pop.)	Education: Tertiary gross enrol. ratio (f/m per 100 pop.)	Intentional homicide rate (per 100 000 pop.)	Seats held by women in national parliaments (%)	CO2 emission estimates (million tons/tons per capita)	Unemployment (% of labour force)	CPI: Consumer Price Index (2010 =100)
UN data base, https://data.un.org/, 2019 numbers	*Social Indicators*							*Environment and infrastructure indicators*	*Economic indicators*	
Argentina	79.0 / 72.2	12,10	109.2 / 109.7	110.4 / 105.7	111.6 / 67.2	5,90	38,80	204.0 / 4.7	7.9%	152
Azerbaijan	74.6 / 68.6	31.4	104.2 / 102.5	NA	29.0 / 25.3	2,10	16,80	37.5 / 3.9	5,00%	153
Belize	76.0 / 71.1	15,00	110.3 / 116.1	86.2 / 82.2	30.5 / 18.9	37,60	9,40	0.5 / 1.4	9.3%	106
Botswana	67.0 / 61.8	35.5	103.5 / 107.2	NA	27.2 / 18.9	15 (2010 data)	9,50	7.0 / 3.2	17,80%	146
Brazil	78.0 / 70.7	15,80	112.4 / 115.4	103.9 / 99.2	58.8 / 42.5	29,50	15,00	529.8 / 2.6	11,60%	161
Colombia	78.9 / 73.1	14,40	110.7 / 114.6	101.3 / 96.0	65.0 / 56.0	25,50	18,70	84.1 / 1.8	9,20%	136
Costa Rica	81.7 / 76.7	9,30	110.3 / 109.7	129.5 / 122.8	62.2 / 49.3	11,90	45,60	7.8 / 1.6	7,90%	126
Dominican Republic	75.9 / 69.6	27,50	104.0 / 111.5	83.0 / 76.5	77.0 / 43.0	15,20	26,80	21.5 / 2.1	5,50%	133
Ecuador	78.5 / 72.8	15,10	104.7 / 103.7	108.6 / 105.6	49.0 / 42.2	5,90	38,00	43.9 / 2.8	4,20%	124
Georgia	76.8 / 67.8	12,90	103.3 / 102.3	108.1 / 106.5	61.3 / 54.0	1,00	14,80	9.0 / 2.2	11,60%	127
Jordan	75.5 / 72.2	17,10	100.0 / 99.0 (2005 data)	65.9 / 63.9	33.9 / 29.6	1,50	15,40	26.5 / 3.6	14,50%	125
Lebanon	80.6 / 76.9	9,60	89.5 / 97.3	63.0 / 63.6	45.8 / 39.6	4,00	4,70	24.1 / 4.3	6,70%	119
Malaysia	77.2 / 73.0	6,40	103.6 / 102.6	88.4 / 84.1	45.5 / 38.6	2,10	14,40	242.8 / 8.1	3,30%	121
Mexico	77.9 / 72.0	15,20	103.9 / 103.3	104.2 / 95.8	38.8 / 37.6	19,30	48,20	480.3 / 3.8	3,30%	137
Paraguay	75.2 / 71.3	20,80	104.3 / 107.6	79.1 / 74.2	41.2 / 29.1 (2010 data)	9,30	15,00	5.7 / 0.9	4,60%	140
Peru	77.7 / 72.6	14,10	102.6 / 102.3	98.3 / 98.5	73.6 / 65.8	7,70	30,00	61.7 / 2.0	3,70%	127
South Africa	64.2 / 57.8	32,90	100.3 / 104.2	104.9 / 96.0	24.0 / 17.1	34,00	42,80	489.8 / 9.1	27,60%	152
Thailand	78.9 / 71.6	10,10	99.5 / 99.7	114.3 / 119.1	57.7 / 41.1	3,20	5,40	316.2 / 4.7	1,40%	112
Turkey	78.7 / 72.4	11,80	100.9 / 101.6	101.7 / 104.2	96.5 / 110.7	4,30	17,40	346.0 / 4.5	11,20%	202

Note: Due to keyboard, computer configurations, point and comma are used interchangeably for mentioning decimals. Methodological notes concerning meta data information, definition of indicator, source of indicator can be found at the UN data base via the link mentioned in the table.

Based on the figures shown in Tables 16, 17 and 18, we can say that the 19 upper-middle-income countries manifest in different deprivations regarding various variables such as access to drinking water, electricity, sanitation facilities, educational attainment. However, considerable inequality and inflation rates (or consumer price index) are common features affecting people living in such countries. But also, low level of social insurance coverage, high level of informal employment, worst living conditions in terms of access to basic services in rural areas are other frequent disadvantages concerning majority of population in these countries. Notably, HIV, undernourishment, no access to drinking water, sanitation facilities emerge substantially in South Africa, and Botswana as well as South Africa faces high level of unemployment and security issues. Even if these countries are part of upper-middle-income group, we can observe that they face also different types of deprivations. In multidimensional poverty analysis based on micro national data, it is also important to gather macro level figures

which are computed on different source. These indicators can be seen as a complementary information, which reflect societal situation which might correspond to achievement or deprivation. Societal achievements and non-achievements will certainly have an impact on citizens' life bringing them advantages or disadvantages in their lifestyle or decision-making process. In other words, societal facts play role in shaping capabilities of citizens.

If we look at specifically on comparison of Brazil, Mexico, South Africa, and Turkey regarding these indicators, we can point out that: (1) life expectancy is around 78 for women and 71 for male in Brazil, Mexico, and Turkey whilst in South Africa it is 64 for female and 58 for male; (2) infant mortality is the least in Turkey almost 3 times higher in South Africa; (3) regarding primary school enrollment, Brazil has the best score, regarding secondary school enrollment male have lower score compared to female in Brazil, Mexico, and South Africa but in Turkey the enrollment is slightly higher for male, and regarding tertiary school enrollment female enrollment is relatively higher in Brazil and South Africa and inverse in Turkey; (4) crime rate in Turkey is considerably low compared to Brazil, Mexico, and South Africa; (5) in terms of seats held by women in national parliaments, the performances of Mexico and South Africa are much higher than Brazil and Turkey; (6) regarding CO2 emission, South Africa has the worst score and Brazil has the best; (7) unemployment rate is quite similar in Brazil and Turkey; however it is very high in South Africa and it is considerably low in Mexico; (8) consumer price index is greatly high in Turkey but still it is an important problem for other 3 countries as well; (9) proportion of population with severe disabilities receiving disability cash benefit is very low in Turkey vs. Brazil has a great performance on that; (10) the proportion of population using basic drinking water services in urban areas are quite similar in 4 countries; (11) regarding the proportion of population using basic drinking water services in rural areas, Turkey has a great score with no significant difference between urban and rural areas, Mexico is also close to Turkey in this aspect; however, in South Africa discrepancy between urban and rural areas is the highest and Brazil also witness such case with a slightly better performance compared to South Africa; (12) regarding the proportion of population using basic sanitation services in urban areas, Turkey has a great performance whilst South Africa has considerably the worst. When it comes to rural areas, Turkey still has the best performance with around 10 per cent of difference in favour of urban areas. In particular, discrepancy becomes significant in Brazil with around 33 per centage points difference in favour of urban areas. In Mexico, the difference is about 12 per centage points in favour of urban areas and in South Africa, in both urban and rural areas between 72 and 76 per centage of population have access to sanitation services; (13) regarding the proportion of population with access to electricity in urban areas, the 4 countries have very good scores given that South Africa is the worst. On the other hand, difference between urban and rural areas is only significant in South Africa where 68 per cent of rural population have only access to electricity; (14) regarding undernourishment, Turkey has the best score; but still in Brazil, Mexico, and South Africa it is a considerable problem; (15) regarding access to clean fuels and technologies for cooking, Mexico and South Africa need to make efforts for ameliorating the situation; (16) in terms of educational attainment among adolescents, Turkey has the worst performance with around 10 per cent of school dropout among adolescents in lower secondary school age; (17) regarding adults (ages 15+) and children (ages 0-14) newly infected with HIV, the situation of South Africa is alarming; (18) Turkey has the best score for Gini index around 42, then Mexico comes with 48.3 and then Brazil with 53.7; (19) in terms of inflation, Brazil has the worst score with 8.7 per cent, then Turkey comes with 7.8 per cent, then South Africa with 6.6 per cent, and South Africa with 2.8 per cent; (20)

literacy rate among people aged 15 or over is the best in Turkey; however, the performances of Brazil and Mexico are very good as well in that aspect.

Table 17. Some of World Development indicators for select 19 upper-middle-income countries based on Word Bank data base

World Bank data base, https://databank.worldbank.org/source/world-development-indicators# , 2016 numbers	Access to clean fuels and technologies for cooking (% of population)	Adolescents out of school, female (% of female lower secondary school age)	Adolescents out of school, male (% of male lower secondary school age)	Adults (ages 15+) and children (ages 0-14) newly infected with HIV	Children in employment, female (% of female children ages 7-14)	Children in employment, male (% of male children ages 7-14)	Coverage of social insurance programs (% of population)	GINI index (World Bank estimate)	Inflation, consumer prices (annual %)	Informal employment (% of total non-agricultural employment)	Literacy rate, adult total (% of people ages 15 and above)
Argentina	98,40	0,10	0,10	6500	NA	NA	NA	42	NA	NA	99,10
Azerbaijan	95,50	7,20	4,90	NA	NA	NA	NA	NA	12,40	NA	99,80
Belize	85,40	13,10	11,60	500	NA	NA	NA	NA	NA	NA	NA
Botswana	64,10	NA	NA	9700	NA	NA	NA	NA	2,80	NA	NA
Brazil	95,60	4,10	3,30	50000	NA	NA	NA	53,7	8,70	NA	92,80
Colombia	91,80	7,80	8,90	7600	NA	NA	NA	50,8	7,50	57,40	94,70
Costa Rica	93,50	5,90	5,00	970	0,60	1,90	NA	48,7	0,00	35,80	NA
Dominican Republic	90,40	6,70	4,70	2800	NA	NA	NA	45,7	1,60	53,20	93,80
Ecuador	95,60	NA	NA	2400	NA	NA	10,60	45	1,70	63,40	94,40
Georgia	77,80	NA	NA	720	NA	NA	NA	36,6	NA	NA	NA
Jordan	99,10	NA	NA	100	0,40	2,00	NA	NA	-0,80	NA	NA
Lebanon	NA	NA	NA	200	NA	NA	NA	NA	-0,80	NA	NA
Malaysia	96,30	12,10	15,40	5400	NA	NA	NA	NA	2,10	NA	93,70
Mexico	85,40	2,70	4,00	11000	NA	NA	NA	48,3	2,80	NA	94,90
Paraguay	66,20	NA	NA	1200	NA	NA	NA	47,9	4,10	65,80	94,70
Peru	75,10	3,30	3,50	3500	NA	NA	NA	43,6	3,60	58,40	94,20
South Africa	84,80	NA	NA	280000	NA	NA	NA	NA	6,60	34,30	NA
Thailand	74,40	12,30	13,10	8300	NA	NA	NA	36,9	0,20	62,20	NA
Turkey	NA	10,50	9,10	NA	NA	NA	36,10	41,9	7,80	NA	96,20

Note: Due to keyboard, computer configurations, point and comma are used interchangeably for mentioning decimals. Methodological notes concerning meta data information, definition of indicator, source of indicator can be found at the UN data base via the link mentioned in the table.

3.2 National poverty measurement practices in select upper-middle-income countries

The table in Appendix 9.8 includes the names and websites of the national statistical agencies of select upper-middle-income countries mentioned in introduction[251]. A snapshot on the following topics will be yielded respectively: (i) different types of surveys (data sources) used to measure poverty; (ii) multidimensional poverty approaches; (iii) surveys used for measuring multidimensional poverty; (iv) multidimensional vs. one-dimensional approaches to poverty; (v) surveys used for measuring monetary poverty; and (vi) publication of poverty estimates related to Sustainable Development Goals of the UN (SDGs). We will then finalize this section by focusing in particular on Brazil, Mexico, South Africa, Turkey, and the countries employing a multidimensional approach to measure poverty including Argentina, Colombia, Costa Rica, Ecuador in terms of different dimensions and indicators used in the construction of multidimensional poverty index or dashboard[252].

Among the select 19 upper-middle-income countries, to measure poverty, a number of different surveys are used: household expenditure survey, household budget survey, living standards measurement survey, household income and expenditure survey, welfare indicators survey,

[251] There are in fact 60 upper-middle-income countries; however, statistical offices of some countries do not provide a comprehensive information such as the case of small population islands, or websites of some of them are not yet sufficiently developed for divulging detailed information, or the access to some of them are limited, or there is no information in English, French, Spanish, or Portuguese. We did not integrate European Union countries and some other Eastern European countries that use the standard EuroStat procedure that we already explained in the literature review; except we show the case of Turkey as she is part of our empirical analysis.

[252] We won't provide methodological information for Peru as the official document available on the website of the INEI (National statistical and informatics agency) is quite old, published in 2000.

multi-topic household survey, general household survey, living conditions survey, income and living conditions survey, census, labour force survey, household income and basic amenities survey, and socio-economic survey. From a multidimensional perspective, Argentina and Peru have opted for the Unmeet Basic Needs (UBN) approach; Costa Rica, Mexico and South Africa have employed the Multidimensional Poverty Index (MPI) approach based on Sen's capabilities theoretical framework and Alkire and Foster's methodology (2007, rev. 2009b, 2011a); and Ecuador and Colombia have measured poverty by employing both UBN and MPI approaches. However, it is important to note that Colombia, Costa Rica, and Mexico are the 3 countries have officially recognized multidimensional poverty (but they generate also monetary poverty outputs); otherwise, the other 4 countries still use monetary poverty measure as the main official poverty statistics and the multidimensional one for providing complementary information on living conditions. Census, socio-economic survey, living conditions survey, general household survey are often used to measure multidimensional poverty.

On the other hand, the other 12 countries currently use monetary based poverty measure(s): absolute poverty thresholds such as the World Bank's 1 or 2 USD per day per person (Brazil, Botswana, Georgia, Turkey); EuroStat's relative poverty thresholds such as 60 per cent of median national equivalized disposable household income (Georgia, Turkey); consumption expenditure patterns based income poverty lines (Botswana, Colombia, Costa Rica, Ecuador, Mexico, Paraguay, Peru) and consumption expenditure thresholds (Argentina, Belize, Dominican Republic, Jordan, Lebanon, South Africa Republic, Thailand, Turkey). As we can see some countries produce more than one type of monetary poverty measures such as Turkey. Concerning consumption expenditure-based methodologies, the basket of food is determined to find out the rate of extremely poor people and the basket of food and non-food such as including clothing, shelter is determined to compute the rate of poor people in the country. Income poverty lines are generally updated as a function of inflation rate and/or adjusted for price differences across the regions within the country. Household expenditure survey, living conditions and income survey, household income and expenditure survey, household budget survey are often used to measure monetary poverty. The information on multidimensional poverty measurement is generally better documented in terms of coherent and systematic content for construction of the measure compared to monetary metric measurement. In the case of Azerbaijan, and Malaysia, income poverty statistics were only made available on their website without any detailed explanation on the methodology. Moreover, Botswana, Brazil, Costa Rica, Ecuador, Georgia, Malaysia, South Africa Republic, and Turkey show also the SDG related statistics. It becomes important to emphasize 3 issues: first, as we can see from the results above the practice of multidimensional poverty at national level is very limited despite a growing number of works in the academic area[253]; second, sociological theories are not used for explaining poverty and/or for constructing a measure by national statistical institutes/agencies according to the documents/press releases on their websites[254]; third, there

[253] United Nations tries to promote the Global Multidimensional Poverty Index (based on Alkire and Foster methodology and Sen's capability approach) in their annual human development report; nevertheless, country cases require specific look on the selection of indicators and dimensions according to socio-economic context of the country. For that reason, we can say that the use of multidimensional poverty by national statistical agencies is still significantly low among upper-middle-income countries.
[254] Statistical agencies of some countries provide a brief note on their website or in official document or press release of poverty statistics about a number of most used theoretical concepts on poverty such as

is no agreement about in which category the poverty statistics enter such as whether it is social, or household income/expenditure, or living conditions, or inequality statistics given the way poverty statistics are shown on the websites of national statistical agencies.

The national statistical agency of ***Argentina*** (INDEC) measure also poverty based on the unmeet basic needs-UBN (called also Necesidades Básicas Insatisfechas-NBI)[255]. INDEC mentions that the main goals are finding out population and households with unsatisfied basic needs, explaining structural poverty using National Census of Population, Households and Housing, and they consider NBI as the accumulated lack of material deprivations an alternative tool besides income poverty. There are 5 conditions of NBI given that if a household does not fulfill at least 1 of them then it is considered poor/unmeet basic needs. The conditions are as follows: (NBI 1) *Housing*: households living in tenant rooms, hotels or pensions, homes not intended for residential purposes, precarious housing and other types of housing. House, apartment and ranch type homes are excluded; (NBI 2) *Sanitary conditions*: households living in a home not having a toilet; (NBI 3) *Overcrowding*: households living in an overcrowded home-more than three people per room in the home; (NBI 4) *School attendance*: households that have at least one school-age child (6 to 12 years old) who does not attend school; (NBI 5) *Subsistence capacity*: there is four or more people per employed member in the household and the head of the household has not completed the third grade of primary schooling. According to the estimations obtained by census 2001 and 2010, it is found that the per centage of people living with unmeet basic needs across the country is decreased from 17.7 to 12.5 between in 2001 and 2010 (i.e., the per centage of households living with unmeet basic needs across the country is decreased from 14.3 to 9.1 between 2001 and 2010).

The national statistical office of ***Brazil***, Instituto Brasileiro de Geografia e Estatistica (IBGE), provide poverty statistics on their website under the topic of multi-domain statistics focusing specifically on living conditions, poverty, and inequality. It is reported that IBGE measures poverty from a monetary approach in the following way: first, the line of $ 1.90 per day, converted by the purchasing power parity (PPC-2011) of $ 1.66 per dollar; second, the value of the 2011 line is adjusted for inflation according to the deflators used by the Continuous PNAD survey (IBGE, Pesquisa Nacional por Amostra de Domicílos Contínua); third, people with per capita household income below the poverty line are considered poor which give the head count ratio of poverty across the country. And the calculations exclude persons in the condition of a pensioner, domestic employee or relative of a domestic employee. For instance, in 2017, it is found that 7.4 per cent of Brazilian population lives under the poverty line[256].

In case of ***Mexico***, the CONEVAL (Consejo Nacional de Evaluacion de la Politica de Desarrollo Social) carries out the computation of multidimensional poverty in collaboration

basic needs, capabilities approach. In case of multidimensional poverty index, theoretical grounds of capabilities approach are briefly mentioned and/or human rights related framework as well.

[255] Please refer to the link below for further information (accessed 9 January 2020) :
https://www.indec.gob.ar/indec/web/Nivel4-Tema-4-47-156
From the literature, we know that it is an important step to conceptualization and measurement of poverty from a multidimensional aspect notably thanks to greatly significant works of CEPAL (see for instance-accessed 9 January 2019, https://www.cepal.org/es/publicaciones/4784-metodo-necesidades-basicas-insatisfechas-nbi-sus-aplicaciones-america-latina).

[256] Please refer to the link below for further information (accessed 7 January 2020) :
https://sidra.ibge.gov.br/tabela/6682#resultado

with INEGI (Instituto Nacional de Estadistica y Geografía)[257]. We will explain the main points of the CONEVAL's methodology for measuring multidimensional poverty in Mexico based on their official document published in 2010[258] (we already explained it briefly in multidimensional poverty section of the theoretical framework). The Mexican Multidimensional Poverty Measure (MMPM onwards) includes 2 main dimensions: economic well-being, and social rights across 8 indicators: current per capita income, average educational gap in the household, access to health services, access to social security, quality and spaces of the dwelling, access to basic services in the dwelling, access to food, degree of social cohesion. The MMPM follows two fundamental steps: identification of the poor, and aggregation of the poor to obtain a macro-level indicator (e.g., the number of poor people in the country) according to the Sen's recommendations (1976). For the identification of the poor regarding the *economic well-being* space, "the standard monetary poverty method" is used "which compares per capita income with the monetary value of two baskets, one for food and other that includes, besides food, any other good or service usually consumed by the Mexican families" (CONEVAL 2010, 12). On the other hand, for the identification of the poor regarding the *social rights* aspect, the 3 following criteria are defined: "(i) deprivations had to be measured by means of dichotomic variables, in the sense that the right is either met or not met, because it does not exist an intermediate value on the fulfillment of any right; hence, there is not an ordinal scale; (ii) no right is superior to any other; therefore, all of them should ―have the same value, in other words, given a linear combination of indicators, all of them should be equally weighted; and (iii), the deprivation of any right makes a person socially deprived" (CONEVAL 2010, 12). To strengthen the choice of using equal weighting, the CONEVAL arguments that (2010, 30) "defining the social deprivation index as the sum of the six deprivation indicators assumes that every deprivation has the same relative importance. This feature, as well as the decision to adopt one single deprivation as the deprivation threshold, are based on the principles of indivisibility and interdependence of human rights, which recognize that the non-fulfillment of any of the human rights infringes upon the fulfillment of the others, and that there is no human right that can be considered more important than any other". Based on these criteria, the index of social deprivation is constructed by counting the deprivations experienced by the person regarding the 7 indicators of social rights. Herewith, the "number of unfulfilled rights for any person" is found thanks to this index (CONEVAL 2010, 12). After that, poverty lines are determined for both two dimensions, economic well-being, and social rights: (i) "the population with an income insufficient to acquire the goods and services required to satisfy its needs[259]" are considered poor regarding the economic well-being[260]; (ii) "the population deprived in at least one of the

[257] Please refer to the link below for further information (accessed 7 January 2020) :
https://www.coneval.org.mx/Medicion/MP/Paginas/Pobreza-2018.aspx

[258] Please refer to the link below for further information (accessed 7 January 2020):
https://www.coneval.org.mx/rw/resource/coneval/med_pobreza/MPMMPingles100903.pdf
And related other documents in English can be found in the link below (accessed 7 January 2020):
https://www.coneval.org.mx/Medicion/Paginas/English-version.aspx

[259] including both food and non-food needs (CONEVAL 2010, 29).

[260] The poverty line for the dimension of economic well-being is called "the wellbeing threshold" (CONEVAL 2010, 29); this is a minimum income level that every single individual should have it for not being considered deprived regarding economic well-being, or in other words for being able to afford the essential goods and services. In addition, to identify the extreme deprivation in economic well-being, another threshold called "the minimum wellbeing" is determined for identifying people who have a sufficient income to buy food given that it does not allow to buy nutritional (adequate) food (CONEVAL 2010, 29).

six social indicators" are considered poor regarding social rights[261] (CONEVAL 2010, 29). In this context, CONEVAL has decided that a person situating under the poverty line regarding both 2 dimensions is *multidimensionally poor*[262]. Moreover, a person is considered *vulnerable due to social deprivation* if at least one of her/his social rights is not guaranteed but at the same time he or she is not poor regarding economic well-being; a person is considered *vulnerable due to income* if he or she has not a sufficient income level to satisfy economic well-being but at the same time he or she is not poor regarding social rights; a person is considered living in *extreme multidimensional poverty*, if he or she has a low income level even this amount allows him or her to buy food but not nutritional food for living a healthy life and is deprived regarding at least 3 of 6 social deprivations; and a person who is multidimensionally poor but not in extreme multidimensional poverty situation is considered living in a *moderate level of multidimensional poverty* (CONEVAL 2010, 32-33). At a final stage, as a part of the aggregation process, the headcount ratio of poverty at country level (i.e., aggregate measure of poverty) is obtained by counting multidimensionally poor people and dividing this sum by the total population of the country. The pilot work on the MMPM is carried out by employing the Socioeconomic Conditions Module of the National Survey of Household Income and Expenditures, 2008 (MCS-ENIGH 2008)[263]. It is important to note that even if social cohesion is part of the social rights dimension, the methodology proposed by the CONEVAL does not yet include it. Concerning the unit of analysis, it is the individual; however, household level information is applied to every single individual living in the household in case individual information cannot be obtained (i.e., some questions in the survey are asked to each individual but some of them are asked at household or dwelling levels).

Now, we will look at more specifically on the content of indicators used in social deprivation index of the MMPM. In the table below, the first column mentions the indicator and the second columns describes the conditions of deprivation regarding the indicator. The information mentioned in the table comes from the CONEVAL document, the description of the deprivation cases and the name of indicators are directly taken from the official document (2010, 43-45):

[261] Given that there are 6 indicators in social rights dimension, to find out who is deprived in at least 1 of the 6 social indicators the CONEVAL methodology determines a *deprivation threshold* for each indicator. For instance, if the person is reported not having access to health services then he or she is considered deprived regarding health services, and he or she takes a score of 1 in case of deprivation. This means that a person has got maximum a score of 6 (the person's social deprivation index score) in case of deprivation in all 6 indicators (total deprivation regarding social rights) and minimum a score of 0 in case of zero deprivation in all 6 indicators at the same time. As we mentioned the CONEVAL has decided that a person is considered poor regarding social rights if w/he has at least 1 of 6 deprivation in this dimension (i.e., her/his social deprivation index score should be at least equal 1), thus we can see that the final deprivation threshold is set at 1 to determine who is poor regarding social rights- if a person's social deprivation index score is equal or greater than 1, then he or she is considered poor regarding social rights.

[262] CONEVAL (2010, 28) states that "a person is considered to be multidimensional poor when the exercise of at least one of her social rights is not guaranteed and if she also has an income that is insufficient to buy the goods and services required to fully satisfy her needs".

[263] CONEVAL (2010, 52) mentions that "the MCSENIGH 2008 makes it possible to obtain information on income, sociodemographic characteristics and indicators for measuring multidimensional poverty at the national and state scale".

Table 18. Indicators and deprivation thresholds of the social deprivation index of the MMPM

Educational gap	▪ For people aged three to fifteen years old. When they lack mandatory basic education and are not attending a formal educational center. ▪ For people born before 1982. If they do not meet the minimum mandatory basic education level that prevailed at the time, they should have attended elementary school. ▪ For people born from 1982 onwards. If they have not completed the minimum current mandatory basic education requirement (secondary school).
Access to health services	▪ She is not enrolled in or entitled to receive medical services from any institution offering them, including the Seguro Popular, the social security public institutions (IMSS, federal or state ISSSTE, 25 Pemex, 26 Army or Navy) or private medical services.
Access to social security	▪ For those who are economically active, it is considered that they are not deprived in this dimension if, through their job, they enjoy the benefits established in the law ▪ For non-paid workers and for workers on their own, given the voluntary nature of enrollment in the system, they are considered to have access to social security when they receive medical services as a job benefit or when they are actually voluntarily enrolled and when, in addition, they bear a retirement plan. ▪ For the general population, people are considered to have access when they benefit from a retirement program or pension or when one of their relatives has access to social security. ▪ In the case of people in retirement ages (sixty-five years or older), people are considered to have access to social security if they benefit from a social pension program for senior citizens. ▪ The population not meeting any of the above criteria is considered deprived due to access to social security
Quality and spaces of the dwelling	▪ If the dwelling has dirt floor. ▪ If the roof of the dwelling is made of cardboard sheets or waste. ▪ If the walls of the dwelling are made of mud or daub and wattle; reed, bamboo or palm tree; cardboard, metal or asbestos sheets; or waste.

	▪ The ratio of the number of members of the household per room is greater than 2.5 (overcrowding).
Access to basic services in the dwelling	▪ Water is obtained from a well, river, lake, stream, or truck; or when piped water is carried from another dwelling or gotten at a public faucet or hydrant. ▪ There is no drainage service, or when the drainage is connected to pipes leading to a river, lake, sea, ravine or crack. ▪ The dwelling has no electricity. ▪ Wood or coal with no chimney are used for cooking or heating food inside the dwelling.
Access to food	▪ There are four possible levels of food insecurity: severe food insecurity, moderate food insecurity, mild food insecurity and food security. Households are considered deprived in access to food if they present *moderate* or *severe* food insecurity[264].

Source: CONEVAL (2010).

The CONEVAL has published the multidimensional poverty results in Mexico between 2008 and 2018. Some major highlights are the following[265]: regarding the 2008 data, 33.4 per cent of the Mexican population are found multidimensionally poor, 11 per cent are living in extreme multidimensional poverty, 4.7 per cent are vulnerable due to income, 32.3 per cent are vulnerable due to social deprivation, and 18.7 per cent are neither multidimensionally poor nor vulnerable whilst regarding the 2018 data, 34.5 per cent of the Mexican population are found multidimensionally poor, 7.4 per cent are living in extreme multidimensional poverty, 6.9 per cent are vulnerable due to income, 29.3 per cent are vulnerable due to social deprivation, and 21.9 per cent are neither multidimensionally poor nor vulnerable. In terms of indicators of the social deprivation index, the per centage of people experiencing educational gap is dropped from 21.9 to 16.9, the per centage of people of suffering deprivation in access to health services is dropped from 38.4 to 16.2, the per centage of people suffering deprivation in access to social security is dropped from 65 to 57.3, the per centage of people suffering deprivation in quality and spaces of the dwelling is dropped from 19.7 to 13.8, the per centage of people suffering deprivation in access to food is decreased from 21.7 to 20.4, and the per centage of people suffering deprivation in access to basic services in the dwelling is dropped from 22.9 to 19.8 between 2008 and 2018. Moreover, the per centage of people experiencing at least one social deprivation is decreased from 76.6 to 70.4 per cent, and the per centage of people experiencing at least 3 social deprivations is decreased from 31.7 to 18.8 per cent between 2008 and 2018. On the other hand, concerning the economic aspect of poverty, the per centage of people living under the economic well-being threshold is slightly dropped from 49 to 48.8 per cent-almost

[264] However, the four degrees of food insecurity are reported along with the other 2 food indicators for reflecting different aspects of the fulfillment of this social right reported by the CONEVAL (2010, 51).
[265] Please refer to the link below for further information (accessed 7 January 2020):
https://www.coneval.org.mx/Medicion/MP/Paginas/Pobreza-2018.aspx

the same tendency over 10 years of period, and the per centage of people living in extreme monetary poverty remained at 16.8 per cent between 2008 and 2018[266].

In addition to the official multidimensional poverty measure, the CONEVAL provides also monetary poverty results by employing two income poverty lines: extreme poverty line income, and income poverty line[267]. The former is based on the food basket per person a month whilst the latter is based on the food and non-food basket per person a month.

Colombia's national statistical agency, DANE, measures multidimensional poverty in 2 different ways: unmeet basic needs (Necesidades basicas insatisfechas-NBI), and multidimensional poverty index. We will start off by the first one widely used in Latin America following the CEPAL's powerful initiatives. DANE qualifies NBI as a tool to find out the share of Colombian population lacking basic needs; for that, the following indicators are used[268]:

- *inadequate housing*: if physical characteristics of housing considered inappropriate for human accommodation (e.g., not adequate wall, floor materials);
- *housing with critical overcrowding*: housing with more than three people per room (excluding kitchen, bathroom and garage);
- *housing with inadequate services*: the lack of access to minimum vital and sanitary conditions;
- *housing with high economic dependence*: in case if there are more than three people per employed member in the household and the head of the household has a maximum of two years of approved primary education;
- *housing with school-age children who do not attend school*: households with at least one child over 6 years old and under 12, relative of the head of the household and not attending a formal education center.

A household experiencing deprivation regarding at least 1 of 5 indicators listed above is considered unmeet basic needs, and a household experiencing deprivation regarding at least 2 of 5 indicators is considered in a miserable situation. Regarding 2018 estimates based on National Census of Population and Housing (Censo Nacional de Población y Vivienda-CNPV), 14.13 per cent of the Colombian population lives with unmeet basic needs.

DANE produces also both monetary and multidimensional poverty measures under the thematic of poverty and inequality. According to methodological issues described in the official document prepared by National Council of Economic and Social Policy (Consejo Nacional de Política Económica y Social-CONPES) under the National Planning Department (Departamento Nacional de Planeación) published in 2012, we will show some key points of

[266] The press release from CONEVAL on the results of the poverty measurement of 2012 can be found in the link below (accessed 7 January 2020):
https://www.coneval.org.mx/Informes/COMUNICADOS_DE_PRENSA/2013/COMUNICADO_PRENSA_003_MEDICION_2012_ING.pdf

[267] Please refer to the link below for further information (accessed 8 January 2020) :
https://www.coneval.org.mx/Medicion/MP/Paginas/Lineas-de-bienestar-y-canasta-basica.aspx

[268] Please refer to the link below for further information (accessed 9 January 2020):
https://www.dane.gov.co/index.php/en/statistics-by-topic-1/poverty-and-life-conditions/necesidades-basicas-insatisfechas-nbi
The detailed information on the indicators can be found in the link above. And the document in the link below presents the key points of NBI (accessed 9 January 2020):
https://www.dane.gov.co/files/censos/resultados/prest_NBI_100708.pdf

multidimensional and monetary measures respectively[269]. The Colombian Multidimensional Poverty Index (CMPI) is based on Alkire and Foster methodology (2007, rev. 2009b, 2011a) including 5 dimensions and 15 indicators in total as follows[270]:

Table 19. Dimension and indicators of the Colombian Multidimensional Poverty Index (CMPI)

Dimension	**Indicator(s)**
Education	▪ Low educational attainment, ▪ Illiteracy
Childhood and youth conditions	▪ School absence, ▪ School lag, ▪ Barriers to access services for early childhood care, ▪ Child labour
Employment	▪ Long-term unemployment ▪ Informal employment
Health	▪ Without health insurance ▪ Access barriers to health service
Home public services and housing conditions	▪ No access to improved water source ▪ Inadequate disposal of excreta ▪ Inadequate floors ▪ Inadequate exterior walls ▪ Critical overcrowding

On the other hand, DANE provides two main monetary poverty measures: poverty rate and extreme poverty rate. People are classified as poor if their average income per month is less than the value of the poverty line; and a person is identified as extreme poor if her/his average monthly income is less than the value of the extreme poverty line. The value of the extreme poverty line represents the monthly cost per capita necessary to acquire the basic food basket that covers the minimum nutritional requirements according to ICBF (Instituto Colombiano de Bienestar Familiar) and FAO (Food and Agriculture Organization) standards; and the value of the poverty line is equivalent to the monthly cost per capita necessary to acquire, in addition to food, other basic goods and services. Between 2016 and 2018, the per centage of multidimensionnally poor people is increased from 17.8 to 19.6. But also, in 2018, monetary poverty was 27 and extreme poverty was 7.2[271].

Costa Rica's national statistical agency (INEC) computes multidimensional poverty index based on the Alkire and Foster methodology (2007, rev. 2009b, 2011a), as a result of the initiative of the Ministry of Human Development and Social Inclusion of the Council Social Presidential (Ministerio de Desarrollo Humano e Inclusión Social, del Consejo Presidencial

[269] Please refer to the link below for further information (accessed 9 January 2020):
https://www.dane.gov.co/files/acerca/Normatividad/CONPES-150.pdf

[270] We already explained in detail this methodology in multidimensional poverty sub-section in the literature review.

[271] Please refer to the links below for further information (accessed 9 January 2020):
https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/2018/bt_pobreza_multidimensional_18.pdf,
https://www.dane.gov.co/files/investigaciones/condiciones_vida/pobreza/2018/bt_pobreza_monetaria_18.pdf

Social) - led by the Second Vice Presidency of the Republic - and the Horizon Association Positive (Asociación Horizonte Positivo), as a complementary measure to traditional monetary metric poverty statistics[272]. But also, the approach of unmeet basic needs (Necesidades basicas insatisfechas-NBI) is accepted as an official multidimensional poverty measure since 2000. The first attempt to NBI was realized by employing Population and Housing Census of the year 2000 (Censo de Población y Vivienda). For that, 4 dimensions were defined: access to decent shelter/housing, access to healthy living, access to knowledge and access to other goods and services, as recommended by CEPAL (see Méndez and Trejos 2004). NBI is again computed in 2011 by using Census. The results have shown that the per centage of people lacking at least one basic need is decreased from 36 to 24.6 per cent between 2000 and 2011. The computation of Costa Rican Multidimensional Poverty Index (CRMPI) has been officialized by INEC in 2015. The National Household Survey (Encuesta Nacional de Hogares-ENAHO) is used for this purpose. The index is constructed upon the 5 dimensions: education, health, housing and internet use, labour/work and social protection. The table below shows the short description of indicators per dimension (the full information on the indicators can be found in INEC (2015, 17-38)[273]:

Table 20. Dimensions and indicators of the Costa Rican Multidimensional Poverty Index (CRMPI)

Education	▪ No participation to formal education, ▪ Educational lag, ▪ No achievement of high school, ▪ Underdevelopment of human capital
Health	▪ No health insurance, ▪ No access to water, ▪ No toilet (no drainage) ▪ No garbage (removal of rubbish)
Housing and internet use	▪ Bad state of roof or floor, ▪ Bad condition of the exterior walls, ▪ Overcrowding, ▪ No use of internet
Labour/work	▪ Long term unemployment or discouraged people, ▪ No compliance of labour rights (breach of minimum salary, other rights), ▪ Performing independent informal job
Social protection	▪ No childcare for first child, ▪ Elderly without pension, ▪ People with disability no receiving transfers, ▪ Outside the workforce due to family obligations

[272] Information here concerning Costa Rican Multidimensional Poverty Index and UBN comes from the official document of INEC (2015). Please refer to the link below for further information (accessed 9 January 2020):
http://www.inec.go.cr/sites/default/files/documentos/pobreza_y_presupuesto_de_hogares/pobreza/metodologias/mepobrezaenaho2015-01.pdf

[273] Please refer to the link below for further information (accessed 9 January 2020):
http://www.inec.go.cr/sites/default/files/documentos/pobreza_y_presupuesto_de_hogares/pobreza/metodologias/mepobrezaenaho2015-01.pdf

According to the computations based on the National Household Survey 2018 (Encuesta Nacional de Hogares), 19 per cent of households were found living in multidimensional poverty.

On the other hand, concerning the official monetary measure of poverty, INEC measure the 3 following groups: (i) households with extreme poverty: those households with a per capita income equal to or less than the cost per capita of the Basic Food Basket (CBA) or extreme poverty line; (ii) households in non-extreme poverty: those households that have equal or lower per capita income to the poverty line but higher than the per capita cost of the CBA or extreme poverty line; (iii) non-poor households: those households with a per capita income that allow them to cover their basic food and non-food needs (i.e., income per capita is greater than value of the poverty line). For instance, in 2018, 14.72 per cent of the population was living under the poverty line, and 6.34 per cent was living under the extreme poverty line (based on the National Household Survey 2018).

The national statistical agency of ***Ecuador*** (Instituto Nacional de Estatística y Censos-INEC) provides 3 different measures of poverty: monetary metric (income poverty, consumption poverty), unmeet basic needs (NBI), and multidimensional poverty index. For the first and third, the National Survey of Employment, Unemployment and Under-Employment (Encuesta Nacional de Empleo, Desempleo y Subempleo) and Living Conditions Survey (Encuesta de Condiciones de Vida); for the second, Population and Housing Census (Censos de Población y Vivienda) and Living Conditions Survey are used. For computing poverty based on the unmeet basic needs (NBI), the following dimensions and indicators are used[274]:

Table 21. Ecuador's Unmeet Basic Needs (NBI) dimensions and indicators

Dimension	**Indicator(s)**
Economic capacity	Household is considered deprived in this dimension if: ▪ The head of the household has less or maximum 2 years of schooling; *and* ▪ There are more than 3 people to take care of per each employed person in the household
Access to basic education	Household is considered deprived in this dimension if: ▪ Child(ren) aged between 6 and 12 yrs. old in the household do not attend school
Access to housing/dwelling (housing quality)	Household is considered deprived in this dimension if: ▪ The floor material of the dwelling is land/dust/soil/mud or another not adequate material making dirt the floor; *or*

[274] Please refer to the links below for further information (accessed 10 January 2020) : https://www.ecuadorencifras.gob.ec/pobreza-por-necesidades-basicas-insatisfechas/, and https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2018/Diciembre-2018/201812_Pobreza.pdf

	▪ The wall material of the dwelling is cane or another similar material
Access to basic services (sanitation services in the dwelling)	Household is considered deprived in this dimension if: ▪ Dwelling does not have sanitation services, or if the quality of them is not adequate (e.g., latrine); *or* ▪ There is no piped water coming from the general distribution network
Overcrowding	Household is considered deprived in this dimension if: ▪ The ratio of people per bedroom is greater than 3

A person is considered poor/unmeet basic needs if he or she lives in a household deprived regarding at least 1 of 5 dimensions. In december 2018, 33.5 per cent of the population was poor in terms of umeet basic needs.

The Ecuador Multidimensional Poverty Index (ECMPI) is based on the Alkire and Foster methodlogy (2007, rev. 2009b, 2011a) and contains 4 main dimensions and 12 indicators (see the table below[275]):

Table 22. Ecuador Multidimensional Poverty Index (ECMPI) dimensions and indicators

Dimension	**Indicator(s)**
Education	▪ Absence from basic and high school education, ▪ No access to higher education because of financial issues, ▪ Incomplete education
Work and social security	▪ Child labour, ▪ Unemploymen, or inappriate employment, ▪ No contribution to the pension system
Health, water, and food	▪ Extreme poverty, financial incapacity ▪ No water service provided by public network
Housing and healthy environment	▪ Overcrowding, ▪ Housing deficit, ▪ No toilet, no drainage ▪ No collection of rubbish

[275] Please refer to the link below for further information (accessed 10 January 2020, see page 11): https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2017/Pobreza_Multidimensional/ipm-metodologia-oficial.pdf. The description of indicators can be found in the document available in the link above. Please see also, https://www.ecuadorencifras.gob.ec/documentos/web-inec/POBREZA/2018/Diciembre-2018/201812_Pobreza.pdf

The two main statistica are computed based on this methdoology: Multidimensional Poverty Rate (MPR): percentage of people living in households that have deprivations in a third or more of the weighted indicators; and Extreme Multidimensional Poverty Rate (TPEM): percentage of people living in households that have deprivation in at least half of the weighted indicators. In december 2018, 37.9 per cent of the population was multidimensionally poor.

In terms of monetary poverty, individuals whose total income per capita is less than poverty line are considered poor[276]. The poverty line in 2006 was US $ 56.64 per month per person, and the extreme poverty line in 2006 was US $ 31.92 per month per person. Currently, these poverty lines were upadted regulary as a function of consumption price index. As of December 2018, poverty rate at the national level was 23.2% and extreme poverty rate was 8.4%.

According to the latest presse release (April 2019) of Statistics ***South Africa*** (Stats SA)[277], poverty is measured based on the 2014/15 Living Conditions Survey (LCS) and its statistics are published in the report of Men, Women and Children: Findings of the Living Conditions Survey 2014/15[278]. It is found that 49.2 per cent of the adult population (18 yrs. or over) live below the upper-bound poverty line (UBPL). In terms of gender, 46.1 per cent of adult males vs. 52 per cent of adult females experience poverty based on the UBPL. The latest statistical release (2019) on the poverty lines used by Statistics South Africa explain in detail the way poverty is measured[279.] There are 3 official poverty lines determined by Statistics South Africa: (i) food poverty line/extreme poverty line: the level of income the person needs for being able to afford the "minimum required daily energy intake" ; (ii) lower-bound poverty line (LBPL): "the food poverty line plus the average amount derived from non-food items of households whose total expenditure is equal to the food poverty line"; and (iii) upper-bound poverty line (UBPL): "the food poverty line plus the average amount derived from non-food items of households whose food expenditure is equal to the food poverty line" (see Statistics South Africa 2019, 3). These 'inflation-adjusted national poverty lines' (per person per month in Rands) are annually updated by using Consumer Price Index for taking in consideration variations in cost of living/purchasing power (i.e., price changes for goods and services). The methodology behind the construction of these poverty lines is based on the cost-of-basic-needs approach taking into account household's consumption expenditure on both food and non-food items. The recent poverty lines are inflation adjusted on the values obtained from the Income and Expenditure Survey (IES) 2010/11[280].

Stats SA has also attempted to compute poverty from a multidimensional perspective since 2000. The first report was Measuring poverty in South Africa (Statistics South Africa 2000) which includes a household infrastructure index and a household circumstances index created

[276] For technical information, please refer to the document in the link below (accessed 10 January 2020): https://www.ecuadorencifras.gob.ec//documentos/web-inec/POBREZA/Metodologia+de+pobreza+por+ingresos.pdf

[277] Please refer to the link below for further information (accessed 7 January 2020): http://www.statssa.gov.za/?p=12075

[278] The report is available in the link below (accessed 7 January 2020): http://www.statssa.gov.za/publications/Report-03-10-02%20/Report-03-10-02%202015.pdf

[279] Please refer to the link below for further information (accessed 7 January 2020): http://www.statssa.gov.za/publications/P03101/P031012019.pdf

[280] Statistics South Africa published the country's official national poverty lines in 2012 (Statistics South Africa 2019, 6).

by employing Census 1996 data[281]. The second report was Provincial Indices of Multiple Deprivation (Statistics South Africa 2006)[282] aimed at showing some indices of deprivation at provincial level and a ward-level analysis was carried out for each province including income and material, employment, education, health, and living environment aspects of life by employing Census 2001 data. The third report was Men, Women and Children (2013) which presented a multidimensional poverty index based on the Bristol method across 8 dimensions using Living Conditions Survey (LCS)2008/09 data (see Statistics South Africa 2014, 2). And, the most recent is The South African MPI: Creating a Multidimensional Poverty Index (Statistics South Africa 2014)[283]. The major goal of the creation of this index (called SAMPI) was to provide poverty profiles at provincial and municipal levels. The SAMPI is based on the Alkire and Foster methodology (2007, rev. 2009b, 2011a) and computed by using Census 2001 and Census 2011 data. It is important to note that multidimensional poverty index is not the official South African measure and so far, another report did not release. Statistics South Africa mention that "the SAMPI is not intended to replace the poverty headcounts and gaps produced by Stats SA using the three national poverty lines that have been developed (i.e., the food poverty line, the lower-bound poverty line and the upper-bound poverty line). As mentioned above, the SAMPI should rather be seen as a complementary measure to these money-metric measures" (Statistics South Africa 2014, 8).

The SAMPI is composed of 11 indicators across 4 dimensions (health[284], education, standard of living, economic activity). The table below shows the description of the indicators and their deprivation thresholds. The descriptions are directly taken from the document (Statistics South Africa 2014, 6).

Table 23. Indicators and deprivation thresholds of the SAMPI

Child mortality (health)	▪ If any child under the age of 5 has died in the past 12 months
Years of schooling (education)	▪ If no household member aged 15 or older has completed 5 years of schooling
School attendance (education)	▪ If any school-aged child (aged 7 to 15) is out of school
Fuel for lighting (standard of living)	▪ If household is using paraffin/candles/nothing/other
Fuel for heating (standard of living)	▪ If household is using paraffin/wood/coal/dung/other/none

[281] The document is available in the link below (accessed 7 January 2020):
https://www.gov.za/sites/default/files/gcis_document/201409/povertyrep.pdf
[282] The document is available in the link below (accessed 7 January 2020):
http://www.statssa.gov.za/census/census_2001/PMID/PIMDReport2006.pdf
[283] The document can be found in the link below (accessed 7 January 2020):
http://www.statssa.gov.za/publications/Report-03-10-08/Report-03-10-082014.pdf
We already mentioned the South African MPI (SAMPI onwards) in multidimensional poverty sub-section of the literature review. We will just provide here some complementary information.
[284] In the report (Statistics South Africa 2014, 8) it is mentioned that as the census questionnaire does not include the questions related to nutrition, food security, hunger, then the actual SAMPI is then considered providing a relatively limited information on the health dimension.

Fuel for cooking (standard of living)	▪ If household is using paraffin/wood/coal/dung/other/none
Water access (standard of living)	▪ If no piped water in dwelling or on stand
Sanitation type (standard of living)	▪ If not a flush toilet
Dwelling type (standard of living)	▪ If an informal shack/traditional dwelling/caravan/tent/other
Asset ownership (standard of living)	▪ If household does not own more than one of radio, television, telephone or refrigerator and does not own a car
Unemployment (economic activity)	▪ If all adults (aged 15 to 64) in the household are unemployed

Source: Statistics South Africa (2014, 6).

The per centage of multidimensionally poor people (i.e., the head count ratio) is decreased from 17.9 to 8 between 2001 and 2011. Besides, the adjusted headcount ratio, namely the SAMPI (the headcount ratio adjusted by the intensity of poverty) is decreased from 0.08 to 0.03 between 2001 and 2011 (Statistics South Africa 2014, 9).

The national statistics institute of ***Turkey*** provides detailed information on poverty statistics in terms of both methodology and estimations on their website. According to that information, we see that poverty is computed in 2 different ways based on income and expenditure patterns via the basket of food and non-food items[285] by employing Household Budget Survey (HBS) from 2002 and Statistics on Income and Living Conditions (SILC) from 2006. In a relative deprivation perspective by employing SILC, 4 poverty thresholds are determined: 40, 50, 60, 70 percent of the national median equivalized household disposable income. Between 2008 and 2018, the per centage of poor people (having an income level under 60 per cent of the national median equivalized household disposable income) is decreased from 24.1 to 21.1 per cent. When the price level differences/purchasing power differences between regions are taken in consideration, the per centage of people having an income level under 60 per cent of the national median equivalized household disposable income is decreased from 22.8 to 20.5 per cent between 2008 and 2018. On the other hand, concerning poverty statistics based on consumption expenditure patterns, the basic needs approach is employed. In that vein, the Turkish Statistical Institute has computed poverty in 2 different ways from 2002: (i) below 2,15 $ per capita per day, and (ii) below 4,3 $ per capita per day (the current PPP values in TL are used for the equivalents of 1 $). According to the first criteria, poverty rate is decreased from 0.47 to 0.06 per cent between 2008 and 2015; and according to the second criteria, poverty rate is decreased from 6.83 to 1.58 per cent between 2008 and 2015.

3.3 National poor relief and poverty eradication programs in South Africa, Turkey, Mexico, and Brazil

In previous section, we saw the latest poverty measures adopted at national level in upper-middle-income countries. In this section, we will more specifically focus on the four upper-middle-income countries, namely South Africa, Turkey, Mexico, Brazil which are included in our empirical analysis. The goal is two-fold: exploring national poor relief (social welfare) and

[285] The link is (accessed 7 January 2020): http://www.turkstat.gov.tr/PreTablo.do?alt_id=1013

poverty eradication programs implemented by the governments of South Africa, Turkey, Mexico and Brazil since the Millennium era. This information will help us first to understand at conceptual level which approach of poverty is used, second to understand at measurement level which target group and dimensions of poverty are used in national poverty reduction and poverty relief programs to identify the poor, as a function of the coherent available information on these issues. We will, therefore, be able to observe the evolution of these programs in terms of their content. We note that discussions in terms of efficiency of the results of programs and technical details concerning the management of programs on selection of the target groups, money transfers, and link between government and municipalities are not part of this section.

South Africa

In the history of South Africa, as many facts and studies have already shown, apartheid has played a crucial role in social, economic, political, institutional organization of the country, which also influenced the structure of social surveys and poverty measurement techniques[286]. Apartheid[287] (i.e., separate development) is the name that used for describing the legalized racial discrimination and segregation-based political regime applied by the government between 1948 and 1994, in which the priority in many aspects was given to minority White people (i.e., White supremacy) in opposition to Black African, Colored, and Asian South Africans (i.e., racial classification)[288]. Schlemmer and Møller (1997, 16) has described apartheid as "'re-fragmentation' of the country into the white-controlled core society and independent or semi-autonomous reserves or 'Bantustans' in the less-developed areas of subsistence economy". For instance, by the Population Registration Act of 1950 each South African citizen should be registered regarding her/his race (four racial groups were determined[289]), by the Prohibition of Mixed Marriages Act of 1949 it was banned marriage between a White and a Non-White, by the Reservation of Separate Amenities Act of 1953 some public premises and facilities including transports were exclusively reserved for White people. Moreover, by some other acts such as the Native Labour Act of 1953, the Bantu Education Act of 1953, the Extension of University Education Act of 1959, the Group Areas Act of 1950, the Natives Land Act of 1913, the Bantu Homelands Citizenship Act of 1970, for the Non-White population, political, civil rights, education, employment, land-tenure opportunities were become limited and this population was forced to live in specific areas determined by the government such as Colored townships (e.g., South Western Township (Soweto) near Johannesburg in Gauteng province), Bantustans. Taylor (2005, 149) has stated that "historically, the black majority experienced exclusions from occupational welfare (in the form of unemployment insurance benefits,

[286] For a detailed history of poverty in South Africa between 1855 and 2005, see Davie (2015). Crais (2011), Hanoman (2018), Wilson and Ramphele (1989) provide also historical information regarding poverty and hunger in South Africa.

[287] Afrikaans: "apartness," or racial separation. Please refer to the link below for further information (accessed 12 September 2019): https://www.britannica.com/place/South-Africa

[288] It is important to note that before the apartheid, since the Dutch colonialism and then under the British and Afrikaners governments, racial inequality was in fact imposed by different acts into the South African society. For instance, the Mines and Works Act No 12 of 1911, the Natives Land Act of 1913, the Native Trust and Land Act of 1936, the Minimum Wages Act of 1925, the Mines ad Works Act of 1926 were entered into force before the apartheid (see Hanoman 2018).

[289] First, three racial groups were determined that are Black (or Bantu referring all black Africans), White, and Colored (i.e., mixed race). Asian (Indian and Pakistani) were added later as another category. Please refer to the link below for further information (accessed 12 September 2019): https://www.britannica.com/topic/apartheid

minimum decent wages, health care and retirement pensions), from direct welfare provision of public goods and services (such as education, healthcare and housing) and from welfare, which includes income-support measures in the form of cash benefits through the tax system"[290]. In a similar way, Møller (1998, 29) has mentioned that "the social mobility of the black sector of the population was severely restricted as a result of inferior 'bantu' education, reservation of non-menial jobs, influx controls which regulated access to centers of employment, and restrictions on entrepreneurship and access to capital markets". Moreover, studies have showed that Black population has experienced lower (subjective) well-being and higher material deprivation compared to White population (Møller 2007, 390). According to 1983 numbers, half of the Black population were found happy or very happy vs. 93 per cent of the White population were happy or very happy (Møller 2013, 922). As we can see Non-White people were clearly discriminated, their human rights and liberty were clearly violated by the South African government which promoted domination of White in economic, social, political, institutional aspects in the country. However, at international platform, against this racial discrimination, for instance the United Nations Declaration on the Elimination of All Forms of Racial Discrimination has recognized that "…any doctrine of racial differentiation or superiority is scientifically false, morally condemnable, socially unjust and dangerous, and that there is no justification for racial discrimination either in theory or in practice…"[291].

During the apartheid, in social surveying and poverty measurement, the *Poverty Datum Line* (PDL onwards) was one of the most known practices in South Africa. Prof. Edward Batson (former student of Bowley in London School of Economics, but also influenced by the works of Rowntree) from the University of Cape Town is considered as the pioneer for using the PDL in poverty measurement (see Batson 1941, 1945, 1950[292]). For instance, in July to September 1944, Batson carried out a social survey in Salisbury, Southern Rhodesia, to measure poverty line. At that time, Non-white people were generally blamed by the government for being poor because of their behaviors. In opposite, based on different living items, Batson has showed that Non-White population were also suffered poverty as other poor White population. Regarding Batson's technique, in 1966, Watts has measured poverty in the cities of Johannesburg, Pretoria, and Durban, and in the towns of Benoni, Springs, Witbank, and Greytown[293]. Watts (1967, ii) has described the PDL as an estimation of "lowest possible cost for maintaining a household in health and decency under Western conditions, in the short run only. Food, clothing, cleansing materials, and fuel and lighting are taken into account, bearing in mind the varying needs of persons of different ages and sexes". Influenced by Rowntree's poverty concept, Watts also has distinguished primary and secondary poverty. In this line, his methodology consists of summing each household member's cost of living with the household's total cost of fuel and lighting, for calculating *primary* PDL for the household. At second stage, the cost of rent, transport fees from home to workplace for workers are added to the primary PDL to obtain the *secondary*

[290] Based on Titmuss 1969 [1958].
[291] Please refer to the link below for the items (accessed 26 January 2019):
http://www.un-documents.net/a18r1904.htm
[292] Please refer to the link below (accessed 26 January 2019):
http://www.historicalpapers.wits.ac.za/inventories/inv_pdfo/AD1715/AD1715-14-14-24-001-jpeg.pdf.
See also Davie (2015, chapter 3).
[293] Please refer to the link below for the documents (accessed 26 January 2019):
https://opendocs.ids.ac.uk/opendocs/bitstream/handle/123456789/5353/Watts_1967_Poverty%20datum%20line.pdf?sequence=1

PDL for the household. Watts has mentioned that his findings are not related to racial grouping because he assumed that minimum cost of living is theoretically the same for everyone regardless of her/his race. On the other side, in 1978 and 1979, the team of Møller and Schlemmer has conducted a quality of life survey in Durban, which is a very important city in terms of trade and sugar industry (see Møller and Schlemmer 1983, 232). The results have showed that 93 per cent of White, 88 per cent of Indian, but only 41 per cent of Black South Africans were satisfied with their life in general. This rich survey in terms of content has focused on many aspects of life satisfaction (i.e., "Would you think of your life as a whole, everything in it that makes up your life, and tell me how satisfied or dissatisfied you are with your life. On the whole would you say you were satisfied, just satisfied, not satisfied or dissatisfied, dissatisfied, very dissatisfied?"), happiness about life, manifold issues of concern (e.g., family health, own health, love relationship, own education, education for children, wealth, job security, accessible community services, housing supply, self confidence).

May[294] (2010, 1) has mentioned that based on first representative national survey (i.e., income and expenditure survey) carried out in 1993 just one year before the end of apartheid in collaboration with the World Bank and the SALDRU (Southern Africa Labour and Development Research Unit, University of Cape Town), 52 per cent of all African households living in rural areas were poor (see also Carter and May 1999, Klasen 1997, Wilson 1996)- the poverty line was settled at about a monthly US$250 at the time for a family with two adults and three children living in an urban area (cited in Seekings and Nattrass, 2). More dramatically, in the regions Eastern Cape and Northern Transvaal, about two-thirds of the population were living in poverty (Whiteford et al. 1995, cited in Seekings and Nattrass, 2). Based on the same data, Van der Berg (1997, 482) has mentioned that in 1993, 38.3 per cent of Black, 20.8 per cent of Colored, 11.3 per cent of Indian population in South Africa were unemployed; however, this rate was significantly lower for White population with 4.3 per cent. Regarding personal income per capita, the most advantaged group was the White with R25 344 whilst the Black people was receiving on average R2 520 and the Colored population was receiving R5 316[295] (Van der Berg 1997, 482). The income of the most well-off 10 per cent of households was corresponding to at least 50 per cent of national income (Seekings and Nattrass 2015, 4, see also Whiteford et al. 1995). In addition, 57 per cent of Black population, 20 per cent of Colored population, 7 per cent of Indian[296] population against only 2 per cent of White population were living in poverty (Seekings and Nattrass 2015, 3). Klasen (2000) has measured deprivation through the 14 aspects of life (education, income, wealth, housing, water, sanitation, energy, employment, transport, financial services, nutrition, health care, safety, and perceived well-being) by employing the SALDRU 1993 data[297].This study has showed that 49.7 per cent of the

[294] Please refer to the link below for the access to the document:
https://pdfs.semanticscholar.org/801f/ffcec9ba3ea9425252996db4e36453635f3b.pdf

[295] Regarding Bloomberg's exchange rates, currently, 1 USD is equal R13.62 (South African Rand). Please refer to the link below for further information (accessed 27 January 2019):
https://www.bloomberg.com/markets/currencies

[296] In some other sources, Indian population was included also among the Colored population.

[297] Indicators are: average years of schooling of all adult (16+) household members, expenditure quintiles, number of household durables including vehicles, phone, radio, TV, geyser, stoves, kettle, bicycles), housing characteristic, type of water access, type of sanitation facilities, main source of energy for cooking, share of adult members of households employed, type of transport used to get to work, ratio of monthly debt service to total debt stock, share of children stunted in household, use of health facilities during last illness, perception of safety inside and outside of house, compared to 5 years ago, and level

population is deprived vs. 25.4 per cent of the population is most deprived; such per centages led to the fact that around 75 per cent of the population suffer various capabilities deficit. Qizilbash (2002) has measured poverty in a multidimensional aspect through the 7 indicators (household expenditures, educational attainment, water source, the frequency or type of refuse removal, energy source used for cooking, the number of rooms per household, and employment) based on Cerioli and Zali (1990) and Cheli and Lemmi (1995) fuzzy set approaches by employing 1996 South Africa Census. Each employed indicator has 5 categories; for instance, the categories of electricity, gas, coal/paraffin, dung, and wood are used as the 5 different energy sources for cooking. And each category is associated with a poverty degree number as a result of a particular fuzzy set method varying from 0 to 1 where 0 represents non-poor, 1 represent definitely poor, and between 0 and 1 represent different degrees of vulnerability. In this work, he particularly showed disparities across provinces; for instance, 63.5 per cent of North Province are found poor regarding energy source for cooking, 41 per cent of Eastern Cape are found poor regarding water source, 44.61 per cent of KwaZulu Natal are found poor regarding job situation.

The apartheid regime was not only characterized by racial inequality in terms of poverty, child mortality, life expectancy, educational attainment, land possession, income and employment[298], but also in terms of social policy including limited access to social welfare programs for Non-White population. But after the apartheid, South African social security system had begun to gradually be modified in favor of Non-White population for reaching egalitarian benefit levels and access for all citizens no matter race. For instance, the item 27 of the Act No. 108 of the Constitution of the Republic of the South Africa of 1996[299] states that "(1) Everyone has the right to have access to health care services, including reproductive health care; sufficient food and water; and social security, including, if they are unable to support themselves and their dependents, appropriate social assistance, (2) the State must take reasonable legislative and other measures, within its available resources, to achieve the progressive realization of each of these rights, (3) No one may be refused emergency medical treatment".

The first development oriented social welfare program after the end of apartheid[300] was the *Reconstruction and Development Program* (RDP onwards) which was launched by the first Black president Nelson Rolihlahla Mandela's political party the African National Congress

---

of satisfaction of household. Klasen (2000, 40) attributed a score to each indicator from 1 to 5 (1 is for the most deprived and 5 is for the least deprived), and then constructed the deprivation index (most deprived people have an index value less or equal 3 vs. the less deprived people have an index value superior than 3). Each indicator has 5 categories within; for instance, for the main source of energy used for cooking, wood has a score 1, dung has a score 2, paraffin or coal has a score 3, gas from bottle, dry battery has a score 4, and electricity from grid, town gas has a score 5. See also Klasen (1997) for the comprehensive results of SALDRU survey in terms of poverty, inequality and deprivation.

[298] Regarding Whiteford et al. 1995, May et al. 2000, and Seekings and Nattrass 2005, 4): minority White population was living approximately 10 years more than Black and Colored population, the probability of Black and Colored children to die by the age of one was seven times higher compared to White children, and 70 per cent of Black and Colored adolescents born between 1966 and 1975 did not get their high school diploma against only 20 per cent of White adolescents.

[299] Please refer to the link below for the full Act:
https://www.gov.za/sites/default/files/images/a108-96.pdfs

[300] See Taylor (2014; 2016, 151-153) for the list of social policies and their main goals in the context of social assistance grants, undertaken from 1994.

(ANC) in 1994[301]. The main aim of RDP was to improve socio-economic conditions in the country, especially reducing poverty by increasing social assistance programs with a larger cover targeting the Non-White population who suffered from the apartheid (see Corder 1997, Wessels 1999). After the election of 1994, Mandela and Nobel peace laureate Archbishop Desmond Tutu have expressed the importance of becoming "rainbow nation" for South Africa, as of understanding diversity and respect of every citizen's right regardless race, origin (Møller 2013, 918). During his speech at the 2005 Campaign to Make Poverty History in Trafalgar Square, Mandela has expressed his view on poverty as follows: "Overcoming poverty is not a task of charity, it is an act of justice. Like slavery and apartheid, poverty is not natural. It is man-made and it can be overcome and eradicated by the actions of human beings"[302]. In RDP, five major development policy axes and six main goals are defined. The axes are: "meeting basic needs, developing human resources, building the economy, democratizing the State and society, and implementing the RDP" (Hanoman 2018, 9). Notably, a number of social programs are planed such as "school feeding schemes, child support grants, free health services for children between 0-6 years, for pregnant and lactating women, pension funds for the elderly, working for water, community public works programs, Provincial community food garden initiatives like Kgora and Xoshindlala, land reform and farmer settlement, production loans scheme for small farmers, infrastructure grant for smallholder farmers and the Presidential tractor mechanization scheme" (Republic of South Africa 2002, 5[303]). Concerning basic needs, providing food security nationwide was considered the priority in the program. In 1996, in the new Constitution of the Republic of South Africa, access to healthy and sufficient food and water was considered as a fundamental right to each citizen in the country (Republic of South Africa 1996, Provision 27[304]). But also, the Taylor Report[305] (called also the Report of the Committee of Inquiry into a Comprehensive System of Social Security for South Africa) published in 2002 has proposed a comprehensive social protection programme considering a diverse range of social assistance topics such as health, unemployment, poverty, to implement efficient social policies in the context of empowering national security in favor of the disadvantaged citizens in South Africa. The recommendations of this report were then adopted by the South African government.

[301] Mandela mentioned the aim of the RDP during his speech to a Joint Sitting of Parliament, 24 May 1994 (Preface to the RDP White Paper 1994) as follows: "My Government's commitment to create a people-centered society of liberty binds us to the pursuit of the goals of freedom from want, freedom from hunger, freedom from deprivation, freedom from ignorance, freedom from suppression and freedom from fear. These freedoms are fundamental to the guarantee of human dignity. They will therefore constitute part of the centerpiece of what this Government will seek to achieve, the focal point on which our attention will be continuously focused. The things we have said constitute the true meaning, the justification and the purpose of the Reconstruction and Development Programme, without which it would lose all legitimacy".
Please refer to the link below for further information (accessed 15 May 2020): https://www.gov.za/sites/default/files/governmentgazetteid16085.pdf

[302] Please refer to the link below for further information (accessed 29 January 2019): http://www.mandela.gov.za/mandela_speeches/2005/050203_poverty.htm

[303] Please refer to the link below for further information (accessed 2 February 2019): https://www.gov.za/sites/default/files/gcis_document/201409/foodpol0.pdf

[304] Please refer to the link below for further information (accessed 29 January 2019): https://www.wipo.int/edocs/lexdocs/laws/en/za/za107en.pdf

[305] Please refer to the link below for further information (accessed 29 January 2019): https://sarpn.org/CountryPovertyPapers/SouthAfrica/taylor/

At the same year, in the framework of Growth, Employment and Redistribution Strategy (GEAR) and in accordance with the 1996 World Food Summit accords, the *Food Security Policy* is developed by the South African government, for ensuring "access to adequate, safe and nutritious food" (Food Security Working Group 1997, pt. 1.1). Later, in 2002, the *Integrated Food Security Strategy (IFSS)* was launched to "attain universal physical, social and economic access to sufficient, safe and nutritious food by all South African at all times to meet their dietary and food preferences for an active and healthy life" (Republic of South Africa 2002, 13[306]) with the main objectives of increasing household food production, income, new job opportunities, nutrition and food safety, food social allowance. Most recently, in 2013, in the framework of National Development Plan (NDP)[307], the *National Policy on Food and Nutrition Security* together with *The Fetsa Tlala Integrated Food Production Initiative* was established to focus on increasing food production and access of it for households, communities and notably for small-scale farmers as a critical element for reducing poverty, unemployment, inequality by 2030 (DAFF 2013). Meanwhile, there is also a specific *National School Nutrition Program* providing one healthy meal to students in primary and secondary schools[308]. However, due to massive unemployment, specifically long-term unemployment, a considerable number of South Africans suffer food deprivation; and despite economically not active status, vulnerable unemployed 18-59 yrs. old (and medically not disabled) people do not receive cash transfers because they are not eligible to receive cash help (Taylor 2016, 147). Many studies have shown that food insecurity is likely to be happened in rural areas, highly populated urban areas, and places where Africans were severely suffered by the apartheid (cited in Hanoman 2018, see also Twine and Hunter 2011, Crush et al. 2010). In addition, Hanoman (2018, 12-13) has emphasized that food insecurity problem has also occurred in one of the breadbaskets of the country, the Free State Province[309], where poverty rate was found at 61.9 per cent according to Statistics South Africa 2014 which was higher than the national poverty level (i.e., 56.8 per cent), and 22.7 per cent of the population were deprived of adequate access to food according to the Department of Agriculture, Forestry and Fisheries (DAFF)[310]. In 2012, the Department of Agriculture and Rural Development of the Free State Province has reported that "poverty, inadequate provision and maintenance of basic infrastructure, informal housing and insecurity

[306] Please refer to the link below for further information (accessed 2 February 2019): https://www.gov.za/sites/default/files/gcis_document/201409/foodpol0.pdf

[307] Please refer to the link below for further information (accessed 2 February 2019): https://www.gov.za/issues/national-development-plan-2030

[308] Please refer to the link below for further information (accessed 2 February 2019): https://www.education.gov.za/Programmes/NationalSchoolNutritionProgramme.aspx

[309] Hanoman (2018) states that still in the Free State Province, typical apartheid era agriculture structure, namely, White farm owner and Black farm labourer, did not change so much. Racial dynamics appear omnipresent; for instance, the Black and Colored population do still prefer live in their townships and for who are in prosperous situation they do not want settle in wealthier White area because of insecurity they feel in White neighborhoods. Author, therefore, suggests taking in consideration such situations (e.g., racial socio-economic analysis) including race crimes in implementation of food security and agriculture policies in the country.

[310] The DAFF has several on-going programs such as the African Agricultural Development Program (AADP), Agricultural Broad-Based Black Economic Empowerment, Comprehensive Agricultural Support Program, Integrated food Security and Nutrition Program, Vulnerable Workers on Farms, in Forestry and Fisheries Program. Please refer to the link below for further information (accessed 2 February 2019): https://www.daff.gov.za/daffweb3/Programme/Integrated-food-Security-and-Nutrition-Programme

of tenure, inadequate public transport, lack of economic opportunities, high illiteracy and innumeracy, non-payment of services, droughts, HIV/AIDS, unemployment and crime" (cited in Hanoman 2018, 51)[311] were the main problems witnessed by habitants in the province. Such situation shows clearly multidimensional side of poverty and complex deprivation status.

Furthermore, in the framework of national Department of Public Works, the specific goals were determined regarding housing, access to water, access to electricity, health care, land reform, and job creating through public works for diminishing poverty (Mfusi and Govender 2015, 119). Regarding *housing*, the aim was "providing well-located and affordable shelter for all by 2003 and building one million houses in five years", regarding *water* the aim was "supplying 20 to 30l of clean water each day to every person in two years and 50 to 60; per day within five years from a point no more than 200 meters from each dwelling", regarding *electricity* the aim was "supplying 2.5 million more households and all schools and clinics with electricity by the year 2000", regarding *health care* the aim was "giving free medical care to children under the age of six and to homeless children; improve maternity care for women; organize programs to prevent and treat major diseases like TB and AIDS", regarding *land reform* the aim was "implementing land reform based on redistribution of residential and productive land to those who need it but cannot afford it and restitution to those who lost land because of apartheid laws", and regarding *job creation through public works* the aim was "introducing a national public works program to provide basic needs such as water supply, sewerage and roads and at the same time create jobs, particularly in poor and rural areas". For diminishing unemployment, several programs are developed concerning both urban and rural areas such as the *Integrated Sustainable Rural Development Program* (ISRDP onwards) applied in selected 18 rural areas, the *Urban Renewal Program* (URP onwards) applied in selected 8 urban townships for increasing access to basic services, ameliorating infrastructure, increasing connection to neighborhoods by public transports and roads, providing adequate housing, creating new economic activities for the disadvantaged population suffered from the governance of apartheid regime, increasing security, promoting education, training and skills development[312].

In the early 2000s, the government has attempted to identify the poor areas (called as *Nodal* areas) which are mainly based on 'second economy' or 'the marginalized economy' for living. During his speech to National Council of Provinces in November 2003, the former President Thabo Mbeki has defined these areas as underdeveloped places "contributing little to the GDP, absorbing the largest percentage of the country's population, incorporating the poorest of the rural and urban poor, structurally disconnected from both the First World and the global

[311] For instance, see the Agricultural land and projects/programme/support profile for effective planning and implementation, in the district of Thabo Mofutsanyana. Retrieved from www.ard.fs.gov.za/files/fs-profile/Thabo-Mofutsanyana.pdf (accessed 2 February 2019).

[312] In fact, during the apartheid regime, (i) no economic and infrastructural development was done for rural areas due to the Group Areas Act, No. 36 of 1966; and (ii) the Black and Colored population in urban areas were placed at townships, which were highly populated and were lacking adequate access to basic services. Especially, after the end of apartheid regime, a huge migration from rural to urban areas was happened for finding new jobs in the cities, which also increased the population of townships (see Statistics South Africa 2016a in the link below, accessed 2 February 2019):
http://www.statssa.gov.za/publications/Report-03-10-14/Report-03-10-142011.pdf
But also see the document by Statistics South Africa 2016b in the link below (accessed 2 February 2019):
http://www.statssa.gov.za/publications/Report-03-10-13/Report-03-10-132011.pdf

economy, and incapable of self-generated growth" (see Republic of South Africa 2007b, 5). For diminishing poverty, government has particularly focused on the elimination of 'second economy'. Notably, unemployment and underemployment have been important issues; for instance, Møller (2007a) has mentioned that "South Africa's sound macro-economic policy has not created sufficient jobs to absorb the new entrants into the labour market each year. More than two-thirds of South Africans between 18 and 35 years are unemployed. The official unemployment rate that excludes discouraged workseekers stood at 26.7% in early 2006, in spite of the 658,000 jobs that were created in the year to September 2005". The *Accelerated and Shared Growth Initiative for South Africa* (ASGISA onwards) which was launched by Deputy President Phumzile Mlambo-Ngcuka in 2006, has aimed at increasing the importance of first economy in the market, promoting small business, halving unemployment and poverty between 2004 and 2014. Thus, specific measures for women and youth population have been taken in consideration. Concerning women, the main goal was to increase their economic opportunities including new human resource training, better access to finance, more active participation in agriculture and creative industries, better access to basic needs, more active participation in the Expanded Public Works Program (EPWP). Concerning youth population, government has targeted to diminish unemployment rate of graduates by providing them jobs and learnerships[313]. In this line, in 2010, the government with the initiative of the Economic Development Minister Ebrahim Patel has launched the *New Growth Path* (NGP) aiming at promoting inclusive economic growth, providing equity, creating five million jobs and therefore reducing unemployment rate to under the 15 per cent by 2020. The program has targeted specifically six sectors for new job opportunities, which are infrastructure (i.e., energy, transport, communication, water, housing), agriculture (i.e., ameliorating living and working conditions of 660 000 farm workers), mining, manufacturing, and the green economy (i.e., solar, wind, biofuel technologies), and tourism[314].

Van der Berg (1997) has evoked that the South African social security system is mainly composed of two elements: social insurance (i.e., target population is economically active population practicing formal jobs), and social grants (i.e., social assistance, social transfers, social benefits) addressing the poor lacking a social insurance. Compared to Brazil and Mexico's specific poverty eradication programs composed of various social assistance, South African government has instead implemented more specific programs such as for improving education, health access which constitute different items of poverty in a multidimensional perspective. For instance, the Old Age Pension, Child Support Grant, Expanded Public Works Program (EPWP), The South African Schools Act of 1996, Heath Systems Trust, Operational Plan for Comprehensive HIV and AIDS Care, Management and Treatment for South Africa can be considered some of the social assistance programs implemented by the government in post-apartheid era, which are related to poverty alleviation. Taylor (2015, 150) has recommended social grants "as a policy instrument to achieve a welfare function when they provide temporary relief designed as a short-term intervention to address shocks, such as unemployment and related risks, hunger and malnutrition, and other vulnerabilities", and cash

[313] Please refer to the documents in the links below for further information (Republic of South Africa 2007a, accessed 2 February 2019): https://www.daff.gov.za/docs/GenPub/asgisa.pdf, https://www.gov.za/sites/default/files/gcis_document/201409/asgisa-2007.pdf

[314] Please refer to the information given at ILO platform in the link below (accessed 2 February 2019): https://www.ilo.org/jobspact/news/WCMS_151955/lang--en/index.htm, https://www.gov.za/about-government/government-programmes/new-growth-path

grants to "protect people living in extreme poverty or those defined as living in income-poor households, by ensuring a minimum level of well-being for people who live in extreme destitution due to income poverty" in the core of social policies in line with the social protection principles evoked in Taylor Report 2002. In present time, South Africa government's social benefits cover old-age pension, child support grant, social relief of distress, care dependency grant, grant in aid, war veterans grant, foster child grant, disability grant, admission of older persons to residential facilities, and international social services[315].

Ralston et al. (2016) have emphasized that elderly people living in rural areas are notably vulnerable regarding income, health status (notably in terms of HIV), and employment. Regarding the study done by Stewart and Yermo (2009), they have found that for nearly three-quarters of the elderly population aged 60 yrs. or more were dependent on old-age grant as their main income source provided by the government for pursuing their life. In this context, social allowances for older people becomes one of the important issues in combatting poverty. For instance, Schatz et al. (2011) have mentioned that among female-headed households affected by HIV, access to old-age grant is a substantial support for overcoming financial and psychological difficulties. The actual *Old Age Pension* implemented by the government ensures a cash transfer of R1 700 maximum per month except for people aged 75 yrs. or older it is R1 720. And the target group consists of at least 60 yrs. old South African citizens or permanent resident living in South Africa with no receipt of any other social grant for him or herself, no cover by the national social insurance, no earnings more than R78 120 if the person is single or R 156 240 if he or she is married, and with no assets worth more than R1 115 400 if the person is single or R2 230 800 if he or she is married[316]. Another social grant called *Admission of older persons to residential facilities* aims at providing 24-hour care for the elderly who wish to live in an older person's residential facility. The eligibility for this program requires being South African citizen, at least 60 yrs. old and notably in the position of destitution (i.e., extreme poverty) despite receiving an old-age grant or a pension.

Regarding *Child Support Grant* (CSG onwards)[317] which is first implemented in1998 after the recommendations of the report of the Lund Committee on Child and Family Support (1996), the actual program of government supports R410 a month per child. There are six conditions to be fulfilled for being eligible to receive the grant: (1) for the adult: having South African nationality or permanent residency living in South Africa, being child's primary caregiver, not earning more than R48 000 per year if the person is single or not more than R96 000 as the income of household for married persons; (2) for the child: be 17 yrs. old or younger, should live in South Africa, no covered by the national social insurance, and should live with her/his primary caregiver who should not be paid for her /his caregiving. This social welfare is available only for maximum six children in the household. As a result of impact assessment program carried out by Unicef on the South African CSG (see DSD, SASSA and UNICEF 2012)[318], the

[315] Please refer to the link below for further information on social benefits (accessed 2 February 2019): https://www.gov.za/services/services-residents/social-benefits

[316] Please refer to the link below for further information (accessed 27 January 2019) : https://www.gov.za/services/social-benefits-retirement-and-old-age/old-age-pension

[317] Please refer to the link below for further information (accessed 28 January 2019) : https://www.gov.za/services/child-care-social-benefits/child-support-grant

[318] Please refer to the report in the link below (accessed 27 January 2019): https://www.unicef.org/evaldatabase/files/CSG_QUANTITATIVE_STUDY_FULL_REPORT_2012.pdf

report has showed that in 2012 the program was covering 10 million South African children a month. The benefit of such program on children's life in different perspectives is shown in the report, concerning children's early life including nutrition, performance at school, development of cognitive skills, health status, child labour, adolescent's risky behaviors. Some major results are: the CSG was found having a significantly higher impact on children enrolled in the CSG at birth compared to children enrolled in the CSG at the age of six. Notably, it is found that the CSG ameliorated academic performance of the girls and children with less educated mother. Concerning health status, notably the boys are found significantly affected (at the 10% level) in a positive way by the CSG. In other words, the likelihood of becoming ill is reduced by early enrollment in the CSG for the boys (e.g., 21 per cent of likelihood of being ill for boys enrolled in the CSG at the birth versus 30 per cent of likelihood of being ill for boys enrolled in the CSG later). But also, for the adolescents, it is found that receipt of the CSG reduces absences from school. In addition, the CSG has a significant impact on reducing pregnancy, sexual activity, the use of alcohol and drug, criminal activity and being part of gang among adolescents. Thus, the CSG has a positive impact at three levels which are reducing poverty and vulnerability of the household in general sense, diminishing drop-out and absences at school, increasing good health status notably for the children enrolled at early age such as at birth, reducing the likelihood of having HIV for the adolescents.

There are also another specific program focusing on children with disability requiring a special care. The *Care dependency grant*[319] ensures taking care of children with severe disability needing a full-time special care from birth until they turn 18 yrs. old. The main person who is responsible of the child receives R1 700 a month. There are some eligibility criteria for both the adult and child which are: (1) the adult should be South African citizen or have permanent resident, he or she should be parent, or primary caregiver, or a foster parent of the child appointed by the court, and he or she should not earn more than R202 800 a year if he or she is single or he or she should not have more than R405 600 a combined household income if he or she is married[320]; (2) the child should live in South Africa, be maximum 18 yrs. old, should not receive any permanent care grant from any state institutions, should have a sever disability and need full-time special care. On the other hand, concerning adults with disability, there is also a special program called *Disability grant*[321] provided by the government. For receiving this grant, the applicant should be a South African citizen or permanent resident or refugee living in South Africa at the time application, be between 18 and 59 years old, not receive any cover from any state institutions, not earn more than R78 120 if he or she is single or not more than R156 240 if he or she is married, not have assets worth more than R1 115 400 if he or she is single or not more than R2 230 800 if he or she is married, and most importantly the person should be in a position of not being able to work for more than six months, due to her/his physical or mental disability[322]. There are two types of disability grant: permanent versus temporarily. The permanent disability grant is given if the applicant's disability will continue for more than a

[319] Please refer to the link below for further information (accessed 28 January 2019):
https://www.gov.za/services/services-residents/parenting/child-care/care-dependency-grant
[320] Income criterion does not apply for foster parents.
[321] Please refer to the link below for further information (accessed 28 January 2019):
https://www.gov.za/services/social-benefits/disability-grant
[322] In addition, there are also three criteria: having a 13-digit, bar coded identity document, pass by a medical examination by a doctor appointed by the State, and keep all previous medical records and reports.

year. The temporarily disability grant is given if the applicant's disability will continue between six and twelve months without any rupture. However, the government emphasizes that a permanent disability grant may not be for whole life but at least it will be provided for more than twelve months.

In order to diminish vulnerability of her citizens in some specific situation, the South African government provides a *Social Relief of Distress*[323] allowance for who in need of assistance regarding material and basic needs, for a short period as a temporary help (generally for three months but an extension of another three months is possible as well). The eight different situations are determined in the program for being able to benefit the assistance: (i) the period between after having demanded for a child support grant and the day of receiving the grant; (ii) when a crisis or disaster happened, until the person get back to her/his normal situation; (iii) the person could not get the grant but he or she is in desperate situation and he or she does not know how to process without the grant; (iv) because of health issues, the person is unable to work for less than six months; (v) the person cannot get the monetary help from the other parent of her/his child/children; (vi) the main responsible person for earning money in the household (i.e., breadwinner) is died; (vii) the breadwinner of the household is sent to prison for a period of less than six months; (viii) the person is affected by a disaster but her/his community, are where he or she is living did not state a disaster situation to the government. As a cover, the eligible person receives a food parcel or a voucher to buy food, or a cash transfer in some provinces.

In the context of decent work and better employment opportunities, the *Expanded Public Works Program* (EPWP onwards) is launched in 2004[324]. This program has aimed at creating job opportunities in infrastructure (i.e., government-funded infrastructure projects), non-state (i.e., work opportunities in the non-profit organization program, community work program), environment and culture (i.e., work opportunities in public environment and culture programs), and social sectors (i.e., work opportunities in public social programs), increasing labour market inclusion by creating work experience, training opportunities, and especially at decreasing poverty by providing temporary work (i.e., income opportunities in short or medium period) for unemployed persons in nationwide. The target population for this program includes poor, less skilled people, notably poor people registered in social security system, but they are not eligible for grants and in fact they wish to work but unfortunately, they cannot find a job (e.g., Public works, Republic of South Africa 2017 reported that 62 per cent of all participants were unemployed with a willing of working before participating the EPWP project), women, people living with disabilities, and youth population with special focus on youth unemployed (Mfusi and Govender 2015, 124). Most recently, the government has reported that since 2012/2013 in total 4 185 426 work opportunities were already created with 4 205 730 forthcoming opportunities by the end of 2018/19. Via this program, government targets also increasing social inclusion with stronger community participation, providing sustainable livelihoods.

[323] Please refer to the link below for further information (accessed 28 January 2019): https://www.gov.za/services/social-benefits/social-relief-distress

[324] Please refer to the links below for further information (accessed 28 January 2019): http://www.epwp.gov.za/, https://www.gov.za/about-government/government-programmes/expanded-public-works-programme, https://www.parliament.gov.za/storage/app/media/ProjectsAndEvents/2017_ncop_provincial_week_october/docs/EPWP_as_a_Catalyst_for_Development-Mr_Dladla.pdf

From the end of apartheid and the start of democratic regime, South Africans especially the most disadvantaged and discriminated black, colored population have got new hopes for their future. According to Møller (2013, 924), the 1988 numbers (during the apartheid) show that 34 percent of Black, 57 per cent of Colored, 51 per cent Indian, and 55 per cent of White population were reported optimistic for the future. However, in 2010 (post-apartheid), 50 per cent of White and 53 per cent of Colored population found the life will be getting better in the future. Among the Black population, 50 per cent of them was reported positive about their future (i.e., an increase of 16 per centage points between 1988 and 2010) and for the Indian population, an increase of 7 per centage points was found. Moreover, Møller (2013, 929) has found out that 43.7 per cent of Black, 70.9 per cent of Colored, 70.7 per cent of Asian/Indian, and 82.7 per cent of White South Africans were satified with their life in 2010. Despite progress, when we look at quality of life at different aspects, we can see that there is still much to do for eradicating extreme poverty including food security and dealing with a number of socio-economic problems such as situation of townships, inequality, mass unemployment, HIV, no accessibility to drinking water, no proper sanitary services, no electricity, no adequate dwelling, and safety issues. Seekings and Nattrass (2015, 2) have mentioned that "between 1994 and 2004 some poor people enjoyed an increase in their life opportunities and a rising standard of living…for people with few formal skills, democracy may have brought the vote, dignity, and a growing welfare state that provided income, housing, and health care. But the quality of schooling remained low, crime escalated, and formal employment opportunities shrank". And, most recently, the national planning commission in the Department of Planning, Monitoring and Evaluation has launched a National Development Plan (NDP) aiming at "ensuring the achievement of a decent standard of living for all South Africans by 2030" with the 2 main goals to achieve: (i) eliminating poverty ("a reduction in the number of people who live in households with a monthly income below R419 per person from 39 percent to zero, by 2030") and (ii) reducing inequality ("a reduction in inequality as measured by the Gini coefficient, from 0.69 to 0. 60) by taking in consideration the following 10 thematic: (1) housing, water, electricity and sanitation; (2) safe and reliable public transport; (3) quality education and skills development; (4) safety and security; (5) quality health care; (6) social protection; (7) employment; (8) recreation and leisure; (9) clean environment; (10) adequate nutrition[325].

Turkey

In case of Turkey, we don't see a particular poverty eradication programme launched nationwide for electoral campaigns or for other specific political interests. Rather, governments have been implementing development plans for specific periods covering different aspects of life such as unemployment, social services, poverty, inequality. The reports published by government, for instance by development department or strategy and budget department, inform the structure, content of development programs as well as the progress done vis à vis previously launched development goals at national level. In next paragraphs, we will briefly explore the latest social relief programs, and some of development plans applied by government across the country.

[325] Please refer to the link below for further information on National Development Plan, Vision 2030 of the government (accessed 3 January 2020):
https://www.nationalplanningcommission.org.za/National_Development_Plan

In terms of social welfare programs, according to information provided on the website of the department of strategy and budget[326], for 2019, it is planned to provide allowance to the following subjects: (i) health/medical service grant for people suffering financial problems; (ii) grant for disabled citizens for home care, and grant for disabled citizens in a broad sense as well as to their relatives who take care of them; (iii) grant for women victims of violence including shelter in women's guesthouses, psycho-social support, legal support by provincial lawyers and bar associations, nursery facilities for children between the ages of 0-6, and pocket money around half of the monthly net minimum wage; (iv) grant for foster-parents who take the responsibility of raising children in need; (v) elderly grant for people aged over 65: (vi) grant for children of families suffering financial poverty; (vii) maternity benefits; (viii) within the scope of conditional health assistance, pregnancy grants on the condition that pregnant women give birth at the hospital and go to the doctor regularly; (ix) within the scope of conditional health care, child grants are given to the families on monthly basis on the condition that families regularly take their 0-6 year old children to health checks.

Between 1963 and 2018, 10 development programs were already applied, which are generally designed for a period of 5 years. The most recent one is 11th development program for 2016-2023[327]. Under the umbrella of social services, social relief and fight against poverty, the current plan mentions that the aim is increasing access to opportunities, ameliorating social and economic participation, raising life standards, decreasing income inequality, reducing poverty for those are at risk of poverty and social exclusion. Some of the policies and measures defined in the program are: increasing employment opportunities, and specific formations for those are suffering poverty; increasing employment opportunities for those are receiving social relief and capable of being active on the labour market; reinforcing the link between social service and social relief programs; increasing access to social relief in rural areas; developing standards of caring services for elderly and disabled citizens; increasing social service programs to socially disqualified, fragile population in terms of both quality and quantity; increasing daily care services for those are in need; developing the access to social services for mobile agricultural workers.

Previously, in 9th development program applied for 2007-2013[328], under the topic of strengthening human development and social solidarity, fighting against poverty was addressed. In the same way of Eurostat defines poverty, this program refers to people who are at risk of poverty and social exclusion and aims at improving people at this risk through sustainable policies regarding labour, education, and health. The report mentions also the important role of social security and redistribution of income to those in need. In particular, it is mentioned that "services to reduce poverty will be in a manner of preventing the emergence of a poverty culture and encouraging the poor to become producers" (Turkey State Planning Organization 2006, 103). Moreover, in sustainable development aspect, the program aimed at (i) diminishing informal employment; (ii) focusing on rural and underdeveloped zones in order to increase entrepreneurship and productivity, empowerment of women, notably via vocational

[326] Please refer to the link below for further information (accessed 13 February 2020) : http://www.sbb.gov.tr/sosyal-yardimlar/

[327] The report is available in the link below (accessed 13 February 2020): http://www.sbb.gov.tr/wp-content/uploads/2019/07/On-Birinci-Kalkinma-Plani.pdf

[328] The report is available in the link below (accessed 14 February 2020): http://www.sbb.gov.tr/wp-content/uploads/2018/11/Ninth_Development_Plan_2007-2013.pdf

educational programmes/formations; (iii) increasing social awareness vis à vis gender violence, i.e., violence against women, and fighting against it across the country; (iv) increasing education opportunities in particular for girls living in rural areas, disabled people, children from low income families, and diminishing and preventing child labour; (v) increasing home care services for elderly population; (vi) creating new employment opportunities for disabled people; (vii) reducing unemployment risk for the unskilled and disadvantaged people immigrated from rural to urban areas; and (viii) making better educational, employment services as well as efforts for ameliorating dwelling conditions for those are under the risk f poverty and social exclusion. The report elaborates also some goals for increasing the effectiveness of the social security system. This lats point includes: (i) extension of target groups across the country to reach the entire population; (ii) increasing the quality of services; (iii) diminishing informal employment for reaching the entire economically active population, given that in case of informal employment, employees cannot benefit from social security services.

In 8th development plan applied for 2001-2005[329], under the topic of enhancement of welfare, alleviation of poverty was mentioned. In that vein, the plan was aimed at increasing economic growth, and based on that eliminating absolute poverty and reducing relative poverty. But also, the topics of income redistribution in favor of poor population, focus on rural areas in terms of productivity and in particular enrichment of agricultural products, small and medium sized enterprises, entrepreneurship, discrepancies among regions in terms of access to and availability of educational, health, social services, housing conditions of the poor as well as their situation vis é vis access to educational, health, and social services, and food security for the poor notably vulnerability of the poor in case of natural disasters and economic crisis were determined in the plan to combat poverty.

It is important to note that development plans include a large variety of aims and specific programs that are not necessarily mentioned above. We can say that elimination of absolute poverty was at priority in early development programs. And, in recent plans, the focus was shifted towards relative poverty and social exclusion in the same spirit of European Union look at poverty. Poverty has been measured in a monetary aspect by the national statistical office; however, multidimensional aspects of poverty is implicitly evoked in development plans in order to increase the quality of life of citizens in the country. Through years, informal employment, increasing access to social security, ameliorating access to education, health, and social services as well as living conditions in rural areas were regularly determined in the reports. Most recently, gender violence, social exclusion, and amelioration of care services for elderly, disabled population were emphasized.

Mexico

As Brazil, Mexico has encountered a substantial level of poverty and inequality, notably due to economic crisis and fluctuations happened since the 1980s (e.g., debt crisis in the 1980s, structural adjustment programs involving a restructuration of the economy with privatization, fiscal reform, austerity policies, the Peso crisis of 1994-1995). Between 1984 and 1992, the share of ninety percent of the population in national income was sharply decreased in favor of

[329] The report is available in the link below (accessed 14 February 2020): http://www.sbb.gov.tr/wp-content/uploads/2018/11/Eight-Five-Year-Development-Plan-2001-2005.pdf

the rest most prosperous ten percent of the population (see Trejo and Jones 1998, Diaz-Cayeros et al. 2016). Regarding Consejo Nacional de Evaluacion de la Política de Desarrollo Social's computation (CONEVAL[330] onwards, cited in Diaz-Cayeros et al. 2016, 46-47): (i) in 1992, 21.4 per cent of Mexican population was living under the nutritional poverty line (i.e., based on caloric intake) given that the per centage increased to 37.4 in 1996 and decreased to 19.7 in 2012; (ii) in 1992, 53.1 per cent of Mexican population was living under the basic needs poverty line (i.e., including assets, clothing, shelter, transportation) given that the per centage increased to 69.1 in 1996 and decreased to 52.3 in 2012. As we can see over the period of twenty years, the two poverty rates remained high level with some fluctuations. Concerning poverty in basic needs, we can say that Mexico experienced four phases between 1992 and 2012: (1) between 1992 and 1994, the poverty rate was 53 per cent on average; (2) the rate sharply increased to 69 per cent in 1996; (3) from 1998 the rate started to decrease reaching the lowest level around 43 per cent in 2006; (4) from 2008 the rate started to increase again reaching 52 per cent in 2012. The twenty years period, despite some ameliorations around 2006, witnessed a non-amelioration in general for the poor. Concerning nutritional poverty, the same trend was emerged: (1) between 1992 and 1994, the poverty rate was 21 per cent on average; (2) the rate sharply increased to 37.4 per cent in 1996; (3) from 1998 the rate started to decrease from 33.3 per cent to 14 per cent in 2006; (4) from 2008 the rate took a negative direction reaching 19.7 per cent in 2012. The decrease in poverty rates is supposed due to the two social welfare programs called *Progresa* and *FISM* (Fondo de Aportaciones para la Infraestructura Social Municipal) which are introduced in 1997.

The *Progresa*[331] (Programa de Educación, Salud y Alimentación) is a conditional cash transfer program at federal level aiming at reducing extreme poverty, school drop-out, child labour and increasing access to education, health, nutrition, income, which launched in 1997 under the presidency of Ernesto Zedillo (and lasted until 2012)[332]. Moreover, this program takes in consideration vulnerable situation of women and children in the sense that following priorities are determined in the program: (i) increasing health status and providing food security for children under the age of five years old, pregnant and nursing women; (ii) monitoring educational, health and nutritional status of children simultaneously for avoiding malnourishment, starvation, dropout because of illness and lack of access to health services; (iii) motivating and ensuring adequate conditions for illiterate or dropout adults to finish their elementary or secondary studies, and keeping school attendance of children; (iv) providing cash transfers via redistribution program to families living in extreme poverty struggling with hunger and deprived in basic needs; (v) providing to poor families comprehensive information on the importance of education, health and nutritional status in order to make them more active and conscious about healthy life and human capital (see Levy 2006). The International Food Policy Research Institute considered the Progresa as one of the best public safety-net-programs implemented by a government as a poverty reduction framework, to reduce notably chronic poverty and undernutrition.

[330] The National Council for the Evaluation of the Policy of Social Development and Poverty Measurement.

[331] Please refer to the link below (accessed 23 January 2019):
http://www.ifpri.org/publication/mexico-progresa

[332] Another program called *Solidaridad* was applied between 1988 and 1997.

Later, in 2000, the government under the presidency of Vicente Fox continued to apply the *Progresa* with some modifications under a new name called the *Oportunidades* (launched in 2002) which included new benefits and a larger coverage with conditions of regular school attendance for children and of regular health care visits for family members in order to benefit cash transfers. The main beneficiaries were female household responsible, mothers. It is reported that by 2005, five million families (which corresponds to more than half of families living in extreme poverty) are covered by the *Oportunidades* (See Diyaz-Cayeros et al. 2016, 49). Notwithstanding, despite a large coverage of the program and high expectations, Diyaz-Cayeros et al. (2016, 50) mention some shortcomings of the program such as since the cash transfers are conditioned by regular school attendance of children in family and regular clinic visit for monitoring heath status, poor people living in locations without school and health services are not able to benefit the program.

As we can see the two conditional cash transfer programs, the *Progresa*, and the *Oportunidades*, take in consideration multiple aspects of poverty through various indicators. Besides, before the implementation of these 2 programs, the Marginality Index (*Índice de Marginación*)[333] initiated by the Mexican Population Council (CONAPO-Consejo Nacional de Población)[334] in 1990 (with the first report on regional inequality and marginality at municipal level in Mexico[335] including 4 dimensions, education, dwelling conditions, distribution of population, and income, but also the marginality of the indigenous population in the country) until 2010 has employed multiple indicators to measure marginality in terms of various geographical units such as municipality, federal units, urban vs. rural settings. Marginalization is defined as a multidimensional and structural phenomenon emerging due to inequal distribution and leads to exclusion of some groups form the mainstream society both from the process and the benefits of development (CONAPO 2011)[336]. For instance, the 1990 index includes the per centage of illiterate people aged 15 or over, the per centage of people aged 15 or over without a primary school graduation, the per centage of people living in households lacking a toilet and a proper sewerage system, the per centage of people living in households/dwellings lacking electricity, the per centage of people living in households/dwellings lacking a piped water system, the per centage of households living in an overcrowded house/dwelling, the per centage of people living in households/dwellings having a dirt floor/not adequate floor conditions, the per centage of people living in locations with less than 5,000 habitants, and the per centage of economically active population earning an income up to 2 minimum wages.

In January 2013, under the presidency of Enrique Pena Nieto, a new social welfare program called *Cruzade Nacional Contre el Hambre*[337] (CNCH onwards) composed of 70 sub-programs at federal level is launched. The main goals of this program were reducing poverty, hunger,

[333] The dictionary of the variables used in the construction of the index of marginalization between 1990 and 2010 can be reached out via the link below (accessed 3 January 2020):
https://www.gob.mx/cms/uploads/attachment/file/459300/DICCIONARIO_DE_DATOS.pdf
[334] Please refer to the link below for more information on CONAPO (accessed 3 January 2020):
https://www.gob.mx/conapo#2943
[335] The official document in Spanish is available in the link below (accessed 3 January 2020):
http://www.conapo.gob.mx/work/models/CONAPO/indices_margina/margina90/IM1990_docprincipal.pdf
[336] Please refer to the link below for further information (accessed 3 January 2020):
http://www.conapo.gob.mx/work/models/CONAPO/Resource/1755/1/images/01Capitulo.pdf
[337] The National Crusade Against Hunger.

undernutrition, and social marginalization. In line with the MDG1, a more specific program focusing on eradication of hunger, called *Sistema Nacional contra el Hambre* (SINHAMBRE) is created by the government for increasing food security and ameliorating food production in the country. Five goals are defined in the framework of CNCH, which are: (1) eliminating hunger by providing the food to the most disadvantaged groups and the poor with limited resources; (2) eliminating malnutrition among children, and preventing malnutrition in general; (3) increasing national food production with a priority in areas of extreme poverty and of high deprivation of food, and especially increasing earnings of small peasants and farmers[338]; (4) reducing post-harvest losses; (5) promoting community participation[339]. More specifically, first goal includes the ten points following: ameliorating physical and economic conditions to reach health and nutritional foods, increasing food supply in areas of higher concentration of PEA, ameliorating health status, reinforcing social security, decreasing school dropout and educational lag, ameliorating dwelling conditions, ameliorating access to basic housing services (e.g., electricity, water), ameliorating construction of infrastructure, creating new jobs and new income sources, improving access to funding producers and organizations in social sector.

In 2014, another social inclusion program called *Prospera* replaced the Oportunidades, under the presidency of Enrique Pena Nieto. The *Prospera*[340] remains a social inclusion program for the poorest and disadvantaged population in the country, in same line with his precedents, focusing on reducing extreme poverty and improving access to nutrition "through direct money benefits to the beneficiary families to contribute to the improvement of the volume, quality and diversity of the food they eat", to education "by means of a larger educational coverage, granting scholarships to encourage school retention and progression", and to health "by means of actions to promote health and disease prevention, as well as drive access to quality healthcare services" (SEDESOL 2014, cited in Lárraga 2016, 7-8[341]). Target groups are defined through two stages: first, localities are selected (i.e., geographic targeting)[342], second, beneficiary households are identified in the selected localities. In second stage, the most eligible person in the household for answering questions regarding household situation is invited for an interview which is generally held at home. This interview serves for identifying whether the household is eligible or not for the coverage, and it is based on the Survey of Socio-Economic and Demographic Characteristics of Households[343]. After the interview based on the responses, authorities compute total current monthly income per capita for the household by employing

[338] In Mexico, geographical segregation plays important role in poverty analysis because extremely poor people are generally living in rural areas, far from convenient transports and roads, and they are deprived in basic needs such as health, education services, transports, drinking water, and food (See Diaz-Cayeros et al. 2016, 51).
[339] Please refer to the link below (accessed 24 January 2019):
https://www.mppn.org/wp-content/uploads/2016/11/19.-MEXICO_ESP_WEDS.pdf
[340] With this program, Mexican government initiated to improve transparency and building an integrated information system to reach more easily who need the social assistance and they need what exactly. Such system developed by the Social Development Secretariat (SEDESOL) aims at better identifying the target group for an optimal coverage and resolving duplications.
[341] Or, please refer to the link below (accessed 24 January 2019):
http://dof.gob.mx/nota_detalle_popup.php?codigo=5421756
[342] For more details on it, see Lárraga (2016).
[343] Encuesta de Características Socioeconómicas y Demográficas de los Hogares (ENCASEH).

proxy means test[344]. The two models are defined regarding settling (urban vs. rural). In *rural* model, the following variables are used (Lárraga 2016, 15): (1) the ratio of the number of household members ages 0–15 and 65+ to the number of household members ages 16–64; (2) the logarithm of the total number of household members; (3) the average level of schooling of the head of the household and spouse, secondary incomplete; (4) the average level of schooling of the head of the household and their spouse, secondary complete or more; (5) the number of household members, dependent workers; (6) the number of household members, self-employed; (7) the combined food insecurity indicator[345]; (8) the health coverage access (i.e., at least one household member has access to health coverage through his or her employer); (9) the health coverage access (i.e., head of household is self-employed and at least one household member has access to health coverage through his or her employer); (10) whether the household receives remittances; (11) household rents; (12) the total number of rooms in the dwelling; (13) whether dwelling includes a bathroom for the exclusive use of the household; (14) whether household has access to water; (15) whether in dwelling there is a concrete floor; (16) whether household has access to fuels such as firewood, charcoal or oil for cooking; (17) whether household has a refrigerator; (18) whether household has a vehicle; (19) whether household has a VCR or DVD player; (20) whether household has an electric oven or microwave; (21) the social Gap Index at the municipal level. In *urban* model, the following variables are used as additional to rural model (except the variable of whether the household has a concrete floor through the dwelling); (21) the number of women ages 15–49 in the household; (22) the number of household members with Seguro Popular, (23) whether the household owns the dwelling it occupies; (24) whether the household own a computer; (25) whether the household has a fixed telephone line; (26) whether the household lives in a town with a population of 100,000 or more; (27) whether the household lives in a town with a population greater than 15,000 but less than 100,000. Thus, the process of estimation of household income is based on a multidimensional view on poverty including material, dwelling living standards, social such as education, health, social participation aspects. Once the household income level is estimated, it is compared with the minimum economic well-being threshold (i.e., Línea de Bienestar Mínimo) defined by the CONEVAL. If household's estimated income level is inferior than the reference threshold, then the household is considered poor and eligible for the coverage[346].

[344] When the reported income is considered unreliable or income data is unavailable, proxy means test is used to estimate household income level based on observable characteristics of the household through socio-economic surveys for instance. Please refer to the link below as the reference (accessed 24 January 2019):
https://olc.worldbank.org/sites/default/files/1.pdf

[345] In *urban* model, the food insecurity indicator is composed of the following two indicators: (1) due to a lack of resources, whether an adult in the household skipped breakfast, lunch or dinner at least once (during the last three months), (2) due to a lack of resources, whether an adult in the household ate just one meal or did not eat for a whole day (during the last three months). However, in *rural* model, the use one of these two indicators is valid.

[346] However, households having an income level superior than the threshold may go for another verification (called *Línea de Verificaciones Permanentes de Condiciones Socioeconómicas*) if the household is composed of at least one member of the household under 12 years of age, or one who can remain a scholarship holder of the program, or a woman under 49 years of age, for being able to benefit the program (see SEDESOL 2014, cited in Lárraga 2016, 8; or, please refer to the link below (accessed 24 January 2019)):
http://dof.gob.mx/nota_detalle_popup.php?codigo=5421756).

According to statistics, at the end of 2014, 6 million of households are covered by the *Prospera* (see Lárraga 2016[347]), which reached 6,9 million households in 2018 including higher education scholarship attributed to 1,093, 659 adults for completing their primary and secondary education[348]. In addition, the program has stimulated access to vocational training, formal employment through the National Employment Service, financial services[349] and proposes a generalization of popular insurance among the poor households and *Liconsa* milk supply program[350] providing powdered milk to poor families. Francesca Lamanna, a social protection specialist at the World Bank (November 19, 2014), mentions that "delivering cash directly to families encourages beneficiaries to send their children to school and to the health center; it alleviates extreme income restrictions to cover food costs; and allows them to make important investments to break the cycle of poverty"[351]. As we can see, the *Prospera* has primarily focused on three social rights which are education, health, and nutrition. For promoting social inclusion of the poor and the most disadvantaged groups including those are marginalized or living in remote areas, social protection framework extended the coverage of these people for facilitating access to employment and financial services (i.e., labour market insertion, "by means of institutional coordination agreements seeks to improve the access of members of the homes served by the program to training or employment activities or programs which will allow their favorable insertion in the formal labor market", financial inclusion, "giving priority access to the beneficiaries of the program to financial services such as financial education, savings, life insurance, and credit" (SEDESOL 2014, cited in Lárraga 2016, 7-8[352])) in order to include them in mainstream activities of society. Most of families have become able to send their children to school instead of letting them work, such an improvement has created a substantial hope for increasing human capital, better childhood, and decreasing inter-generational poverty in the country. Notably, direct cash transfers for getting food and having access to health services helped families to spend their money for the schooling of their children, savings, obtaining microcredits, increasing production.

As we have seen Mexico has known several poverty reductions, cash transfer programs; however, the recent efforts of the CONEVAL to measure multidimensional poverty are quite substantial. As a first country in the world, Mexico, have adopted in 2004 the Article 36 under the General Social Development Law-LGDS (*Ley General de Desarrollo Social*) which mentions specifically the legal basis for considering and measuring poverty from a multidimensional perspective based on the 7 indicators across the country, and has led to the creation of the CONEVAL as a main responsible of the Mexican Multidimensional Poverty

[347] Please refer to the link below (accessed 23 January 2019):
http://www.worldbank.org/en/news/feature/2014/11/19/un-modelo-de-mexico-para-el-mundo
[348] Please refer to the link below for 2018 update (accessed 23 January 2019):
https://www.gob.mx/prospera/articulos/en-prospera-celebramos-nuestro-4-aniversario?idiom=es
[349] Please refer to the link below (accessed 23 January 2019):
http://www.worldbank.org/en/news/feature/2014/11/19/un-modelo-de-mexico-para-el-mundo
[350] Please refer to the link below for more information (accessed 23 January 2019):
https://www.thecasecentre.org/main/products/view?id=65459
[351] Please refer to the link below for further information (accessed 14 May 2020):
https://www.worldbank.org/en/news/feature/2014/11/19/un-modelo-de-mexico-para-el-mundo
[352] Or, please refere to the link below (accessed 24 January 2019):
http://dof.gob.mx/nota_detalle_popup.php?codigo=5421756

Measure[353] (current income per capita, average gap between compulsory and actual education at the household level, access to health services, access to social security, decent housing in terms of space and quality, access to basic elements in decent housing such as access to piped drinking water, adequate sewerage system, access to electricity, access to nutritious and quality food, degree of social cohesion-see also Gordon and Nandy 2016, 28-29). Since its birth in 2004, the CONEVAL has published the estimates of poverty in each two years across the country both at national and federal levels, and in each five years at municipal level, based on the data provided by the National Institute of Statistics and Geography (*Instituto Nacional de Estadística y Geografía*). In that vein, the Mexican law orders a multidimensional measure of poverty and according to CONEVAL computations a person deprived of at least one of the social rights and possess an income level inferior than the wellbeing line (i.e., income line) is considered *multidimensionally* poor. Moreover, a person deprived of 3 or over social rights and lacks a sufficient amount of income is considered *extremely* poor (CONEVAL 2010, 28).

Brazil

Over the last two decades, Brazil has performed an important willing of fighting against poverty through different poor relief policies under the umbrella of various poverty eradication programs implemented by government or international organizations, such as Bolsa Escola (School Bag), Fome Zero (Zero Hunger), Bolsa Familia (Family Allowance), Brasil Sen Miséria (Brazil without extreme poverty), UN's Millennium Development Goals, UN's Sustainable Development Goals. Brazil has known for being one of the most unequal societies in the world and has witnessed significant disparities in terms of access to health and education services, notably for the disadvantaged groups with low earnings, indigenous and afro-descendants (Quilombola) populations. Regarding World Bank's calculations, the rate of extreme poverty in terms of $1.25-dollar PPP (purchase power parity) per day has known some fluctuations between 1981 and 1993. In that period, the rate was moved between 14 per cent and 18 per cent with a drastic reduction at 9 per cent in 1986[354]. However, in 1995, the rate came again at the level in 1986. Later, between 2004 and 2009 the rate began to follow a smoothly decreasing curve. As an important indicator of economic inequality, the share of the poorest 10 per cent population in national income holds an extremely low level in Brazil. In the 1980s, the share was between 0.8 and 0.9 per cent and in 1987 it decreased to 0.7 percent. Between 1987 and 2003 over around sixteen years, the share was fluctuated between 0.7 and 0.5 per cent. In 2003, it began to increase reaching 0.8 per cent in 2009. However, over thirty years, between 1980 and 2008, the extremely poor Brazilians did not earn much. Their share in

[353] Please refer to the link below (accessed 3 January 2020, see pages 8-9):
http://www.diputados.gob.mx/LeyesBiblio/pdf/264_250618.pdf

[354] The economic program called *Plano Cruzado* launched by the government under the presidency of José Sarney on February 28, 1986, is aimed at stabilizing the prices of food, fuel, cleaning products and other services in order to reduce the rate of inflation, which was significantly declined in the same year (approximately from 12 per cent to 1.5 per cent). This program had an effect on reducing extreme poverty and increasing access to food for a short period. In 1987, the rate of extreme poverty was already jumped to 14 per cent, and in 1992 it reached 18 per cent. In the 1990s, high inflation rate and economic regression came back, which led to reduction of food assistance programs. From 1993, with the implementation of the economic program called *Plano Real*, poverty situation has been alleviated. The Plano Real is implemented by the government under the presidency of Itamar Franco in 1994 for ameliorating economic situation and is applied also under the presidency of Fernando Henrique Cardoso. This program increased access to food and the rate of extreme poverty rate remained around 11-12 percent between 1994 and 2002.

national income varied between 0.5 and 0.9 percent which remains extremely low despite few ameliorations.

As many studies have already shown income inequality in late 1990s is mainly due to regressive pension system, unequal access to education, lack of opportunities for the poor and most disadvantaged population to reach higher education, high skills training programs. The World Bank has advised that positive changes in education system is a good tool in long-term but meanwhile the country needs additional social welfare programs and a new pension system (i.e., a more equitable pension system) for reducing poverty and inequality in Brazil. In 2001, under the presidency of Fernando Henrique Cardoso, the education related social welfare program called *Bolsa Escola* is applicated at federal level within the *Social Protection Network*[355]. Earlier, in 1995, the governor Cristovam Buarque was first created the Bolsa Escola in Brasilia (federal district). The aim of this conditional cash transfer program was to reduce drop out and child labour in order to ensure schooling of children and improve human capital in the country from generation to generation. The poor families[356] are supposed to receive additional resource (i.e., monthly cash as a scholarship, which is R$15 per child attending the school, up to 3 children in a family which means a maximum R$45 per household) from Bolsa Escola through magnetic cards (e.g., beneficiaries should present their Citizen Card called *Cadastro Unico* for obtaining access to *Social Protection Network* programs, and transfers are generally paid to the mother) which in turn help children to go regularly to school and to not to work, with a condition that the child should be aged between 6 and 15 yrs. old, be present at school (i.e. 85% as a minimum rate of school attendance, and schools should report attendance information to municipal governments) and the family should have an income level inferior than the determined threshold (i.e., inferior than R$ 90 per capita per month).

The *Fome Zero* is launched in 2003 under the presidency of Luiz Inacio Lula de Silva (President Lula onwards), later monitored by the Ministry of Social Development and Fight against Hunger (MSD which is created in January 2004). The goal of this social welfare program was three-fold: first, maintaining food security and reducing hunger and extreme poverty in half by 2015 (i.e., the program is addressed to poor households suffering of malnourishment, by

[355] The *Social Protection Network* encapsulates twelve social programs designed as a function of income level and household composition, which are implemented under the presidency of Fernando Henrique Cardoso in 1995, for helping the disadvantaged people in Brazil. The proposed programs are as follows: Bolsa Alimentação (Food Bag-target group was children up to 6 yrs. old, pregnant women, mothers in the period of breastfeeding up to 6 months), Programa de Erradicação do Trabalho Infantil (PETI) (Child Labour Eradication Program-target group is children aged 7-17 yrs. old), Bolsa Escola (School Bag-target group is children aged 6-15 yrs. old), Auxílio Gás (Gas Assistance-target group is low income families, as a form of monetary subsidy for affording cooking gas in dwelling) , Programa Brasil Jovem (Young Brazil Program-target group is children aged 15-17 yrs. Old), Abono Salarial PIS/PASEP (Salary PIS/PASEP-target group is workers), Bolsa Qualificação (Scholarship Qualification- target group is workers with suspended contract, enrolled in specific training program by employer), Seguro Desemprego (Unemployment Insurance-target group is domestic workers, fishermen), Garantia-Safra (Garantee for drought area-target group is families living in the drought area or at the risk of drought), Aposentadoria e Pensões Rurais (Retirement and Rural Pensions-target group is rural workers), Benefício de prestação continuada (BPC/Loas) (Continuous Benefit-target group is elderly people from 65 yrs. old and disabled people), Renda Mensal Vitalícia (Monthly Life Anuity-target group is elderly people from 70 yrs. old with no any subsistence). The necessary conditions for being able to benefit these programs depend also on various income related terms. We just provided a very brief information on target group of programs. The provided allowances from these programs are in a form of income.

[356] The terms 'household' and 'family' will be used interchangeably in Brazil case.

allocating them an amount of $20 per month), second, increasing economic growth, wages and stimulating creation of new jobs, third, building new water cisterns in arid regions in order to provide clean water, ameliorating accessibility to electricity through the two specific programs called *Zero Sede* (Zero Thirst) and *Luz para Todos* (Light for All)[357]. For having a healthy diet, the program recognized that a decent income is necessary. In line with this objective, as a part of the Fome Zero, another federal government social welfare program called *Bolsa Familia* is inaugurated in 2003 by the President Lula[358], combining Bolsa Escola (i.e., a minimum income allocation as a scholarship in primary education), Bolsa Alimentação and Cartão Alimentação (i.e. two different programs providing income allocations related to food security), and Auxílio Gás (i.e., providing subsidy to poor households to buy cooking gas) which are implemented earlier[359].

The *Bolsa Familia* is a conditional[360] cash transfer (CCT) program aimed at (i) providing monetary help by cash transfers (i.e., from R$15-95, depending on household's income and composition) to meet basic needs such as decent access to public health, education, transport services for the families living in poverty or extreme poverty (e.g., especially, to families composed of unemployed members or members working at precarious, temporary jobs in order to help children, grandchildren)[361] which are calculated based on monetary perspective.; for instance, in 2003 the extreme poverty rate in terms of World Bank's $1.25 dollar PPP per day criterion was around 11 per cent; (ii) increasing human capital; and (iii) diminishing inter-generational poverty and inequality. In order to benefit the program, (i) children between 6-15 years old should attend the school at least 85 per cent, and children up to six years old must maintain 85 per cent school attendance; ii) children up to six should obtain their immunization certificate, and together with pregnant women they should regularly do health check-ups (Soares et al. 2007). The target beneficiary group was composed of: first, priority group (considered 'extremely poor') which is the families having a per capita monthly income inferior than R$50, second, additional group (considered 'moderately poor') which is the families

[357] Please refer to the link below (accessed 17 January 2019):
https://www.mds.gov.br/webarquivos/publicacao/Fome%20Zero%20Vol1.pdf

[358] A bunch of social welfare programs are created regarding the eradication of hunger and promoting healthy diet, such as National School Feeding Program - PNAE; Distribution of Food Baskets to Specific Population Groups; Promotion of Life Habits and Healthy Eating; Network of Public Equipment and Food and Nutrition Services - Redesan; Worker's Feeding Program - PAT; National System of Food and Nutritional Security - Sisan; Food and Nutrition Surveillance System - Sisvan; Promotion of Breastfeeding and Appropriate Childhood Practices in Childhood; National Vitamin A and Iron Supplementation Programs (please refer to the link below for the original name of the programs (accessed 17 January 2019):
http://bolsa-familia.info/fome-zero.html

[359] These programs are considered inefficient because each of them has got her own management style in terms of identifying who will benefit the program and of determining payment system (i.e., banking system). Such system in turn led to many duplications in coverage. This is why, the new government has put together these assistance programs into a single social welfare program called *Bolsa Familia*, which in particular aims at increasing social safety net in terms of better use of public health, education services, of better identification of target group for better coverage, of better monitoring and evaluation, and of better synergies between sub-national and federal programs.

[360] In order to benefit the program, family should be enrolled in the Single Register for Social Programs and the amount of per capita income in family should be below the poverty line.

[361] Before the Bolsa Familia, there were other social programs called *Cartão Alimentação* (Food Card), *Auxílio Gás* (Gas-Aid programs), *Bolsa Escola* (School Bag). But the aim was to bring together all these programs under a single umbrella.

having a per capita monthly income inferior than R$100[362]. The financial allowance is generally given to the female household head, through Citizens Card with beneficiary payments provided directly at the household head's bank account- such process was thought to be useful for diminishing corruption. The government was intended to reach 11.2 million families by 2006. The Bolsa Familia is also linked with the Millennium Development Goals, such as with the MDG1 (reducing extreme poverty), MDG2 (achieving universal education), MDG4 (reducing child mortality), MDG5 (ameliorating maternal health). In 2010, the Bolsa Familia achieved to cover more than 51 million people (Paes-Sousa and Jannuzzi 2016). This program was considered as "a highly praised model of effective social policy" by the former World Bank president Paul Wolfowitz[363], and an example for reducing hunger and extreme poverty in Latin America and the Caribbean region. In 2009, the rate of extreme poverty in terms of $1.25-dollar PPP per day diminished to 6 per cent which is the lowest figure that the country witnessed since the 1980s. Between 2003 and 2010, the number of beneficiary families is increased from 3.6 million to 12.5 million (Paes-Sousa and Jannuzzi 2016, 118-119). And, notably, the poorest regions north-east and north of the country experienced a significant decrease in extreme poverty thanks to the Bolsa Familia which covered an important part of that population.

The *Brasil sem Miséria* program is established under the presidency of Dilma Rousseff in June 2011 as an extension of the Bolsa Familia, to eradicate extreme poverty (i.e., first goal of the MDGs) by 2014, covering 16 million extremely poor people (Paes-Sousa and de Martino Jannuzzi 2016), with a special focus on northeast region which is the poorest region over years in Brazil[364]. The program concerns poor families who are not yet registered in Bolsa Familia, people living up to R$70 a month. In Bolsa Familia, a maximum three children were covered; however, in new Brasil sem Miséria program the number is increased to five. Through this program, the families accepted in this program are expected to benefit from access to water, health, education, social care services, to obtain new employment opportunities with access to special training services, and to increase per capital household income (i.e., the responsible of the family receives allocation for her family). The two sub-programs, one for rural areas, another for cities are created. This feature of the program is very important because it is reported high level of extreme poverty in rural areas despite majority of population lives in urban areas. For instance, based on 2010 Population census, 46 percent of people living in extreme poverty are from rural areas. So, this program determined special measures for rural areas to improve living standards. Notably, the program focuses on amelioration of farmer's production. On the other hand, concerning urban areas, the program aimed at drawing a poverty map and creating new employment opportunities for people aged 18-65 yrs. old, including new training programs, access to micro credit.

[362] Soares et al. (2007, 1) mention R$60 as the eligibility threshold for the families living in extreme poverty, and R$120 for other poor families.

[363] Please refer to the link below (accessed 18 January 2019):
http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:20702063~pagePK:64257043~piPK:437376~theSitePK:4607,00.html

[364] Based on the 2010 population census, 8.5% of population lives in extreme poverty and of these, 59% live in northeast regions. Please refer to the link below for poverty profile in northeast regions (accessed 18 January 2019):
https://www.ipc-undp.org/pub/eng/PRB50_Poverty_profile_the_rural_North_Northeast_regions_of_Brazil.pdf

4. Methodology

This chapter is composed of three sections, which are respectively: (i) introduction; (ii) data and indicators; (ii) the steps to follow in the analysis. We use Breiman et al.'s (1984) classification trees method to examine capabilities patterns (multidimensional poverty). We explain in detail this statistical method in the appendix 9.3. However, we note that the details concerning the statistical package Rpart we used for the classification trees analysis are given in chapter 5 in South Africa case study. The following first section will introduce the four main elements of a multidimensional poverty measure. The second section focuses on the content of the national data sets we use, the indicators we create for the four case studies and for the comparative analyses. And, in third section, we will explain the steps we follow for examining multidimensional poverty.

4.1 Introduction

Before going into detail on data, indicators, models we applied in this thesis to measure multidimensional poverty, we will briefly review the four main elements of a multidimensional poverty measure: indicator, dimension, aggregation, and decomposition.

Indicators, dimensions

The choice of indicators and dimensions represent the most substantial step in the entire procedure of multidimensional poverty measurement. We consider a dimension as an aspect or a component of life such as education, income, shelter. The selected indicators will be then reflecting the content of different living aspects the multidimensional measure takes into account (e.g., household monthly disposable income within the dimension of income, or the highest achieved grade in school within the dimension of educational attainment). A dimension can be represented by many indicators. It is very important to mention the role of an indicator (i.e., the indicator aims to reflect which feature of the dimension), its utility and how it is selected (based on which criteria). The repetition of the same meaning/the same living condition by more than one indicator within the dimension will be problematic. In classical unidimensional studies, only one sphere of life which is often income or expenses (i.e., monetary aspect) is included as the principal dimension. In opposite, multidimensional analysis is based on at least two aspects of life. In each dimension, one or more indicators can be employed. Maggino (2017b, 87) mentions that "indicators comes from the late Latin world 'indicator' that means 'who or what indicates'; in statistics, it represents a more recent term indicating indirect measures of economic or social phenomena not directly measurable. In this perspective, an indicator is not simple crude statistical information but represents a measure organically connected to a conceptual model aimed at describing different aspects of reality". She also adds that "developing indicators is a normative exercise per definition, since indicators are related to a conceptual definition of the phenomenon of interest and since a phenomenon can be defined in different ways. Consequently, in order to describe a phenomenon different group of indicators can be selected".

Maggino (2017b, 110-11) has explained four main methodological issues that a researcher should bear in mind and consider when she defines and selects indicators: (i) *multidimensionality*: identifying which facets of the object of study (e.g., poverty as the object

of study) will be considered ; (ii) *nature*: objective vs. subjective[365], and quantitative vs. qualitative; (iii) *levels of observation*: micro (e.g., individuals, groups) vs. macro (e.g., communities, regions, countries)[366] ; (iv) *dynamics*: external (in terms of objective indicators) vs. internal evaluations (in terms of subjective indicators) of the aspect/dimension of the object of study, trends in terms of indicators' scores such as linear, non-linear distribution, relationships between phenomena.

Alkire (2007a), in her paper, has explained different possible arguments about the selection of indicators in measuring multidimensional poverty based on the capabilities approach. She listed five different methods that researchers use generally to select indicators (2007a, 7): (i) *existing data or convention*: the choice depends on the available data or on the data sets given by the national statistical offices to researcher, or depends on international conventions/accords/standards (e.g., for international cross-country comparison); (ii) *assumptions*: the choice depends on researcher's perception on the object of study, or depends on theories, studies, reports available in the literature or depends on the numbers and empirical evidence reported by national statistical offices or by other government reports; (iii) *public consensus*: the choice depends on the public poll or on the other public based studies' results, or on the commonly shared ideas among public, or on the ideas rectified/applied by the universal human rights chart or by the other international programs such as Millennium Development Goals, Sustainable Development Goals promoted and monitored by noteworthy international/national/regional/community/ local organizations; (iv) *ongoing deliberative participatory processes*[367]: the choice depends on the views made by stakeholders such as opinion of civil society, of experts on government programmes/policies; (v) *empirical evidence regarding people's values*: the choice depends on "the basis of expert analyses of people's values from empirical data, or data on consumer preferences and behaviors, or studies of the values that are most conducive to mental health or social benefit". Alkire has emphasized that often these methods are used consecutively or by combinations, in selecting indictors and dimensions, and she has pointed out that the method of *existing data or convention* is needed almost in each analysis.

In assessing the quality of an indicator, Maggino (2017b, 108) has listed several criteria under the umbrella of 9 thematic, namely (i) "*accuracy and validity*": the indicator should be "clear, meaningful, consistent" with the theoretical framework and the goals of the research, and the indicator should be "appropriate, exhaustive, pertinent" with its functions such as monitoring, reporting, forecasting, accounting; (ii) "*precision*": the indicator should be "repeatable and robust" in measuring the concept; (iii) "*reliability*": the indicator should be "reproducible and stable" in measuring the concept; (iv) "*objectivity*": the indicator should be "transparent and ethically correct" in collecting and disseminating the data; (v) "*appropriateness*": the indicator

[365] Maggino (2017b, 90) distinguishes *objective* vs. *subjective* indicator in the following way: the former is "based upon explicit criteria, shared by external observers" whilst the latter is "based upon subjective evaluations and implicit criteria which can vary from individual to individual".

[366] In our theoretical framework, we defined the actions and activities accomplished at household, family level as the *macro* level functionings because we consider *individual* as the main unit analysis.

[367] For more information on this technical term, please see for instance the document written by Gauvin (2009) in the link below (accessed 16 August 2019):
https://www.ncchpp.ca/docs/DeliberativeDoc1_EN_pdf.pdf
For an application, please see Howard and Leavy (2013) in the link below (accessed 18 August 2019):
https://www.ids.ac.uk/files/dmfile/ParticipateSynthesis-WhatMattersMost.pdf

should be "relevant and credible" in terms of utility; (vi) "*parsimony*": the indicator should be "practicable, up-to-datable, and thrifty" in determining feasibility and cost of the collection of data; (vii) "*availability*": the indicator should be "well-timed, timely, punctual" in reporting the results and their communication, and "periodic, and regular" in observing the object of study/phenomenon over time; (viii) "*comparability*": the indicator should be "discriminant, and disagregable" in detecting differences between social groups, settings, regions, etc.; (ix) "*usability*": the indicator should be "accessible, interpretable, comprehensible, simple, and manageable".

Aggregation technique

A *typical* aggregation process in multidimensional poverty analysis consists of computing *headcount ratio* (i.e., the total number of poor people in a given population) based on the micro level information (individual or household or family). An aggregation process corresponds to obtain a synthetized output (e.g., index[368]) by employing an appropriate mathematical function and by taking into account (i) more than one indicator's values/modalities for a macro/micro level observation; or (ii) one indicator's values/modalities among more than one macro/micro level observations; or (iii) more than one indicator's values/modalities among more than one macro/micro level observations in the data[369]. It is important to keep in mind that an aggregation process can be also composed of more than one aggregation process, which means that more than one mathematical function can be employed through different steps according to the methodology. For instance, Alkire and Foster's (AF's) multidimensional poverty headcount ratio contains two aggregation process, and the AF's multidimensional poverty index corresponds to the multiplication of two aggregate outputs, namely multidimensional poverty headcount ratio and intensity of multidimensional poverty (Alkire and Foster 2007, rev. 2009b, 2011a). It is widely accepted that the index as an aggregate output is easily understandable among public and policymakers who are not necessarily specialized in statistics. However, it is of paramount importance to verify the loss of information due to aggregation process and thus to keep an important variety of information in a multidimensional aspect. The aggregation process should not lead to reductionism; as Maggino (2017a, 115) has mentioned "from the conceptual point of view, any synthesis should preserve the consistency between the single elements and their synthesis and represent a '*stylization*' and not a compression of the represented reality". An aggregate output may be also based on good quality aggregation process in terms of keeping a substantial information on the component(s); but it won't mean necessarily that the output represents a meaningful picture of the goal/object of the study. This point bring also discussions on how the output is visually illustrated and which type(s) of decomposition/sub-group analysis is done in order to reflect better the extent of the object of study.

---

[368] Maggino (2017b, 88) has mentioned that "the word Index comes from the Latin word 'index', which means 'anything that is useful to indicate'. In statistics, it represents historically a very generic term applied with multiple meanings". She also adds that "a generic index value becomes an 'indicator', when its definition and measurement occur in the ambit of a conceptual model and is connected to a defined aim".

[369] We assume that a data is composed of rows and columns, where rows represent observations and columns designate indicators/variables. An observation may be at micro level such as individual, household, or family but also an observation may be at macro level such as country, region, or community.

In case of Alkire and Foster's MPI, indicators' values/modalities (or scores) are obtained for each individual[370] (Alkire and Foster 2007, rev. 2009b, 2011a). Then the values/modalities are compared to the deprivation thresholds defined by the researcher; for instance, if a person's highest level of education is lower than the criterion fixed in the procedure as the must have minimum level of education for not being considered deprived regarding education, then the person will be considered deprived regarding that aspect[371]. After completing the computation of scores for all indicators and comparing the scores with deprivation thresholds accordingly for all members of the household, the next step is counting/summing the number of deprivations experienced in the household. If the total deprivation score of the household is higher than the poverty line[372] determined by the researcher (e.g., if the household has a total deprivation score higher than 33 percent, supposing that higher score means higher deprivation), the household is considered poor[373]. Through this process, we can see that there is a passage from individual level deprivation to household level poverty status. In other words, individual level deprivation scores are aggregated together to obtain a total household deprivation score based on the counting process (or we can also say summing deprivation scores)- given that summing is the aggregation function here. It is important to note that aggregate deprivation score can be computed at individual level as well. For instance, by counting the deprivations faced by the person, his or her total deprivation score can be computed as well. After that, if his or her total deprivation score is compared with a poverty threshold, the information on his or her poverty status can be obtained. On the other hand, at country, regional, community, or other societal level, the aggregate poverty score can be obtained by counting the persons living in poor households or simply by counting people considered poor if the study is not based at household level. Thus, as we have seen *deprivation threshold* is used for identifying whether the person is deprived or not regarding the indicator; *first aggregation function* is used for obtaining a total deprivation score for each individual or household; *poverty threshold* is used for identifying who is poor; and *second aggregation function* is used for obtaining a total number of multidimensionally poor people in society.

An aggregation function does not always correspond to summing or counting, and there is no obligation of employing a dual cut-off system as the MPI applies including both deprivation and poverty thresholds. In the literature there are several aggregation techniques in the umbrella of three main aggregation approaches, namely linear aggregation, geometrical aggregation (e.g., geometrical mean), and non-compensatory aggregation (e.g., Multi-Criteria Analysis)[374]. For instance, the human development index (HDI) composed of three dimensions, namely long and healthy life, knowledge (education), decent standard of living, includes 4 metric variables. Life expectancy in years, expected years of schooling, mean years of schooling, and gross national income per capita (2011 PPP$) are already aggregate macro level (country level)

[370] The scores of some indicators are the same for all members living in the same household such as the monthly household income level. We note that the *global* MPI based on Alkire and Foster's counting approach considers the household as the unit of analysis.

[371] In other words, if the person has a score below than the deprivation threshold of the educational attainment, then he or she will be considered deprived of sufficient educational attainment.

[372] It is important to remark that depending on the methodology, it can be below or above the poverty line for being considered poor. For instance, in typical monetary poverty analysis, if the household has an income level below the poverty line, then he or she is considered poor.

[373] For more information on the construction detail of the MPI, see the multidimensional poverty section in the theoretical framework.

[374] See Nardo et al. 2005 for different aggregation techniques, cited in Maggino 2017a.

indicators computed based on individual level information. Notably, within the dimension of knowledge, the mean of the two indicators is computed to obtain an aggregate score of knowledge. But also, within each dimension, via the following formula,

$$Dimension\ index = \frac{actual\ value - minimum\ value}{maximum\ value - minimum\ value},$$

three dimensional indices, namely life expectancy, education, GNI, are computed. At last stage, the geometric mean of these three (dimensional) indices is computed to obtain a country level human development score[375]. Notably, there is a growing body of research evoking the inconveniences of *compensatory* aggregation methods (see Maggino 2017a, Fattore 2017, Mazziotta and Pareto 2016) in the sense that the scores of some indicators are compensated by the scores of other indicators, and such method is not adequate when it comes to ordinal variables. Maggino (2017a, 126-129) has criticized aggregative approach for its *no suitability* for "reflecting the complexity of a socio-economic phenomenon" and for "capturing the complexity of the relationship between variables".

It is important to select the appropriate aggregation function depending on the scale of indicators (e.g., metric, categorical, ordinal), but of course also depending on the purpose of the study. In the MPI, for instance, most of variables are dichotomous based on yes/no questions (e.g., if the dwelling of the household has or not electricity? yes or no), or they are categorical based on more than two possible response options (e.g., what is the highest educational attainment? Primary school, elementary school, high school, university or equivalent as the response options).

Herewith we can say that an aggregation function, often, serves for three purposes in multidimensional poverty studies: (i) putting together indicators' scores within the dimension to obtain an overall deprivation score of the dimension (i.e., dimension index); (ii) putting together indicators' scores through all dimensions to obtain an overall deprivation score at individual and/or household level (i.e., household/individual total deprivation score); (iii) counting people having a total deprivation score inferior than the poverty line to obtain an overall poverty ratio (i.e., headcount ratio) in a given society. The headcount ratio can be also adjusted according to intensity or severity of people (e.g., MPI is the adjusted headcount ratio which is obtained multiplying the headcount ratio by the intensity of poverty).

In aggregation process, there is also another function called *weighting* that allows assessing a particular importance (called 'weight') to: (i) an *indicator*; or (ii) a *dimension*; or (iii) a *group of individuals* (e.g., most disadvantaged, or vulnerable people regarding an aspect, or an ethnic group, or some people living in a particular area, people living in rural areas, etc.). It should be mentioned whether weighting concerns all individuals or a group of people in the data but it should be also clarified whether a group of people will be weighted regarding which indicator(s) and dimension(s). In other words, a weight can be seen as a metric variable which may change as a function of indicator, dimension, or observations in the data. There are mainly two types of weighting: *equal* vs. *differential* weighting. For instance, in the global MPI, equal weighting (or called nested weighting procedure) is applied through and within dimensions, which means that each dimension has the same weight and within each dimension the dimensional weight is equally shared by its indicators. Maggino (2017a, 123) has emphasized that differential

[375] Please refer to the link below for further information (UNDP 2018, accessed 17 August 2019): http://hdr.undp.org/sites/default/files/hdr2018_technical_notes.pdf

weighting process should take into account "(i) theoretical reflections that endow a meaning on each indicator or consider its impact on the synthesis; and (ii) methodological concerns that helps to identify the proper techniques, consistently with the theoretical structure". However, it is important to note that there is no obligation to add weights in the calculus. Depending on the methodology of the researcher, weights can be defined via two approaches (Maggino 2017a, 123-124): *subjective* (such as expert, civil society opinions[376]), vs. *objective* (by employing "statistical methods such as correlation, Principal Component Analysis, Data Envelopment Analysis, Unobserved Components Models" or by "multi-attribute models, like Multi-Attribute Decision Making"). A weight can increase or decrease the contribution of an indicator, of a dimension, or of a group of individuals into the overall (aggregate) poverty rate, or in other words a weight can make more or less relevant the magnitude of an indicator, or of a dimensional index, or of a household/community/group total deprivation score, or of an individual total deprivation score in the computation of the aggregate poverty score (e.g., MPI, composite index)[377]. Weighting adjusts indicator, dimension, or deprivation score based on the rationale explained by the researcher. In most cases, weights are attributed to dimensions and indicators, but we remark that there is a large variety of discussions on this topic concerning whether weighting should be applied or not and how weighting should be applied based on which methodology.

Decomposition analysis

Before presenting the framework of our sub-group/decomposition analysis, we will start off by mentioning which standard breakdown analysis are generally applied in the capabilities-based multidimensional poverty studies. More specifically, we will take in consideration the Multidimensional Poverty Index (MPI) based studies because as we have seen previously this index methodology (developed by Alkire and Foster 2007, rev. 2009b, 2011a) is the most common way of measuring multidimensional poverty in the literature of capabilities (as we have already mentioned, we use capabilities for conceptualizing poverty in our study). We also note the two following points: (i) we consider the distribution of multidimensional unadjusted and adjusted headcount ratios as the basis and not as a part of sub-group analysis; and (ii) we don't focus on dynamic analysis (i.e., changes in poverty over time) as our study is based on cross-sectional data sets. Traditional sub-group analysis concern often the distribution of poverty rates by gender, region, household types, employment status, educational attainment, industry and occupational groups[378].

In a multidimensional poverty study carried out by Santos et al. (2015, 35) concerning the region of Latin America[379], poverty in terms of setting (rural vs. urban areas) is decomposed. Authors have also provided multidimensional vs. income poverty headcount ratios (e.g., "the

[376] They can be based on or influenced by prior scientific works done in the topic, by international law, by national law, by government initiatives, by government based or independent reports, by recommendations of international organizations, by internationally or regionally or nationally accepted standards.

[377] See multidimensional poverty section in the theoretical framework for more information on weighting approach used in the global Multidimensional Poverty Index.

[378] We remark that some of statistical and conceptual terms employed in following paragraphs are directly taken in their original names used in the books/articles by their authors, we preferred to not change the original names for giving the right context/information.

[379] Including Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Paraguay, Peru.

proportion of those who are income poor but MPI non-poor as inclusion discrepancy", "the proportion of those who are income non-poor but MPI poor as exclusion discrepancy", contribution of each dimension and of each indicator to the MPI. In another multidimensional poverty study realized for six Latin American countries[380] (Battiston et al. 2009), deprivation rates by dimension and rural and urban areas are yielded. In Mexican multidimensional poverty analysis, Foster (2010) has done decomposition analysis by rural and urban areas, across states with an additional contribution analysis of each dimension to the poverty indices (adjusted head count, adjusted gap, adjusted FGT). In Colombian multidimensional poverty analysis (Angulo et al. 2016), multidimensional poverty index, multidimensional poverty gap and multidimensional poverty severity by rural and urban areas are examined. In addition, authors also computed contribution of each dimension to the MPI, proportion of population suffering income and subjective poverty, and proportion of deprived population in terms of indicators (e.g., among the multidimensionally poor, how many of them are illiterate). In another classical multidimensional poverty analysis realized by Alkire et al. (2014) concerning European countries based on EU-SILC data, uncensored headcount ratios of all indicators (including all sample, and not only the poor), adjusted headcount ratio by regions (eastern, northern, southern, western Europe), proportion of deprived population across indicators, contribution of each indicator to the MPI by countries, and multidimensional poverty by gender, age groups (18-24, 25-34, 35-49, 50-64, 65+) are calculated.

In our study, for sub-group analysis, we will be inspired by the works of Gino Germani in his work entitled *Marginality* (1980), Mack and Lansley's *Poor Britain* (1985), Townsend's *Poverty in the United Kingdom* (1979), Bowley and Burnett-Hurst's *Livelihood and Poverty* (1915), and Rowntree's *Poverty: A Study of Town Life* (1901)[381]. Germani, as descriptive statistics[382], has provided distribution of population in terms of sectors of activity (secondary vs. tertiary sector), settings (rural vs. urban), types of dwellings (share of room per person), living conditions in terms of access to water, electricity, sanitary installations[383], socio-occupational strata by natives vs. migrants (domestic service, non-skilled or semi-skilled workers, skilled workers, middle and upper non manual strata, without information or without specified occupation), socio economic status[384] by developed vs. peripheral provinces and place of birth (rural, towns, cities big cities, other). The author has also focused more specifically on the situation of the poor and marginal population compared to the situation of workers population living in the urban conglomerate of Buenos Aires, regarding the following characteristics: rural vs. urban origin, gender of the head of the household, age of the head of the household, marital status of the head of the household, household size, average number of children per non single women, educational attainment of the head of the household (no attainment, incomplete primary, complete primary and secondary or university, other such as technical, professional), employment situation of the head of the household (employed, retired

[380] Including Argentina, Brazil, Chile, Mexico, El Salvador, Uruguay.

[381] For more information on these studies, see the literature review section. Notably, Townsend's, Rowntree's, Bowley and Burnett-Hurst's studies include a large amount of information on construction of budget, household income decomposition but also other sub-group analysis. Here, we mention some of them. Please see original works for full coverage.

[382] The numbers are from 1960 and 1961.

[383] He created an index of installations (1980, 77).

[384] Based on the SES index (1980, 82) that corresponds to the average of four indicators (occupation, education, housing, and other indicators of the standard of living).

and pensioned, without employment for various reasons), percentage of working population by age groups (<14, 15-17, 18+), percentage of working women, occupational status of the employed head of the household (middle and upper occupational strata, manual working for own account, specialized manual, manual non specialized, never had any occupation), branch of activity of the head of the household (primary, secondary, tertiary), percentage of household earning less than the legal minimum wage, dwelling conditions (shanty town, slum, two or more persons per room, occupy only one room), percentage of no affiliation to any association, percentage of the head of the household who do not read magazines, percentage of the head of the household who do not read daily newspapers[385].

As we have seen Mack and Lansley in their study of poverty in Britain (1984) based on the Breadline Britain survey, computed Townsend's deprivation index (1979, 250)[386] based on twelve elements[387] and found out and ranked public's perception of necessities in terms of 35 living items (Mack and Lansley 1985, 54). The authors have enriched their investigation via several sub-group analysis. They started off by presenting how per cent of people in each of the five classes, namely professional and managerial, other non-manual workers, skilled manual workers, semi-skilled and unskilled manual workers, and social security recipients, find necessity of the concerned item, given that there are 35 in total (so for each class, there are 35 statistics). For instance, among the professional and managerial, 88 per cent found that public transport is a necessity, which is classified as the 6th most important necessity perceived in this group (in addition, 96 per cent of them perceived heating as the most important necessity among the 35 items in total). They did the same type of computations for other variables such as age groups (15-24, 25-34, 35-44, 45-54, 55-64, 65+), household type (pensioners vs. non-pensioners: single parent, other families with children, household without children, single people), political views (conservative, labour, liberal/SDP). Furthermore, authors provided other tables showing the distribution of 26 items classed by the majority of people in the sample as necessities, and by three groups: (i) per centage of population not having the item; (ii) per centage of population unable to afford the item (due to lack of sufficient amount of money); (iii) per centage of population not wanting the item. For the third group, they computed the per centage of population not wanting the item by: net equivalent household income (poorest 10%, middle 10%, top 10%), multiple lack of necessities (0 to 5 or more) and household in income deciles, pensioners vs. non-pensioners among the households in bottom 40 per cent of income range. For the second group, the distribution by net equivalent household income, but also another table in terms of households on supplementary benefit are yielded. Multiple lack of necessities is also calculated for: children by their household income situation (in deciles), households on supplementary benefit, elderly living in bottom 40% of households and in terms of pensioners vs. non-pensioners. For those in poverty, the per centage of them being sick (long-term illness) or disabled is also computed. We note that there are also similar additional tables for subjective poverty (i.e., feeling poor).

In his marvelous work, Poverty in the United Kingdom, Townsend (1979) has built a deprivation index based on 60 indicators. To show the extent of multiple types of deprivation,

[385] In addition, Germani has also mentioned some statistics in terms of migration, and he used the term 'family' instead of household (see 1980, 86-90).
[386] See also Mack and Lansley (1985, 33).
[387] More specifically, the percentage of population deprived regarding each of the twelve items defined by Townsend.

he enlarged his investigation through a large variety of sub-analysis. He computed: mean scores of deprivations in terms of income intervals and household size, the relationship between deprivation index score and net household income for four types of households (one person; man and woman; man, women and 1 child; man, woman, and 2 children), deprivation index score in terms of range of income and ownership of assets for 2-person and 4-person households. Regarding the incidence of poverty, he realized regional, gender, age groups (0-14, 15-29, 30-44, 45-64, 65+), type of household[388], birthplace, colour, employment status (employed, self-employed, not employed children, or not employed housewives, or not employed retired), occupational status (professional, managerial, higher supervisory, lower supervisory, routine non-manual, skilled manual, partly skilled manual, unskilled manual) based breakdown examination. With the aim of reflecting the impact of poverty, he yielded examples (qualitative accompanied with some numbers) for 21 distinct households : (1) young family, with disabled man and woman and handicapped child; (2) retirement pensioner living alone; (3) chronically sick and disabled man living with elderly mother; (4) young fatherless family; (5) long-term unemployed man with wife and 6 children (2 low-paid); (6) single woman supporting severely incapacitated mother; (7) low-paid man with young family; (8) complex household including 3 disabled adults, a handicapped child, fatherless family, low-paid and long-term unemployed; (9) single woman supporting elderly father (immigrants); (10) motherless family (disabled adult); (11) disabled retirement pensioner, living alone; (12) long-term unemployed man and wife and 4 children (gipsy family); (13) low-paid man and wife and 3 children (also long-term unemployed); (14) large fatherless family, including handicapped child; (15) elderly disabled couple; (16) low-paid man and wife and 5 children (formerly long-term unemployed); (17) chronically sick and disabled man living with elderly mother; (18) chronically sick man and wife and 3 children; (19) elderly retirement pensioners; (20) 3 disabled adults; (21) fatherless family (also immigrant). He also focused as a separate chapter on rural and urban settings, social minorities, unemployed and sub-employed, low paid, older worker, disabled people and long-term sick, handicapped children, one-parent families, elderly people.

As one of the famous older poverty studies in Britain, Bowley and Burnett-Hust's (1915) study is quantitative based including a large variety of sub-group analysis for examining the living conditions among the working-class population. The authors have provided the number of houses in four cities (Northampton, Warrington, Stanley, Reading) based on: dwelling information in terms of rent per week, number of rooms, number of crowded houses; medium rent as percentage of income by different income groups; earnings by gender and age group such as men over 20 yrs., boy 14-18 yrs., boy under 14 yrs., women (married or widowed) over 16 yrs., girls 14-16 yrs., girls under 14 yrs.; family composition in terms of man sole earner, man and other earners, no man earning combined by the number of dependent children; weekly wage rate combined by age groups; relation of households to poverty line (two poverty lines: Rowntree and New Standard-see literature review for more information) in 3 categories (above standard, at standard, below standard). At individual level, they computed the number of persons; living in working-class households below the new standard by earning vs. dependent men, women, boys 18-20 yrs., boys 14-18 yrs., girls 14-16 yrs., schooled children, infants; occupational group (central and local government, defence, professional occupations, domestic

[388] Man aged 60+; man under 60; woman aged 60+; woman under 60; man and woman; man and woman over 60; man and woman one over 60; man and woman both under 60; man, woman with 1 child, or 2 children, or 3 children, or 4+ children; 3 adults; 3 adults and children; 4 adults; other without children; others with children.

service, commercial, transport, agriculture, metal or machines, etc., building, woodwork, skin or leather, etc., printing or bookbinding, etc., dress, boot and shoe, food, gas or water, etc., general workers, others). Furthermore, average weekly wages paid in different occupational groups for adult males, women and young boys by different age groups, weekly rent by different family income levels are shown.

As one of the classics, Rowntree in his study of poverty in city of York (1901) has provided detailed information on household income/family budget composition, which he used for creating 5 different classes (From A to G-see the absolute poverty sub-section in Appendix 9.1). To describe life in each single class, he computed the number of families, average size of families, average family earnings (including male and female heads of household, male and female supplementary earners, payments by lodgers for board and lodging) and average rent. As a separate chapter, he has focused in detail on housing and health conditions. Concerning housing, he computed the size of house in terms of rooms and average number of occupations per room for examining overcrowding, but also the number of houses with open windows to describe house structures. For showing the relation of poverty to the standard health conditions in the city, he obtained full information on closet accommodation, water supply, the sanitary conditions of the courts and yards, back-to-back houses. In searching immediate causes of poverty (under six thematic: death of chief wage earner, ill or elderly chief wage earner, chief wage earner out of work, chief of wage earner not having a regular work, largeness of family, chief wage earner performing a regular job but at low wage), he computed the number of families, the average size of family, the average family earnings, the average rent for people living in poverty. As a complement, he has provided also some macro level statistics (he calls "vital statistics") such as death rate, death rate of children under 12 months, other detailed information on child mortality, birth rate, the frequency of diseases in the city, but also information on the general physical conditions of boys, girls, and army recruits. He has integrated a supplementary chapter including macro-level statistics composed of 7 dimensions: public-houses and clubs (e.g., on licensees vs. off licenses public houses, frequency of public houses by men, women, and children), education (e.g., number of places provided among board, voluntary schools, percentage of school attendance, cost of education), Christian Churches (e.g., average attendance), trade unions (e.g., members by different categories of occupation), co-operation (e.g., number of stores, members, share capital), friendly societies (e.g., the number of members per registered vs. non-registered societies) and life insurance, and poor relief (e.g., the number of persons in workhouses, number of persons receiving outdoor relief, classification of paupers) for giving a more comprehensive view on the living conditions of the working classes in York.

4.2 Data and indicators

In this section, we will inform about the national data sets we used in the analysis and present the indicators we created for the analysis in a table format per national survey (starting from South Africa, Turkey, Mexico, and Brazil respectively). In the table of indicators there are 6 columns representing respectively: name (code_name will be appeared in statistical outputs), content of variable based on the question given in the questionnaire of survey, codes (categories/modalities of the qualitative variables), dimension (e.g., education, health), unit (to know whether variable at individual or household level given that a household level variable's

output should be the same for each single person living in the household)[389], and type of functioning (according to our theoretical framework).

The quantitative analysis will be based on following micro national data sets: for **South Africa**, *General Household Survey (GHS) 2011* with the dimensions of dwelling situation, material situation, educational attainment, social assistance, environment, health, economic activity, and nutrition; for **Turkey**, *Statistics* on *Living Conditions and Income Survey (SILC) 2011* with the dimensions of dwelling situation, material situation, educational attainment, environment, health, economic activity, and financial situation; for **Mexico**, *Socio-Economic Conditions Survey[390] (MCS) 2012* with the dimensions of dwelling situation, material situation, educational attainment, health, economic activity, and nutrition; and for **Brazil**, *National Household Sample Survey[391] (PNAD) 2011*, with the dimensions of dwelling situation, material situation, educational attainment, and economic activity.

We note that poverty dimensions, indicators, and categories of indicators depend on the structure of survey, and some areas of life may require more indicators than others due to variety and complexity. So that, we won't have an equal number of indicators within each dimension. It is important that each indicator reflects a peculiar aspect of life, and each modality/category of the qualitative indicator defines a specific living condition regarding the topic of the indicator. Each indicator will be assessed with an equal weight; in other words, we consider all indicators of the same importance-which means that each aspect and condition of life is valued equally. Similarly, Mexico's national social deprivation index, which considers each item of the index as the indicator of a social right, applies equal weight procedure by recognizing that "no right is superior to any other; therefore, all of them should 'have the same value', in other words, given a linear combination of indicators, all of them should be equally weighted" (CONEVAL 2010, 12). Dimensions are used for better reflecting the content of indicators in terms of which aspect of life the indicator addresses.

We will employ the name of dimensions and the description of indicators in interpreting the results. Furthermore, the quality report of indicators, content of models and employed R-codes (package rpart in R) will be given in case studies (e.g., see South Africa analysis, chapter 5). In statistical term, we note that throughout the analysis: (i) we use the words *indicators* and *variables* interchangeably in the empirical part. For more information, see the statistical method section 4.2. For type of scale of variables (qualitative vs. quantitative variable-binary or dichotomous or categorical or metric), see introduction to statistical terms in Appendix 9.3; (ii) the answers "I don't know", "unspecified", "other" and missing values are recoded as 99 in data (or in some cases 9 depending on the context, see the notifications given in case studies). The variable of age category (AGE) is obtained by transforming the original metric variable in terms of years to the categories. We also computed additional variables, labelled as metric (composed) and dichotomous (composed) based on the available information in the survey questionnaire. The variable educational attainment (EDU) is transformed from her original state for regrouping some categories together. In other words, except variables mentioned as *metric (composed)* or *dichotomous (composed)* or *transformed*, other variables' categories or scale are directly taken in their original format appearing the survey questionnaire. *Metric (composed)* means that a

[389] For type of scale of variables (qualitative vs. quantitative variable-binary or dichotomous or categorical or metric), see introduction to statistical terms sub-section.
[390] Módulo de Condiciones Socioeconómicas
[391] Pesquisa Nacional por Amostra de Domicilios

new metric variable is created based on the original metric information in the survey questionnaire; *dichotomous (composed)* means that a new dichotomous variable is created based on the original categorical information in the survey questionnaire; and *transformed* means that a new categorical/dichotomous variable is created by regrouping original categories existing in the survey questionnaire. Some composed variables include already transformation process. Please see an exemplary of STATA do file (available upon request) for further information (regarding creation of indicators); (iii) we don't have a discriminatory purpose regarding gender issue. We just use the gender categories provided by the National Statistical Offices in the survey questionnaires. We argue that people are not obligated to declare their gender, or they can choose to give another answer such as 'no gender'.

South Africa

For measuring multidimensional poverty in South Africa, we will employ 2011 National General Household Survey (GHS onwards). Before proceeding with the survey, we will provide the map of South Africa for informing about the geographical coverage[392].

The Census 2011 report of Statistics South Africa mentions that some provincial and municipal boundaries were changed between 2001 and 2011, except the boundaries of only two provinces (Western Cape and Free State) were remained unaffected (see Statistics South Africa 2012a, 9-13)[393].

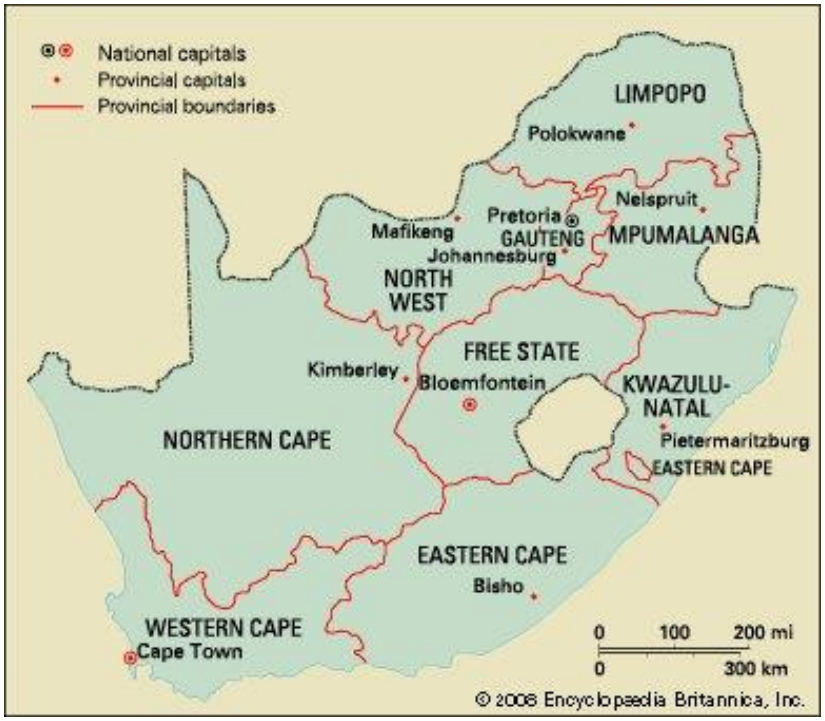


Figure 6. South Africa provinces.
Source (accessed 11 September 2019): https://www.britannica.com/place/South-Africa/Local-government

[392] The indicator, labelled PRO in the analysis (see table of indicators below), includes the 9 provinces: Western Cape, Eastern Cape, Northern Cape, Free State, KwaZulu-Natal, North West, Gauteng, Mpumalanga, Limpopo.

[393] The report (Statistics South Africa 2012a) is available in the link below (accessed 19 July 2019): https://www.statssa.gov.za/publications/P03014/P030142011.pdf

Figure 7. South Africa cities.
Source (accessed 12 September 2019):
https://www.britannica.com/place/South-Africa

National General Household Survey (GHS) 2011

The report released by the Statistics South Africa (Stats SA) mentions that the GHS 2011 interviewed all private households in nine provinces of South Africa and residents in workers' hostels (military barracks and institutions such as old-age homes, hospitals are excluded). This national survey is assessed annually by Stats SA since 2002. A household member is considered as the person "who had stayed here for at least four nights on average per week for at least four weeks"[394], and a household is defined as "a group of persons who live together and provide themselves jointly with food and/or other essentials for living, or a single person who lives alone" (see Statistics South Africa 2012c, 50)[395]. The aim of the GHS is to "determine the level of development in the country and to measure, on a regular basis, the performance of programmes and projects that were implemented to address these needs. The survey is specifically designed to measure multiple facets of the living conditions of South African households, as well as the quality-of-service delivery in a number of key service sectors" (Statistics South Africa 2012c, 1). Our analysis mainly focuses on people aged 18 years or

[394] Source : Questionnaire of GHS 2011 (cover page).
[395] Related to that, "persons who occupy the same dwelling unit but do not share food or other essentials, are regarded as separate households. For example, people who share a dwelling unit, but buy food separately, and generally provide for themselves separately, are regarded as separate households within the same dwelling unit. Conversely, a household may occupy more than one structure. If persons on a plot, stand or yard eat together, but sleep in separate structures (e.g., a room at the back of the house for single young male members of a family), all these persons should be regarded as one household. Multiple households occur when two or more households live in the same dwelling unit." (Statistics South Africa 2012c, 50). Please refer to the link below for further information on GHS 2011 (accessed 19 July 2019):
https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/225

more, except a few analyses are realized at child level (5-17 years old). In total, there are 12 dimensions which are: geographic situation, individual characteristics, household composition, dwelling situation, material situation, educational attainment, social assistance, environment, health, economic activity, nutrition, and school attendance of child. Concerning the *variables* will be used in the analysis, the table below shows their features.

Table 24. Features of selected variables from National General Household Survey 2011, South Africa

Name (code_name)	**Question/Content**	**Codes**	**Dimension**	**Unit**	**Functioning (IF, EF, MAF, MIF, MACMICF)[396]**
Household id (hid)	An identifier code for each household	String variable provided by the Statistics South Africa in the data[397]	-	Household	-
Personal id (id)	An identifier code for each individual	String variable composed of household id and serial number of the person	-	Individual	-
Serial number (serial)	A household member identifier code for each individual in her/his household	String variable provided by the Statistics South Africa in the data	-	Individual	-
Province (PRO)	The name of the province where the household lives	1=Western Cape, 2=Eastern Cape, 3=Northern Cape, 4=Free State, 5=KwaZulu-Natal, 6=North West, 7=Gauteng, 8=Mpumalanga, 9=Limpopo	Geographic situation	Household	Characteristics of province correspond to EF bringing advantage(s) or disadvantage(s) to the person. The person cannot define alone the general conditions of the province. As a result of complex social interactions, behaviors of habitants of province define the context of province.

[396] See the sub-section of capabilities for more explanations on functionings including their description and full name.

[397] The term *string* variable exists in the STATA language referring to *non-numeric* variable. See the link below for more information on that (accessed 14 September 2019):
https://www.stata.com/manuals13/u12.pdf

Setting (SET)	Geographical type-where the household lives[398]	1=urban formal, 2=urban informal, 4=tribal areas, 5=rural formal	Geographic situation	Household	Characteristics of setting correspond to EF bringing advantage(s) or disadvantage(s) to the person. The person cannot define alone the general conditions of the setting. As a result of complex social interactions, behaviors of habitants of setting define the context of setting.
Gender (GEN)	The gender of individual	1=female, 2=male	Individual characteristics	Individual	Individual characteristic (IF). Depending on the context of society, the person may get advantage(s) or suffer disadvantage(s) of her/his gender.
Population group (POP)[399]	The race of individual	1=African/Black, 2=Coloured, 3=Indian/Asian, 4=White, 5=Other	Individual characteristics	Individual	Individual characteristic (IF). Depending on the context of society, the person may get advantage(s) or suffer disadvantage(s) of her/his race.
Age (AGE)[400]	Age group in which the person is affiliated, based on the age of the person	1=18-24, 2=25-34, 3=35-44, 4=45-54, 5=55-64, 6=65+ Categorical (composed)	Individual characteristics	Individual	Individual characteristic (IF). Depending on the context of society, the person may get advantage(s) or suffer disadvantage(s) of her/his age.
Household size (HS)	The number of people living in the household	1=one person, 2=two persons, 3=three persons,	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or

[398] This variable was already existing in the data provided by the Statistics South Africa, there is no related question in the questionnaire. There is a skip from 2 to 4, and we don't have information whether a category coded 3 exists or not. We used directly the categories provided in the data.

[399] This variable is equal to the variable race in the baseline and measurement surveys of Department of Social Development. In the questionnaire of GHS, instead of the *race*, the term *population group* is used.

[400] The original question is: "What is …..'s date of birth and age in completed years?" (transformed from metric).

		4=four persons, 5=five or more Categorical (composed)			disadvantage(s) to the person.
Children (hh_child)	Household types based on the number of children (<18 yrs.) living in the household	1=no child, 2=1 child, 3=two children, 4=three children, 5=four children, 6=five children or more Categorical (composed)	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or disadvantage(s) to the person.
Children school attendance (HH_NCHSA)	Households with children *not* currently attending school/educational institution	0=all children in the household are currently attending school/educational institution, 1=at least one child in the household does not currently attend the school/educational institution Categorical (composed)	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or disadvantage(s) to the person.
Gender_Head (GH)	Gender of the head of the household[401]	1=female, 2=male	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or disadvantage(s) to the person. Except, for the head of the household it is IF.
Age_Head (AH)	Age of the head of household	1=less than 18, 2=18-24, 3=25-34, 4=35-64, 5=65+ Categorical (composed)	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or disadvantage(s) to the person. Except, for the head of the household it is IF.

[401] The father of the household is defined as "the main decision-maker, or the person who owns or rents the dwelling, or the person who is the main breadwinner" (Statistics South Africa 2012c, 51).

Population group_Head (POPH)	Race of the head of the household	same codes as POP (see above)	Household composition	Household	Household characteristic/state (MAF) which may bring advantage(s) or disadvantage(s) to the person. Except, for the head of the household it is IF.
Indigent household (INDG)	Is this household registered with the local municipality as an indigent household?[402]	1=yes, 2=no, 3=NA (e.g., live on farm)	Household composition	Household	Household characteristic/state (MAF) which intrinsically concerns all members of the household. This indicator reflects a sort of overall absolute/extreme poverty situation of the household.
Dwelling type (DWE)	Type of main dwelling	see codes[403]	Dwelling situation	Household	If the person lives in a dwelling built by him or herself, it is MIF (good quality dwelling: achievement vs. bad quality dwelling: deficit). However, bad quality dwellings should not be normally allowed by local authorities. Related to that, if the person lives in a dwelling built by other(s) only for him or her, it is MACMICF; or if it is built for other habitants as well then it is MAF (good

[402] "An indigent household is unable to fully meet payments for electricity, water, sanitation and refuse removal and gets all or some of the basic services free" (source: questionnaire of the survey).

[403] 1= Dwelling/house or brick/concrete block structure on a separate stand or yard or on farm, 2= Traditional dwelling/hut/structure made of traditional materials, 3= Flat or apartment in a block of flats, 4= Cluster house in complex, 5= Town house (semi-detached house in complex), 6= Semi-Detached house, 7= Dwelling/house/flat/room in backyard, 8= Informal dwelling/shack in backyard, 9= Informal dwelling/shack not in backyard, e.g., in an informal/squatter settlement or on farm, 10= Room/flatlet on a property or a larger dwelling/servants' quarters/granny flat, 11= Caravan/tent. A *formal* dwelling corresponds to "a structure built according to approved plans, i.e., house on a separate stand, flat or apartment, townhouse, room in backyard, rooms or flatlet elsewhere"; on the other hand, an *informal* dwelling is "a makeshift structure not erected according to approved architectural plans, for example shacks or shanties in informal settlements or in backyards" (Statistics South Africa 2012c, 51). Formal dwelling is different than both informal and traditional dwellings.

					quality dwelling: advantage vs. bad quality dwelling: disadvantaged)[404].
RDP or subsidized dwelling (DWELS)	Is the dwelling you live in an RDP[405] or state subsidized dwelling?	1=yes, 2=no	Dwelling situation	Household	RDP and state subsidized dwellings are built by the authorities/government for many households. So, it is MAF.
Roof and wall materials (ROF), (WAL)	What is the main material used for the walls and the roof of the main dwelling?	1=bricks, 2=cement block/concrete, 3=corrugated iron/zinc, 4=wood, 5=plastic, 6=cardboard,7= mixture of mud and cement, 8=wattle and daub, 9=tile, 10=mud, 11=thatching, 12=asbestos	Dwelling situation	Household	The same argument that was mentioned for DWE.
Roof and wall conditions (ROFC), (WALC)	In what condition are the roof and the walls of the main dwelling?	1=very weak, 2= weak, 3= need minor repairs, 4=good, 5=very good	Dwelling situation	Household	The same argument that was mentioned for DWE.
Room (RO)	Number of rooms per person in the household[406]	Metric (composed)	Dwelling situation	Household	If the person lives in a household composed of at least 2 persons, then situation of the dwelling such as number of available room(s) per person concerns household as a whole. Thus, it is MACMICF bringing

[404] If the person builds his or her dwelling by him or herself and if the dwelling is of bad quality, then we can say that the person has failed to build a good quality dwelling (which is a deficit). However, as we have mentioned local authorities normally should not allow bad quality dwellings, so that there is also a social responsibility, which should not be denied. People may find themselves living in bad quality dwellings because of many reasons such as economic conditions, availability of dwellings, etc, so that living in a bad quality dwelling does not always correspond to a non-achievement/personal failure.

[405] Reconstruction and development programme. Please refer to the link below for more information (accessed 12 February 2020): https://www.gov.za/faq/finance-business/where-do-i-get-copy-reconstruction-and-development-programme-rdp

[406] including open plan dining/sitting/TV rooms, lounge/dining rooms/sitting rooms/TV rooms and bedrooms.

					advantage or disadvantage to the person. If the person lives alone and s/the takes the action, then it is MIF (achievement for non-deprivation, deficit for deprivation. If the person lives alone but the action is taken by another person, then it is MACMICF.
Kitchen (KIT)	Whether the dwelling includes a kitchen?	1=yes, 2=no Categorical (composed)	Dwelling situation	Household	The same argument that was mentioned for DWE.
Sanitation (SAN)[407]	What is the sanitation type (toilet facility) the household uses?	1=flush toilet connected to a public sewerage system, 2=flush toilet connected to a septic tank, 3=chemical toilet, 4=pit latrine/toilet with ventilation pipe, 5=pit latrine/toilet without ventilation pipe, 6=bucket toilet, 7=none	Dwelling situation	Household	The same argument that was mentioned for DWE[408].
Shared toilet facility (SAN_SH)	Is the toilet facility shared with other households?	1=yes, shared, 2=no shared, 3=no toilet in the dwelling Categorical (composed)	Dwelling situation	Household	The same argument that was mentioned for SAN.

[407] "A hygienic toilet facility refers to flush toilet, chemical toilet or pit latrine with ventilation pipe" (Statistics South Africa 2011, 47).

[408] In normal conditions, only not shared flush toilet connected to a public sewerage system inside the dwelling should be allowed by local authorities. In case of not adequate sanitation facilities, this should not be seen as a personal failure but mainly as a failure of social responsibility. Moreover, if a dwelling does not include a kitchen it is also not a personal failure, but the following question should be asked: how it is allowed by local authorities? For that reason, deprivations and non-deprivations in terms of dwelling conditions items are mostly MAF except if the person has had the full responsibility to build the dwelling for herself.

Location of the toilet facility (SAN_LOC)	Is the toilet facility in the dwelling, in the yard or outside the yard?	1=in the dwelling, 2=in the yard, 3=outside the yard and less than 200m, 4=outside the yard and 201-500m, 5=outside the yard and more than 500m, 6=no toilet[409] Categorical (composed)	Dwelling situation	Household	The same argument that was mentioned for SAN.
Main source of drinking water (WAT_D)	Getting water for drinking for the household, via…	1=piped (tap) water in dwelling, 2=piped (tap) water on-site or in yard[410], 3=borehole-on-site, 4=rain-water tank on-site, 5=neighbor's tap, 6=public tap, 7= watercarrier/tanker, 8=borehole off-site/communal, 9=flowing water/stream/river, 10=stagnant water/dam/pool, 11=well, 12=spring	Dwelling situation	Household	The same argument that was mentioned for DWE. Normally, municipality should provide clean and treated drinking water via piped (tap) system into the dwelling.
Distance of water source (WATDIST)	How far is the water source from the dwelling or yard?	1=less than 200m, 2=201-500m, 3=501m-1km, 4=more than 1 km, 5=in the yard or dwelling	Dwelling situation	Household	The same argument that was mentioned for WAT_D.

[409] 200m is equal to the length of two football/soccer fields (source: questionnaire).
[410] "Piped water in dwelling or on site is piped water inside the household's own dwelling or in their yard. It excludes water from a neighbor's tap or a public tap that is not on site" (Statistics South Africa 2011, 47).

		Categorical (composed)			
Clean drinking water (SW, CW, GTW, FFBSW)	Is water you get clean regarding: 1=Safe to drink? (SW), 2=Clear (has no color/free from mud)? (CW) 3=Good in taste? (GTW), 4=Free from bad smells? (FFBSW)	1=yes, 2=no	Dwelling situation	Household	The same argument that was mentioned for WAT_D.
Drinking water treatment (WAT_T)	Do household members treat the water used for drinking? This may include boiling, adding chlorine or other chemicals, filtering.	1=yes, always, 2=yes, sometimes, 3=no, never	Dwelling situation	Household	The same argument that was mentioned for WAT_D. Normally, treated drinking water should be provided via piped water system into the dwelling by municipality. Except, if the person refuses to use municipality service then it is MFI.
Water supply interruption (WATI)[411]	Has your municipal water supply been interrupted at any time during the last 12 months?	1=interrupted, 2=not interrupted, 3=main source of drinking water is not supplied by a municipality	Dwelling situation	Household	MFA.
Water supply interruption longer than 2 day (WATI_D)	Thinking about the interruptions in your municipal water supply over the last 12 months, was any specific interruption longer than two days?	1=water supply by municipality and interrupted longer than 2 days, 2=water supply by municipality and interrupted but no longer than 2 days, 3=water supply by municipality and not interrupted, 4=main source of drinking water is not	Dwelling situation	Household	MFA.

[411] This variable is used for computing the variable WATI_D.

		supplied by municipality Categorical (composed)			
Electricity (ELC)[412]	Does this household presently have a connection to the mains electricity supply? (combined with the question of free basic electricity)	1=yes, 2=no, 3=free basic electricity Categorical (composed)	Dwelling situation	Household	Normally, electricity should be provided into the dwelling by municipality, so it is MFA. Except, if the person refuses to use municipality service or cannot afford electricity bills then it is MFI.
Electricity cutoff (ELCCUT)	Was the electricity cut off during the past 30 days for nonpayment for this household? If there was no electricity because the pre-paid card was empty it is not considered to be an electricity cut off because of non-payment.	0=no connection to the mains electricity supply, 1=mains electricity supply and pre-paid card system, 2=free basic electricity and pre-paid card, 3=electricity cutoff, 4=free basic electricity and electricity cutoff, 5=electricity cutoff for another reason, 6=mains electricity supply and no cutoff, 7=free basic electricity and no cutoff Categorical (composed)	Dwelling situation	Household	The same argument that was mentioned for ELC.
Energy (ENG_C, ENG_H, ENG_L)	What is the main source of energy/ fuel for this household for cooking (ENG_C), heating	0=no energy, 1=electricity from mains, 2=electricity from generator, 3=gas, 4=paraffin,	Dwelling situation	Household	The same argument that was mentioned for DWE. In normal conditions, local authorities should only allow the use of electricity, natural gas,

[412] This variable is used for computing the variable ELCCUT.

	(ENG_H) and lighting (ENG_L)?	5=wood, 6=coal, 7=candles, 8=animal dung, 9=solar energy			and solar energy, and society as a whole should make an effort together for better environmental conditions for everyone in the society. So that, there is a social aspect behind it. Except of the person refuses to use these energy sources, then it is MIF (personal failure).
Household material assets (TV, RAD, REF)	Does the household own any of the following in the dwelling? 1) television (TV), 2) radio (RAD), 3) refrigerator (REF)	1=yes, 2=no	Material situation	Household	MIF or MACMICF if the household is composed of at least two persons.
Household material assets-TEL	Does the household own an accessible telephone in the dwelling?[413]	1=a functional/working landline telephone but no a functional/working cellular telephone, 2=no landline but a cellular phone, 3=both landline and cellular phone, 4=none of them in the dwelling and nearest accessible phone is 500 m or less, 5=none of them and nearest accessible phone is 501 m to 1 km, 6=none of them and nearest phone is	Material situation	Household	MIF or MACMICF if the household is composed of at least two persons. However, it is important to query availability of landline in the dwelling, in the neighborhood. If public services or local/private authorities do not provide the basic communication network in the neighborhood, or in the dwelling where the person/household lives, then it is MAF in this regard. But in case of not having a telephone in the dwelling due to financial reasons, it is MIF or MACMICF.

[413] Combined of three questions: "Does this household have a functional/working landline telephone in the dwelling?", "Is there a functional/working cellular telephone available within this household", "How far is the nearest accessible telephone".

		more than 1km up to 5 km, 7=none of them and nearest phone is more than 5 km.			
Household material assets-Internet	Do members of this household use any of the following internet services INT_1=internet connection in the household, INT_2=Internet in a library/community hall/Thusong centre, INT_3=internet for students at a school/university/college, INT_4=at place of work, INT_5=internet café 2km or less from the household, INT_6=internet café more than 2km from the household, INT_7=internet access on cellphone	1=yes, 2=no	Material situation	Household	MIF or MACMICF if the household is composed of at least two persons. However, it is important to query availability of internet service/network in the dwelling, in the neighborhood. If public services or local/private authorities do not provide the basic internet network in the neighborhood, or in the dwelling where the person/household lives, then it is MAF in this regard. But in case of not having an internet in the dwelling due to financial reasons, it is MIF or MACMICF.
Education (EDU)	What is the highest level of education that… has successfully ***completed***?	0=no schooling, 1=grade R/0, 2=grade 1-6, 3=primary education diploma (grade 7), 4=lower secondary education (grade 8-9), 5=grade 10-11, 6=NTC1-NTC2, 7=NTC3, 8=NTC4-NTC6, 9=certificate or diploma with	Educational attainment	Individual	MIF (getting diploma) except if the person could not go to school due to household reasons and/or no availability of schools in the neighborhood (in terms of both quality and quantity), then it is MAF or MACMICF.

		less than grade 12, 10=grade 12/high school diploma, 11=higher or post higher diploma, 12=university degree, 13=illiterate Categorical (composed)			
Disability grant[414] (DISG)	Receiving disability grant	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	MAF. Receiving a grant is not a personal achievement or not receiving a grant is not a personal failure because the regulations for grants are determined by government/local authority for some people in the society; so, it is MAF.
Old age grant (OLDG)	Receiving old age pension	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.

[414] According to Dugard (2016, 192), the Child Support Grant (CHSG) is "provided to a parent or primary caregiver who is in need" to "support the basic needs of a child until she or he turns 18"; the Foster Child Grant is "provided to a foster parent who is legally appointed to care for a child below 18 years of age"; the Care Dependency Grant (DEPG) is "intended to assist a parent, primary caregiver or foster parent who is in need, to support a child below 18 years old who has severe disabilities and requires permanent care or support services"; the Disability Grant (DISG) is "available to adults between the ages of 18 and 59 who are in financial need and have such disabilities that leave them unfit to support themselves"; the Old Age Grant (OLDG) "applies to persons over the age of 60 in financial need"; the War Veterans Grant (WARG) is "intended for people over the age of 60 who are in needs and who served in World War II, the Korean War or the anti-apartheid struggle"; the Grant-in-Aid (AIDG) "aims to provide further support to recipients of the old age grant, disability grant and war veterans grant who require full-time care by a third party due to physical or mental disability"; and the Social Relief or distress Grant (SRG) is "a stop-gap measure that aims to provide immediate and temporary assistance to persons who are in social distress and/or unable to meet their own or their family's most basic needs. The assistance is in the form of food or travel vouchers, food parcels, clothing or money". The detailed information on the grants are available in the links below (accessed 3 January 2020): for CHSG, https://www.sassa.gov.za/Pages/Child-Support-Grant.aspx; for DEPG, https://www.gov.za/services/services-residents/parenting/child-care/care-dependency-grant; for DISG, https://www.gov.za/services/social-benefits/disability-grant; for OLDG, https://www.gov.za/services/social-benefits-retirement-and-old-age/old-age-pension; for WARG, https://www.gov.za/services/social-benefits-retirement-and-old-age/war-veterans-grant; for AIDG, https://www.gov.za/services/social-benefits/grant-aid; and for SRG, https://www.gov.za/services/social-benefits/social-relief-distress.

War veterans grant (WARG)	Receiving war veterans grant	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Foster child grant (FOSG)	Receiving foster child grant	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Care dependency grant (DEPG)	Receiving care dependency	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Child support grant (CHSG)	Receiving child support	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Grant-in-Aid (AIDG)	Receiving grant-in-aid	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Social relief or distress (SRG)	Receiving social relief or distress	1=yes, 2=no, 3=no receive social grant	Social Assistance	Individual	The same argument that was mentioned for DISG.
Waste removal (ENVP1)	Do you experience in your community/on your and neighbouring farms irregular or no waste removal?	1=yes, 2=no	Environment	Household	EF.
Rubbish (RUB)[415]	How is the refuse or rubbish of this household collected or removed?	see codes[416]	Environment	Household	Normally, rubbishes should be regularly removed by local authority/private company for the sake of environment and community. So that, it is MAF.
Environment problems (LIT,	Do you experience in your	1=yes, 2=no	Environment	Household	EF.

[415] It is important to note this variable may be included in the dimension of dwelling situation as well. If there are not enough variables in the dimension of environment, including this variable into the dimension of dwelling situation will be an optimal solution.

[416] 1= removed by local authority/private company at least once a week, 2= removed by local authority/private company less often than once a week, 3= removed by community members, contracted by the municipality, at least once a week, 4= removed by community members, contracted by the municipality, less often than once a week, 5= removed by community members at least once a week, 6= removed by community members less often than once a week, 7= communal refuse dump/communal container, 8= own refuse dump, 9= dump or leave rubbish anywhere.

WP, AP, LD, NOI)	community/on your and neighboring farms? 1 = littering (LIT), 2 = water pollution (WP), 3 = outdoor/indoor air pollution (AP), 4 = land degradation/over-utilization of natural resources (e.g., soil erosion, potholes and dongas, overgrazing, cutting of trees for firewood) (LD), 5 = excessive noise/noise pollution (NOI)				
Post (POST)	How does the household receive most of its mail/post?	1 = delivered to the dwelling, 2=delivered to a post box/private bag, 3=through a friend or neighbor, 4=through a shop, 5=through a school, 6=through a workplace, 7=through a tribal/local authority office, 8=do not receive mail	Environment	Household	Same argument that was mentioned for TEL, INTERNET.
Medical coverage (MC)	Is …… covered by a medical aid or medical benefit scheme or other private health insurance?[417]	1=yes, 2=no	Health	Individual	In normal conditions, everyone in a country should be covered by a health insurance, private or public depending on the preference of the person. So that, there is a social

[417] If the person is a dependant and covered by someone else's scheme, the answer is "Yes".

					responsibility behind it in terms of role of government/authorities but also of society as a whole. In that vein, it is MAF.
Disability (DIS)	Does… have difficulty in doing any of the following? 1= seeing (even with glasses if he/she wears them), 2=hearing (even with a hearing aid, if he/she wears one), 3=walking a kilometre or climbing a flight of steps, 4=remembering and concentrating, 5=with self-care, such as washing or dressing, 6=in communicating in his or her usual language including sign language (understanding others and being understood by others)	1=no difficulty, 2=some difficulty, 3=a lot of difficulty, 4=unable to do, 5=do not know, 6=cannot yet be determined	Health	Individual	Individual characteristic, IF.
Disability1 (DIS1)	Is.... limited in his or her daily activities, at home, at work or at school, because of a long-term physical, sensory, hearing, intellectual, or psychological condition, lasting six	1=yes, 2=no	Health	Individual	IF.

	months or more?				
Disability2 (DIS2)	Disability variable including 4 modalities, is composed based on the variable DIS (see above)	1=severe disability (having 'a lot of difficulty' or being 'unable to do' regarding 1 or more of the 6 categories) 2=UN disability (having 'some difficulty' regarding two or more of the 6 categories) 3=other disability (having 'some difficulty' regarding 1 of the 6 categories) 4=no disability Categorical (composed)[418]	Health	Individual	IF.
Severe disability (DISSEV)	The category 1 of the variable DIS2 (see above)		Health	Individual	IF.
Consulting a health worker (HWC)	What is the main reason, why …… did not consult any health worker concerning her/his illnesses/injuries during the past month?	0=consulted, 1=not consulted because it is too expensive, 2=not consulted because it is too far, 3=not consulted because it was not necessary/the problem was not serious enough, 4=not consulted because the person did self-medicated/treated her/his self,	Health	Individual	MIF but some answer options such as 2, 5 reflect MAF.

[418] The category UN disability is constructed based on the UN criteria (see supplementary analysis in case study).

		5=not consulted because of fear of stigmatization Categorical (composed)			
Chronic illness (CILM_1,…,CILM_6)	Has …..been informed by a medical practitioner or nurse that he/she suffers from any of the following chronic illnesses or conditions, and he or she takes the medication for this? 1) asthma (CILM_1), 2) diabetes (CILM_2), 3) cancer (CILM_3), 4) HIV and AIDS (CILM_4), 5) hypertension/high blood pressure (CILM_5), 6) arthritis (CILM_6)	1=yes, he or she takes medication, 2=no, he or she does not take medication Categorical (composed)	Health	Individual	IF.
Time to travel to the health	Time to travel to the health institution[419]	see codes[420]	Health	Household	MAF.

[419] Based on the questions: "How long does it take when using the usual means of transport to get to the health institution that your household normally goes to?- Specify for one direction only, using the usual means of transport", "Is this facility the nearest of its kind (clinic/hospital/health centre etc.) to your dwelling?", and "If not the nearest, why is the household normally not using the nearest facility?" (source: questionnaire of the survey).

[420] 1=less than 29 min., 2=more than 30 min. and facility is the nearest of its kind to the dwelling, 3=more than 30 min., not the nearest one because the nearest facility is not clean, 4= more than 30 min., not the nearest one because in the nearest facility there is a long waiting time, 5= more than 30 min., not the nearest one because the opening times of the nearest facility is not convenient, 6= more than 30 min., not the nearest one because it is too expensive, 7= more than 30 min., not the nearest one because in the nearest facility drugs that he or she needed, not available, 8= more than 30 min., not the nearest one because in the nearest facility staff is rude or uncaring or turns patient away, 9= more than 30 min., not the nearest one because in the nearest facility more than 30 min., not the nearest one because in the

institution (TIMEH2)		Categorical (composed)			
Employment status (ES1)	Current employment status	see codes[421] Categorical (composed)	Economic activity	Individual	MIF.
Sector (SEC)	Sector of work (the organization/business/branch where the person works)	0=not employed. 1=formal sector (registered to perform activity), 2=informal sector (not registered to perform activity)	Economic activity	Individual	In normal conditions, everyone should work in formal sector. Informal sector should be checked by authorities. In this perspective, it is MAF.
Child nutrition (CHN)	In the past 12 months, did any child (17 years or younger) in this household go hungry because there was not enough food?	1=never, 2=seldom, 3=sometimes, 4=often, 5=always, 6=no children in the dwelling	Nutrition	Household	Household characteristic/state, MAF, for the adults of the household.
Adult nutrition (ADN)	In the past 12 months, did any adult (18 years and above) in this household go hungry because there was not enough food?	1=never, 2=seldom, 3=sometimes, 4=often, 5=always, 6=no adults in the dwelling	Nutrition	Household	Household characteristic/state, MAF, except MIF if the household is composed of one person.

nearest facility an incorrect diagnosis has happened, 10=more than 30 min., not the nearest one because it is not on medical aid scheme list of facilities, 11=more than 30 min., not the nearest one because he or she prefers to use a State/Provincial health institution, 12=more than 30 min., not the nearest one because he or she prefers to use a private health institution.

[421] 1=he or she is employed (working for a wage, a commission or a salary; engaged in business activities; temporarily absent but will return to work or business; work without remuneration), 2=he or she sought a job during last 4 weeks and available or ready to start to work, 3=he or she sought a job during last 4 weeks but not available or ready to start to work, 4=he or she did not search but would …. have liked to work during the last calendar week, 5=he or she did not seek a job and would not have liked to work but if a suitable job had been offered or circumstances had allowed, he or she would..... have been able to start work or a business in the last calendar week, 6=he or she did not seek a job and he or she would not have liked to work and even if a suitable job had been offered or circumstances had allowed, and he or she would not..... have been able to start work or a business in the last calendar week, 7=he or she sought a job but missing values concerning availability.

Nutrition (NU1,…,NU4)	1= Did your household run out of money to buy food during the past year? (NU1), 2=Did you cut the size of meals during the past year because there was not enough food in the house? (NU2), 3=Did you skip any meals during the past year because there was not enough food in the house? (NU3), 4=Did you eat a smaller variety of foods during the past year than you would have liked to, because there was not enough food in the house? (NU4)	1=yes, 2=no	Nutrition	Household	Household characteristic/state, MAF, except MIF if the household is composed of one person.
Children not attending educational institution (REFNCSA)	What is the main reason why …… is currently not attending any educational institution?	see codes[422] Categorical (composed)	Child	Individual	Household characteristic/state, MAF, for the adults of the household given that some answer options such as 4, 16 reflect social situation in terms of adequate access to school (various MAFs).
Time to get the educational institution (TIMESC)	How long does it take …… to get to the educational institution	1=less than 15 minutes, 2=15 - 30 minutes, 3=31 - 60 minutes, 4=61 -	Child	Individual	Household characteristic/state, MAF, for the adults of the household. Response options

[422] 1=too old/young, 2= has completed education/satisfied with my level of education/do not want to study, 3=school/education institution is too far, 4=difficulties to get to school (transport), 5=no money for fees, 6=he or she is working at home or business/job, 7=do not have time/too busy, 8=family commitment (e.g., child minding), 9=education is useless or not interesting, 10=unable to perform at school, 11=illness, 12=pregnancy, 13=failed exams, 14=got married, 15=disability, 16=violence in school, 17=not accepted for enrolment, 18=educated at home/home schooled.

	he/she attends? (specify for one direction only, using the usual means of transport)	90 minutes, 5=more than 90 minutes			reflect indirectly societal situation in terms of available school in neighborhood.
Problems encountered at educational institution (BOOK, POORT, LACKT, BADC, FEES, CLAS, TABS, PMAN)	During the past 6 months, what problems, if any, did …… experience at the educational institution he/she attended? 1=lack of books (BOOK), 2=poor quality of teaching (POORT), 3=lack of teachers (LACKT), 4=facilities in bad condition (BADC), 5=fees too high (FEES), 6=classes too large/too many learners (CLAS), 7=teachers are often absent from school (TABS), 8=teachers were involved in strike (PMAN)	1=yes, 2=no	Child	Individual	Household characteristic/state, MAF, for the adults of the household. Response options reflect social situation at school (various MAFs).
Violence at school (VIOLS1,…,VIOLS5)	Has …. experienced any form of violence, corporal punishment or verbal abuse at school between January and June 2011? 1=corporal punishment by teacher (VIOLS1), 2=physical	1=yes, 2=no	Child	Individual	Household characteristic/state, MAF, for the adults of the household. Response options reflect social situation at school (various MAFs).

	violence by teacher (VIOLS2), 3=verbal abuse (being insulted, teased or harassed) by teacher (VIOLS3), 4=verbal abuse (being insulted, teased or harassed) by other learners (VIOLS4), 5=physical abuse (being hit or punched) by another learner (VIOLS5)				

Turkey

Statistics on Income and Living Conditions Survey (SILC) 2011

Our analysis concern only people aged 18 years or over. In total, there are 10 dimensions which are: geographic situation, individual characteristics, household composition, dwelling situation, material situation, educational attainment, environment, health, economic activity, and financial situation. The person him or herself or the head of the household responds to the questions by giving the information for other members of the household- depending on the structure of questionnaire. Before presenting the indicators will be used in the analysis, we will provide the map of Turkey for informing about the geographical coverage.

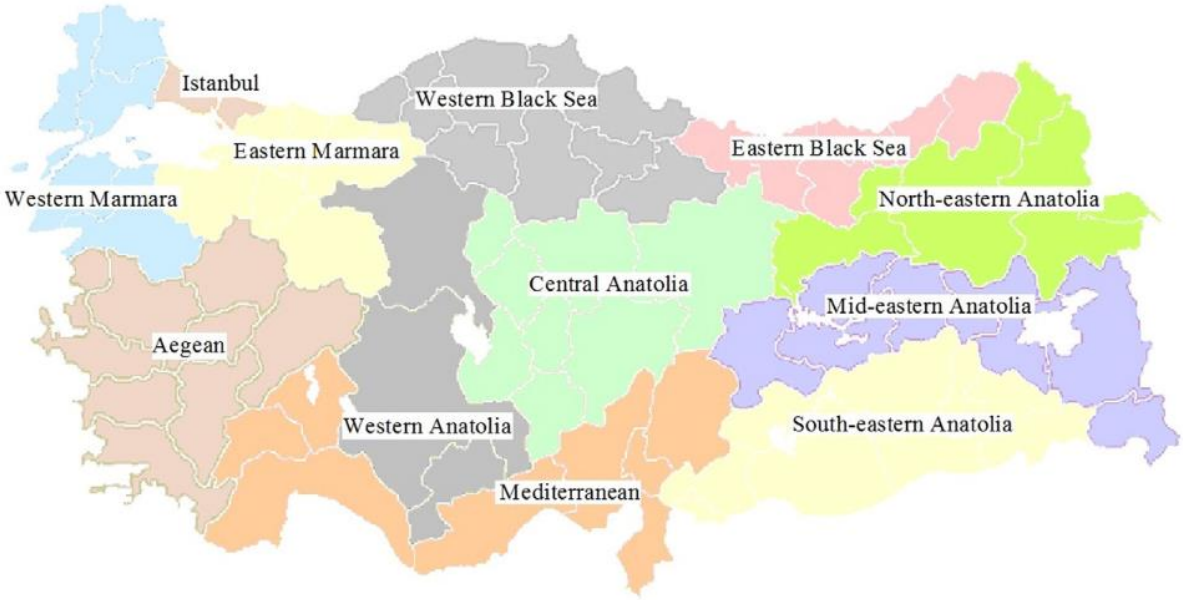


Figure 8. Regions of Turkey based on NUTS1 classification.
Source: Lynn et al. 2015. Intelligence (50) 144-149

Concerning the *variables* will be used in the analysis, the table below shows their features.

Table 25. Features of selected variables from Statistics on Income and Living Conditions Survey (SILC) 2011, Turkey

Name (code_name)	**Question/Content**	**Codes**	**Dimension**	**Unit**	**Functioning (IF, EF, MAF, MIF, MACMICF)**
Household id (hid)	An identifier code for each household	String variable provided by the Institute of Statistics Turkey in the data	-	Household	-
Personal id (id)	An identifier code for each individual	String variable composed of household id and serial number of the person	-	Individual	-
Serial number (serial)	A household member identifier code for each individual in her/his household	String variable provided by the Institute of Statistics Turkey in the data	-	Individual	-
Region (REG)	Geographical area covered in the survey, according to the NUTS1[423] classification, Statistical regions, Level 1[424]	1=(TR1) İstanbul, 2=(TR2) West Marmara, 3= (TR3) Aegean, 4= (TR4) East Marmara, 5= (TR5) West Anatolia, 6= (TR6) Mediterranean, 7= (TR7) Central Anatolia, 8= (TR8) West Black Sea, 9= (TR9) East Black Sea, 10= (TRA) Northeast Anatolia, 11= (TRB) Middle East Anatolia, 12= (TRC) Southeast Anatolia	Geographic situation	Household	EF
Setting (SET)	The category of setting (urban vs. rural) where the household lives-degree of urbanization	1=rural, 2=urban	Geographic situation	Household	EF

[423] Please refer to the link below for more information on NUTS classification implemented by Euro Stat (accessed 28 July 2019):
https://ec.europa.eu/eurostat/web/nuts/background

[424] Classification of Statistical Regions of Europe (SRE Level 1) (source: the questionnaire).

Gender (GEN)	The gender of individual	1=female, 2=male	Individual characteristic	Individual	IF
Age (AGE)	Age group in which the person is affiliated, based on the age of the person	1=18-24, 2=25-34, 3=35-44, 4=45-54, 5=55-64, 6=65+ Categorical (composed)	Individual characteristic	Individual	IF
Household size (HS)	The number of people living in the household	1=one person, 2=two persons, 3=three persons, 4=four persons, 5=five or more Categorical (composed) based on the metric information	Household composition	Household	MAF
Household type (HT)	Household type	see codes below[425]	Household composition	Household	MAF
Gender_Head (GH)	Gender of the head of the household	1=female, 2=male Dichotomous (composed)	Household composition	Household	MAF
Age_Head (AH)	Age group in which the head of the household is affiliated, based on her/his age	1=less than 18, 2=18-24, 3=25-34, 4=35-64, 5=65+	Household composition	Household	MAF
Dwelling type (DWE)	Type of main dwelling	1=detached house, 2= semi-detached or terraced house, 3= apartment (less than 10 flats), 4= apartment (10 or more flats), 5= other	Dwelling situation	Household	MAF
Roof, wall, window (ROWW)	If there any problem such as leaking	1=yes, 2=no	Dwelling situation	Household	MAF

[425] This variable is already composed by the Turkey National Statistical Office and is available in the data. The modalities are: 1=single person, 2= 2 adults, no dependent children, both adults below 65 years, 3= 2 adults, no dependent children, at least one adult 65 years or more, 4= Other households without dependent children, 5= Single person with dependent children, 6= Two adults with one dependent child, 7= Two adults with two dependent children, 8= Two adults with three or more dependent children, 9= Other households with dependent children. For single person households, breakdown by below 64 years, 65 years and more, male, female are suggested by the Turkey National Statistical Office in the dictionary of variables.

	roof, damp walls, rotten window frames?				
Room (RO)	Number of rooms per person in the household[426]	Metric composed	Dwelling situation	Household	MAF
Sanitation I (SAN1)	Whether bathroom or shower exists in the dwelling?	1=yes (not shared, for sole use of the household), 2=yes (shared), 3=no bathroom neither shower	Dwelling situation	Household	MAF
Sanitation II (SAN2)	Whether toilet exists or not in the dwelling? (indoor flushing toilet in the dwelling)	1=yes (not shared, inside dwelling, for sole use of the household), 2=yes (shared), 3=no toilet	Dwelling situation	Household	MAF
Kitchen (KIT)	Whether (independent) kitchen exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Piped water system (PWS)	Whether piped water system exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Hot water system (HWS)	Whether hot water system exists or not in the dwelling? (Central hot water, geyser, water heater, solar energy system, etc.)	1=yes, 2=no	Dwelling situation	Household	MAF
Heating system (HEATS)	What is the heating system in the dwelling?	1=stove (coal, gas, natural gas, electricity, etc.), 2= radiator (joint or central heating), 3= radiator (heating system for only a flat/combi boiler), 4=air conditioner, 5= other	Dwelling situation	Household	MAF

[426] excluding kitchen, bathroom, toilet.

Energy for heating (ENG_H)	What is the main energy source/type of fuel most used for heating in dwelling?	1=wood, 2=coal, 3= natural gas, 4= fuel-oil, 5= diesel oil-gasoil, 6= electricity, 7= dried cow dung (tezek), 8=other	Dwelling situation	Household	MAF
Insulation (INS)	Whether there is a heating problem in the dwelling due to insulation state of the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Lighting (LIGHT)	Whether there is a problem with the rooms being dark or not receiving enough light in the dwelling? (darkness of rooms or lack of daylight)	1=yes, 2=no	Dwelling situation	Household	MAF
Household assets (TV, REF, AIRC, WM, CAR, DISHW, COMPINT, MPHO, TEL)	Does the household own or have access to any of the following? 1) colour television, 2) refrigerator, 3) air conditioner, 4) washing machine, 5) car (for private use only), 6) dishwasher, 7) computer, internet connection for households with computer, 8) mobile phone, 9) landline telephone/telephone line	1=yes, 2=no, cannot afford due to financial reason/financial shortage, 3=no, other reason	Material situation	Household	MIF or MACMICF if the household is composed of at least two persons
Education (EDU)	What is the highest level of education	0=illiterate, 1=literate but no diploma (not a graduate), 2= primary school, 3=secondary,	Educational	Individual	MIF

	that… has successfully ***completed***? (highest level successfully completed with a diploma or certificate)	vocational secondary or primary education school, 4= high school, 5=vocational or technical high school, 6=higher education (university or equivalent) Categorical (transformed)	attainment		
Noises (NOI)	Whether there is noise in the dwelling due to neighbors or street?	1=yes, 2=no	Environment	Household	MAF
Pollution (POL)	Whether there is an air pollution, environmental pollution or other environmental problems caused by traffic or industry, where the person lives?	1=yes, 2=no	Environment	Household	MAF
Crime (CRI)	Whether crime, violence, vandalism is encountered intensively where the person lives?	1=yes, 2=no	Environment	Household	MAF
Overall/general health status (HEALTH)	Health status in general	1=very good, 2=good, 3=fair, 4=bad, 5=very bad	Health	Individual	MIF
Chronic illness (CIL)	Suffering from any chronic illness/long standing illness or condition such as diabetes, hypertension, asthma, kidney	1=yes, 2=no	Health	Individual	MIF

	failure, rheumatic diseases, etc.?				
Unmet need for medical examination (MED)	Unmet need for medical examination or treatment during the last 12 months, and main reason for unmet need for medical examination or treatment	1=no unmet need, 2= could not afford to, 3=waiting list, 4=could not take time (because of work, care for children or for others), 5=too far to travel/no means of transportation, 6=fear of doctor/hospitals/examination/ treatment, 7=wanted to wait and see if problem got better on its own, 8= didn't know any good doctor or specialist	Health	Individual	MIF, MAF
Unmet need for dentist examination (DEN)	Unmet need for dentist examination or treatment during the last 12 months, and main reason for unmet need for medical examination or treatment	1=no unmet need, 2= could not afford to, 3=waiting list, 4=could not take time (because of work, care for children or for others), 5=too far to travel/no means of transportation, 6=fear of doctor/hospitals/examination/ treatment, 7=wanted to wait and see if problem got better on its own, 8= didn't know any good doctor or specialist	Health	Individual	MIF, MAF
Economic activity status (ES1)	Current economic activity status (objective approach)	1=employed, 2=unemployed, 3=no seek a job in last 4 weeks but available for work within 2 weeks in case of a job opportunity, 4=seek a job in last 4 weeks but not available for work within 2 weeks in case of a job opportunity, 5=seek a job in last 4 weeks and available for work within 2 weeks in case of a job opportunity but no information on the method for seeking a job, 6=not economically active Categorical (Composed)	Economic activity	Individual	MIF
Economic activity status (ES2)	Self-defined current economic	see codes[427]	Economic activity	Individual	MIF

[427] 1=wage-worker or casual employee (full-time), 2=wage-worker or casual employee (part-time), 3=employer/self-employed/own account worker (full-time) including unpaid family worker, 4=employer/self-employed/own account worker (part time) including unpaid family worker, 5=seeking job, 6= Pupil, student or unpaid work experience, 7=retired, early retired or has given up business,

	activity status reported by the person (subjective approach)				
Employment status in main job (ES3)[428]	Individual's employment status in the main job	1=paid, salaried (regular wageworker/employee), 2= casual employee, 3= employer, 4= self-employed/own account worker, 5=unpaid family worker, 6=economically not active regarding ES1 definition, 7=economically not active regarding ES2 definition except people classified regarding ES1 criteria	Economic activity	Individual	MIF
Occupational group (OG)	Occupational group/profession code of the individual in her/his actual job according to ISCO 88 classification system[429]	see codes[430]	Economic activity	Individual	MIF
Industry (IND)	The main activity code	see codes[432]	Economic activity	Individual	MIF

8=elderly, permanent disabled and/or unable to work, 9=fulfilling household chores, taking care of children, elderly, sick etc., or busy with care of people (acre responsibilities), 10=other inactive person.

[428] If the joint stock company and its cooperative partners are working effectively, then they are classified as paid, salaried (wageworker), or in other words actually working members of corporations or co-operatives are considered as "employee").

[429] Please refer to the link below for further information (European Communities 2008, accessed 28 July 2019) : https://www.ilo.org/public/english/bureau/stat/isco/isco88/publ3.htm

[430] 1=legislators, senior officials and managers, 2=professionals, 3=technicians and associated professionals, 4=clerks, 5=service workers and shop and market sales workers, 6=skilled agriculture and fishery workers, 7=craft and related trade workers, 8=plant and machine operators and assemblers, 9=elementary occupations, 10=economically not active regarding ES1 definition, 11=economically not active regarding ES2 definition except people classified regarding ES1 criteria.

[432] 1=agriculture, forestry and fishing, 2=mining and quarrying, 3=manufacturing, 4=electricity, gas, steam, air conditioning supply, and water supply; sewerage, waste management and remediation activities, 5=construction, 6=wholesale and retail trade; repair of motor vehicles and motorcycles, 7=transportation and storage, 8=accommodation and food service activities, 9=information and communication, 10=financial and insurance activities, 11=real estate activities, 12=professional, scientific and technical activities, 13=administrative and support service activities, 14=public administration and defence; compulsory social security, 15=education, 16=human health and social work activities, 17=culture, arts, entertainment, recreation, and sport, 18=other service activities,

	of the local unit of the main job (workplace, organization where the individual works in her/his main job) according to NACE REV 2 classification system[431]				
Working hours (WH)	Total weekly working hours in her/his main job[433]	1=1-27 hours per week, 2=28-42 hours per week, 3=43+ hours per week, 4=economically not active regarding ES1 definition, 5=economically not active regarding ES2 definition except people classified regarding ES1 criteria	Economic activity	Individual	MIF
Social security/insurance (SC)	Whether the individual is registered or not to the national social security system/social security institutions due to her/his main job?	1=yes registered, 2=not registered, 3=economically not active regarding ES1 definition, 4=economically not active regarding ES2 definition except people classified regarding ES1 criteria	Economic activity	Individual	MAF
Type of contract (CONTRACT)	Type of contract of employees in the main job	1=permanent job/work contract of unlimited duration, 2=temporary job/work contract of limited duration 3=work occasionally without contract, 4= temporary work held by	Economic activity	Individual	MACMICF

19=economically not active regarding ES1 definition, 20=economically not active regarding ES2 definition except people classified regarding ES1 criteria.

[431]Please refer to the link below for further information (accessed 28 July 2019) :
https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF/dd5443f5-b886-40e4-920d-9df03590ff91?version=1.0

We note that compared to the Board Structure of NACE Rev 2. (the document is available in the link above, on page 57), the categories D and E are regrouped together (corresponding to the category 4 in our list) and the categories T and U are not taken into account in our list. Our list in column 3 is in fact provided by the Turkish National Statistical Office directly in the questionnaire of the survey. Thus, we respect their classification.

[433] Total number of hours usually worked in main job (The regularly paid extra working hours of employees are included.)

		students, pupils, etc. in the vacation time, 5=economically not active regarding ES1 definition, 6=economically not active regarding ES2 definition except people classified regarding ES1 criteria			
Financial state I (FS1)	Monthly living status of the household with the total household income (ability to make ends meet with total monthly household income)	1=with a great difficulty, 2=with difficulty, 3=with some difficulty, 4=fairly easily, 5=easily, 6=very easily	Financial situation	Household	MAF
Financial state II (FS2)	Arrears on mortgage, loan repayments or rent payments in the last 12 months?	1=yes (once), 2=yes (at least twice), 3=no, it did not happen, 4= there is no such kind of payment	Financial situation	Household	MAF
Financial state III (FS3)	Arrears on utility bills (house electricity, water and gas invoices) in the last 12 months?[434]	1=yes (once), 2=yes (at least twice), 3=no, it did not happen, 4= there is no such kind of payment	Financial situation	Household	MAF
Financial state IV (FS4)	Arrears on hire purchase instalments, credit cards or other loan payments in the last 12 months?	1=yes (once), 2=yes (at least twice), 3=no, it did not happen, 4=there is no such kind of payment	Financial situation	Household	MAF
Financial state V (FS5)	Is the household able to afford unexpected financial expenses (approximately 400TL)?	1=yes, 2=no	Financial situation	Household	MAF

[434] Paying through borrowing is considered as paid through own sources.

Financial state VI (FS6)	Is the household able to afford adequately the heating needs of the dwelling? (ability to keep home adequately warm)	1=yes, 2=no	Financial situation	Household	MAF
Financial state VII (FS7)[435]	At least 3 days a week, when required or needed, is the household able to afford a meal with meat, chicken, fish? (equivalent food for vegetarians)	1=yes, 2=no	Financial situation	Household	MAF
Financial state VIII (FS8)	Is the household able to afford new clothes, except secondhand? (replace worn out clothes by some new (not second hand) ones)	1=yes, 2=no	Financial situation	Household	MAF
Financial state XI (FS9)	Is the household able to renew/replace old or worn out furniture?	1=yes, 2=no	Financial situation	Household	MAF
Financial state X (FS10)	Is the household able to afford, when required or needed, a one-week annual holiday away from home	1=yes, 2=no	Financial situation	Household	MAF

[435] This variable includes also the nutrition aspect; however, as there is not any other nutrition related variable then it is included in the dimension of financial situation. For instance, in cases of South Africa GHS 2011 and Mexico MCS 2012, a similar variable is directly included in the dimension of nutrition as at least 2 variables on the nutrition situation of the household are available.

	with all household members? (e.g., hotels, motels, lodgings, summer houses)				
Financial state XI (FS11)	Financial burden of the total housing cost[436]	1= a heavy burden, 2= a slight burden, 3= not burden at all	Financial situation	Household	MAF
Financial state XII (FS12)	Financial burden of the repayment of debts from hire purchases or loans excluding housing costs	1= a heavy burden, 2= a slight burden, 3=not burden at all, 4=there is no such kind of repayment	Financial situation	Household	MAF
Educational attainment of the head of the household (EDUH)	Highest educational attainment of the head of the household	same codes as EDU (see above)	Household composition	Household	MAF
Economic activity of the head of the household (ES1H)	Economic activity status of the head of the household in her/his main job regarding objective criteria of economically active population	same codes as ES1 (see above)	Household composition	Household	MAF
Economic activity of the head of the household (ES2H)	Economic activity status of the head of the household in her/his main job regarding subjective criteria of economically	same codes as ES2 (see above)	Household composition	Household	MAF

[436] Repayment of an interest for a loan or credit on dwelling for owner occupiers or rent paid for tenants, utility bills, heating expenses, penalties for utility bills, collective expenses of apartment and regular repair and maintenance costs are all covered.

	active population				
Employment status of the head of the household (ES3H)	Employment status of the head of the household in her/his main job	same codes as ES3 (see above)	Household composition	Household	MAF
Occupational group of the head of the household (OGH)	Occupational group of the head of the household in her/his main job	same codes as OG (see above)	Household composition	Household	MAF
Working hours of the head of the household (WHH)	Working hours group in which the head of the household is affiliated in her/his main job	same codes as WH (see above)	Household composition	Household	MAF
Social security of the head of the household (SCH)	Whether the head of the household is registered or not to national social security system due to her/his main job	same codes as SC (see above)	Household composition	Household	MAF

Mexico

Módulo de Condiciones Socioeconómicas (Socio-Economic Conditions Survey) 2012

Our analysis concern only people aged 18 years or over. In total, there are 9 dimensions which are: geographic situation, individual characteristics, household composition, dwelling situation, material situation, educational attainment, health, economic activity, and nutrition. The person him or herself or the head of household responds to the questions by giving the information for other members of the household as well- depending on the structure of questionnaire.

Concerning the *variables* will be used in the analysis, the table below shows their features.

Table 26. Features of selected variables from Modulo de Condiciones Socioeconomicas (MCS) 2012, Mexico

Name (code_name)	**Question/Content**	**Codes**	**Dimension**	**Unit**	**Functioning (IF, EF, MAF, MIF, MACMICF)**
Household id (hid)	An identifier code for each household	String variable provided by the INEGI in the data	-	Household	-
Dwelling id (dwid)	An identifier code for each dwelling	String variable provided by the INEGI in the data			-
Personal id (id)	An identifier code for each individual	String variable composed of dwelling id, household id and serial number of the person	-	Individual	-
Serial number (serial)	A household member identifier code for each individual in her/his household	String variable provided by the INEGI in the data	-	Individual	-
Gender (GEN)	The gender of individual	1=female, 2=male	Individual characteristic	Individual	IF
Age (AGE)[437]	Age group in which the person is affiliated, based on the age of the person	1=18-24, 2=25-34, 3=35-44, 4=45-54, 5=55-64, 6=65+ Categorical (Composed)	Individual characteristic	Individual	IF
Elderly (ELD)	Elderly people, aged 65 or over	0=nonelderly, 1=elderly	Individual characteristic	Individual	IF
Ethnicity (ETH)	According to your culture, do you consider yourself indigenous?	1=Yes, 2=No	Individual characteristic	Individual	IF

[437] Based on the question "How old is she?" (Cuantos anos cumplidos tiene). Elderly over 97 years are coded as 97 in the data base according to the questionnaire.

Language (LNG)	For those considering as indigenous, speaking Spanish and Indigenous language or a dialect	1=Speaking only indigenous language or a dialect, 2=no speaking but understanding indigenous language or a dialect, 2= no speaking neither understanding of indigenous language or a dialect	Individual characteristic	Individual	IF
Household size (HS)	The number of people living in the household[438]	1=one person, 2=two persons, 3=three persons, 4=four persons, 5=five or more Categorical (composed) based on the metric information	Household composition	Household	MAF
Children (hh_child)	Household types based on the number of children (<18 yrs.) living in the household	1=no child, 2=1 child, 3=two children, 4=three children, 5=four children, 6=five children or more Categorical (composed)	Household composition	Household	MAF
Gender_Head (GH)	Gender of the head of the household	1=female, 2=male	Household composition	Household	MAF
Age_Head (AH)	Age group in which the head of the household is affiliated, based on her/his age	1=less than 14, 2=14-17, 3=18-24, 4=25-34, 5=35-64, 6=65+	Household composition	Household	MAF

[438] Including domestic workers and guests according to the questionnaire.

Dwelling type (DWE)	Type of main dwelling	1=independent house, 2=apartment in building, 3=housing in neighborhood ("vivienda en *vecindad*[439]"), 4=housing in roof room ("cuarto de azotea"), 5= "Local no construido para habitación" (Not build or not convenient for living)	Dwelling situation	Household	MAF
Wall (WAL)	What material is mostly used for the walls of this house?	1=waste material, 2=cardboard sheet, 3=asbestos or metal foil, 4= reed ("carrizo"), bamboo or palm, 5= "embarro" or wattle and daub ("bajareque"), 6=wood, 7=adobe, 8=partition wall, brick, block, stone, quarry, cement or concrete[440]	Dwelling situation	Household	MAF
Roof (ROF)	What material is mostly used for the roofs of this house?	1= waste material, 2=cardboard sheet, 3=sheet metal, 4=asbestos, sheet 5=palm	Dwelling situation	Household	MAF

[439] In Mexican context, this type of habitation (*vecindades*) corresponds to the subdivision of the abandoned/sold upper class colonial mansions in historic centers of cities (e.g., Mexico City), in a form of several individual rooms with a shared central court, kitchen and/or sanitation facility (with other tenants) cheaply rented to poor, low-earning families. In other words, it is an inner-city slum frequented by poor, low-earning families. For instance, see Lewis's book "The Children of Sanchez" (1961)- The family Sanchez lives in *vecindad* in center of Mexico City.

[440] "Tabique, ladrillo, block, piedra, cantera, cemento o concreto".

		or straw, 6=wood or "tejamanil", 7=rooftop with beams, floor framing ("viguería"), 8=brick red, 9=specific floor or joists with small vault			
Floor (FLO)	What material is mostly used for the floor of this house?	1=land, 2=cement or firm, 3=wood, mosaic or other coating	Dwelling situation	Household	MAF
Sanitation (SAN1)	Whether a shower (regadera) exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Sanitation (SAN2)	Whether a toilet exists or not in the dwelling?	1=yes (not shared, inside dwelling), 2=yes (shared with other dwellings), 3=no toilet	Dwelling situation	Household	MAF
Sanitation (SINK)	Whether a sink (fregadero or tarja) exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Kitchen (KIT)	Whether (independent) kitchen exists or not in dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Water I (WAT1)	Water availability in the dwelling via…	1=piped water system inside the house, 2=piped water system outside the house, but inside the land, 3=water piped public tap (or hydrate, fire hydrant), 4=piped water that they carry from another house, 5=pipe water ("agua de pipa"), 6=water from a well,	Dwelling situation	Household	MAF

		river, lake, stream or other			
Water II (WAT2)	How many days a week does the water arrive at this house?	1=daily, 2=every third day, 3=twice a week, 4=once a week, 5=from time to time when…	Dwelling situation	Household	MAF
Hot water system (HOTW)[441]	Whether a boiler/hot water system (calentador de agua) exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Heating system (HEATS)	Whether a heating system (calefaccion) exists or not in the dwelling?	1=yes, 2=no	Dwelling situation	Household	MAF
Electricity (ELC)	In this house, electric is obtained via…	1=public service? 2=particular plant? 3= solar panel? 4=another source? 5=no electric ?	Dwelling situation	Household	MAF
Energy for cooking (ENG_C)	What is the fuel the household use most for cooking?	1=firewood, 2=coal, 3=tank gas, 4=natural gas or pipe, 5=electricity, 6=other fuel	Dwelling situation	Household	MAF
Drain (DR)	Available drainage system	1=this house has a drain or drain connected to the public network, 2=a septic tank, 3=a pipe that will give a ravine or crevice, 4=a pipe that will give a river, lake or sea, 5=no drainage	Dwelling situation	Household	MAF

[441] The variable HOTW is the same of the variable HWS used in Turkey SILC 2011 analysis. These two variables mean the same thing. Due to a mistake, in Mexico MCS 2012 case, the code HOTW is used instead of HWS.

Rubbish (RUB)[442]	The garbage of this house…	1=picked up by a truck or garbage cart? 2=thrown in the public dump? 3=thrown in a container or deposit? 4=burn? 5=bury? 6=thrown in a field wasteland or street? 7=thrown in the ravine or crevice? 8=thrown into the river, lake or sea?	Dwelling situation	Household	MAF
Household assets (AIRC, TEL, MPHO, INT, CAR, RAD, TV, REF, STONE, WM, IRON, VC, COMP)	Does the household own any of the following? 1) air conditioning (AIRC), 2) landline telephone (TEL), 3) mobile phone (MPHO), 4) internet (INT), 5) car (CAR), 6) radio (RAD), 7) television (tv), 8) refrigerator (REF), 9) gas or electric stove (STOVE), 10) washing machine (WM), 11) iron (IRON), 12) vacuum cleaner (VC), 13) computer (COMP)	1=yes, 2=no	Material situation	Household	MIF or MAF
Education (EDU)	What is the highest level of education that… has successfully ***completed***?	0=illiterate, 1=literate but no diploma, 2=preschool, 3= primary school, 4=secondary school, 5=high school ("preparatoria" or "bachillerato"), 6=superior education as	Educational attainment	Individual	MIF

[442] This variable is included in SA GHS 2011 analysis as a component of the dimension of environment. However, in case of Mexico MCS 2012 as the dimension of environment is not available then the variable RUB is included in the dimension of dwelling situation.

		"normal" [443], 7=superior education as "carrera technica or commercial", 8=university degree ("professional"), master or PhD Categorical (transformed)			
Education_ Head (EDUH)	Educational attainment of the head of the household	Same codes as EDU (see above)	Household composition	Household	MAF
Insurance I (INS1)	Are you currently affiliated or enrolled in Popular Insurance?	1=yes, 2=no	Health	Individual	MAF
Medical care I (MC1)	Are you affiliated or enrolled to receive medical care from any institution?	1=yes, 2=no	Health	Individual	MAF
Health status (HSS)	In the last 12 months, have you got sick or suffered from any pain, physical discomfort or accident that prevented/blocked you from doing your daily activities?	1=yes, 2=no	Health	Individual	MIF
Medical care II (MC2)	For those answered yes to the question of HSS: did you receive medical care?	1=every time, 2=almost every time, 3=almost never, 4=never, 5=a few times	Health	Individual	MAF
Medical care III (MC2_1,…, MC2_19)	For those did not answer every time to the question of MC2: why you did not receive medical care?	see codes[444]	Health	Individual	MAF

[443] A specific education for becoming teacher after obtained high school diploma.

[444]1=no place to go for receiving the service, 2=service is expensive, 3=the person has not enough money to benefit the service, 4=medical unit, clinic, or hospital is far from the dwelling, 5=service was not provided even if the person attended the medical unit, 6=no trust to service, 7=the service place treats badly the person, 8=people in the service do not speak same language as you, 9=no want to go, or no find it necessary, 10=have to wait a longtime to be treated, 11=not get the medicine the person needs, 12=not enough time to attend the medical unit, 13=the medical unit was not open, 14=got an appointment for long time later, 15=no one helped to take the person to the medical unit, 16=no doctor in the medical unit, 17=a member of household prevented the person to go to the medical unit, 18=self-medicated, self-treated, 19=other reason.

Employment status 1 (ES1)	Current employment status (reference period: the past month)	1=employed, 2=employed but temporarily absent, 3=unemployed, 4=not economically active because he or she has rent or rent some property (rento or rent some property), 5=not economically active because he or she is retired, 6=not economically active because he or she is responsible for household tasks at the dwelling, 7=not economically active because he or she is studying, 8=not economically active because he or she is disable, 9=not economically active due to other reasons Categorical (composed)	Economic activity	Individual	MIF
Employment status_Head (ES1H)	Current employment status of the head of the household	Same codes as ES1	Household composition	Household	MAF
Job quantity (JOB)	How many jobs did you have last month?	1=one job, 2=more than one job	Economic activity	Individual	MIF
Job quantity_Head (JOBH)	Whether the head of the household has a one job or more	1=one job, 2=more than one job	Household composition	Household	MIF
Access to food I (AF1)	In the last three months, due to lack of money or resources, have the	1=yes, 2=no	Nutrition	Household	MAF

	household ever run out of food?				
Access to food II (AF2)	In the last three months, due to lack of money or resources, did you ever worry about the food running out in the household?	1=yes, 2=no	Nutrition	Household	MAF
Access to food III (AF3)	In the last three months, have the household ever run out of money or resources to get a healthy and varied diet?	1=yes, 2=no	Nutrition	Household	MAF

Brazil

National Household Sample Survey (PNAD[445]) 2011

Our analysis concern only people aged 18 years or over. In total, there are 7 dimensions which are: geographic situation, individual characteristics, household composition, dwelling situation, material situation, educational attainment, and economic activity. The person him or herself or the head of the household responds to the questions by giving the information for other members of the household- depending on the structure of questionnaire. Before proceeding with the survey, we will provide the map of Brazil for informing about the geographical coverage. Brazil is a Federative Republic composed of 26 states divided into municipalities, and one federal district without municipalities. The federal district located in the Center-West Region of the country, includes the federal capital of the country 'Brasilia', and the seat of government of the Federal District.

[445] Pesquisa Nacional por Amostra de Domicílios

Figure 9. 26 states and 1 federal district of Brazil.
Source (accessed 25 December 2019): https://en.wikipedia.org/wiki/States_of_Brazil

Concerning the *variables* will be used in the analysis, the table below shows their features.

Table 27. Features of selected variables from National Household Sample Survey (PNAD) 2011, Brazil

Name (code_name)	**Question/Content**	**Codes**	**Dimension**	**Unit**	**Functioning (IF, EF, MAF, MIF, MACMICF**
Household id (hid)	An identifier code for each household	String variable (composed)	-	Household	-
Personal id (id)	An identifier code for each individual	String variable (composed)	-	Individual	-
Serial number (case_string)	A household member identifier code for each individual in her/his household[446]	String variable (composed)	-	Individual	-

[446] excluding persons whose condition in the household unit was a pensioner, domestic servant or relative of the domestic employee.

Federation Unit (FED)	The name of the federation unit (composed of 26 states and 1 federal district) where the household lives	see codes below[447]	Geographic situation	Household	EF
Region (REG)	The name of region where the household lives (the variable is composed based on the federation unit variable/FED above)	1=north region, 2=northeast region, 3=southeast region, 4=south region, 5=central-west region Categorical (composed)	Geographic situation	Household	EF
Setting (SET)	The category of setting (urban vs. rural) where the household lives	1= urban – city or town, urbanized area, 2= urban – city or town, underdeveloped area, 3= urban-isolated urban area, 4= rural – rural agglomeration (cluster) of urban extension, 5= rural-rural, isolated, populated, 6= rural – rural, isolated, core, 7=rural – rural household, isolated, other agglomerates, 8= rural – exclusive rural cluster rural area[448]	Geographic situation	Household	EF

[447] 11= Rondônia, 12= Acre, 13= Amazonas, 14= Roraima, 15= Pará, 16= Amapá, 17= Tocantins, 21= Maranhão, 22= Piauí, 23= Ceará, 24= Rio Grande do Norte, 25= Paraíba, 26= Pernambuco, 27= Alagoas, 28= Sergipe, 29= Bahia, 31= Minas Gerais, 32= Espírito Santo, 33= Rio de Janeiro, 35= São Paulo, 41= Paraná, 42= Santa Catarina, 43= Rio Grande do Sul, 50= Mato Grosso do Sul, 51= Mato Grosso, 52= Goiás, 53= Distrito Federal. The first group including 7 states (11, 12, 13, 14, 15, 16, and 17) are part of the North Region; the 9 states, coded as 21, 22, 23, 24, 25, 26, 27, 28, and 29, are part of the Northeast Region; the 4 states, coded as, 31, 32, 33, and 35, are part of the Southeast Region; the 3 states, coded as 41, 42, and 43, are part of the South Region; and the last group including 3 states and 1 federal district, coded as, 50, 51, 52, and 53, are part of the Central-West Region of the country. The numeration of states and federal district is directly kept in the way provided by the IBEGI in the data. More information on cities and states in terms of socio-economic, demographic statistics, please refer to the link below (accessed 25 December 2019):
https://www.ibge.gov.br/en/cities-and-states.html?view=municipio

[448] 1 = URBANA - Cidade ou vila, área urbanizada, 2 = URBANA - Cidade ou vila, área não urbanizada 3 = URBANA - Área urbana isolada, 4 = RURAL - Aglomerado rural de extensão urbana, 5 = RURAL - Agregado rural, isolado, povoado, 6 = RURAL - Agregado rural, isolado, núcleo, 7 = RURAL - Agregado rural, isolado, outros aglomerados, 8 = RURAL - Zona rural exclusiva aglomerado rural

Setting (SET1)	The category of setting (urban vs. rural) where the household lives, based on the variable SET above	1=rural, 2=urban	Geographic situation	Household	EF
Gender (GEN)	The gender of individual	1=female, 2=male	Individual characteristic	Individual	IF
Population group/Race (POP)	The race of individual	1=White, 2=Black, 3=Yellow (Japanese, Chinese, Korean origin, etc.), 4=Brown, 5=Indigenous	Individual characteristic	Individual	IF
Age (AGE)[449]	Age group in which the person is affiliated, based on the age of the person	1=18-24, 2=25-34, 3=35-44, 4=45-54, 5=55-64, 6=65+ Categorical (composed)	Individual characteristic	Individual	IF
Household size (HS)	The number of people living in the household	1=one person, 2=two persons, 3=three persons, 4=four persons, 5=five or more Categorical (composed) based on the metric information	Household composition	Household	MAF
Children (hh_child)	Household types based on the number of children (<18 yrs.) living in the household	1=no child, 2=1 child, 3=two children, 4=three children, 5=four children, 6=five children or more Categorical (composed)	Household composition	Household	MAF
Gender_Head (GH)	Gender of the head of household	1=female, 2=male Dichotomous (composed)	Household composition	Household	MAF
Age_Head (AH)	Age group in which the head of the household is affiliated,	1=less than 14, 2=14-17, 3=18-24, 4=25-34, 5=35-64, 6=65+	Household composition	Household	MAF

[449] The original question is: "What is …..'s date of birth and age in completed years?" (transformed from metric).

	based on her/his age				
Dwelling type I (DWE)	Type of main dwelling	1=house, 2=apartment, 3=room	Dwelling situation	Household	MAF
Wall (WAL)	What is the predominant material in the construction of the exterior walls of this dwelling?	1= masonry, 2=rigged wood, 3=uncoated wall of mud (stucco), 4=harnessed wood, 5=straw, 6=other material	Dwelling situation	Household	MAF
Roof (ROF)	What is the main material used for the roof of this dwelling?	1=tile, 2=stone slab, 3=rigged wood, 4=zinc, 5=harnessed wood, 6=straw, 7=other material	Dwelling situation	Household	MAF
Room (RO)	Number of rooms per person in the household	Metric (composed)	Dwelling situation	Household	MAF
Piped water I (PW1)	Does this household have piped water for at least one room?	1=yes, 2=no	Dwelling situation	Household	MAF
Piped water II (PW2)	The piped water used in this dwelling comes from:	1=general distribution network, 2=well or spring, 3=other	Dwelling situation	Household	MAF
Access to water (WATD)	Access to water, composed of the variables PW1 and PW2	1=piped water system for at least one room in the house and comes from the general distribution network, 2=no piped water system in the dwelling, 3=water for the household needs comes from well or spring, 4=water for the household needs comes from other source Categorical (composed)	Dwelling situation	Household	MAF
Sanitation facilities:	In this household or on the property there is	1=yes, 2=yes but shared with other dwellings, 3=no	Dwelling situation	Household	MAF

bathroom, toilet (SAN)	a bathroom or toilet?	bathroom neither toilet in the dwelling Categorical (composed)			
Drainage of bathroom/toilet (DR)	How is the drainage of this bathroom or toilet made?	0=no bathroom or neither a toilet in the dwelling, 1=sewer or pluvial collection network, 2= septic tank connected to the collecting network sewage or rain, 3=septic tank not collected to the collecting network sewage or rain, 4=rudimentary pit, 5=ditch, 6=directed to the river, lake or sea, 7=other	Dwelling situation	Household	MAF
Waste/rubbish (RUB)[450]	The waste of this household is…	1=collected directly, 2=indirectly collected, 3=burnt or buried in property, 4=thrown on wasteland or patio, 5=thrown in river, lake or sea, 6=other	Dwelling situation	Household	MAF
Energy for lighting (ENG_L)	What is the form of lighting in this dwelling?	1=electric (grid, generator, solar), 3=canister oil, kerosene, gas, 5=other	Dwelling situation	Household	MAF
Energy for cooking I (ENG_C1)	Does this household have a two-burner stove?	1=yes, 2=only one burner stove, 3=none	Dwelling situation	Household	MAF
Energy for cooking II (ENG_C2)	The stove in this household uses predominantly:	1=gas cylinder, 2=piped gas, 3=firewood, 4=coal, 5=electric, 6=other	Dwelling situation	Household	MAF
Energy source for cooking (ENG_C)	Energy source for cooking in the dwelling (composed of the variables	1=one or two-burner stove do not exist in the dwelling, 2=two-burner stove and energy source is piped gas or electric, 3=one-burner stove and	Dwelling situation	Household	MAF

[450] In case of South Africa GHS 2011, this variable is included in the dimension of environment. However, here and in case of Mexico MCS 2012 it is included in the dimension of dwelling situation as there is not other variable(s) available on the environment.

	ENG_C1 and ENG_C2)	energy source is piped gas or electric, 4=two-burner stove and energy source is gas cylinder or firewood or coal, 5=one-burner stove and energy source is gas cylinder or firewood or coal Categorical (composed)			
Water filter (WFIL)	Does this dwelling have any type of water filter?	1=yes, 2=no	Dwelling situation	Household	MAF
Household assets (TEL, MPHO, RAD, TV, REF, WM, COMPINT, CAR)	Does the household own any of the following in the dwelling? 1) a functional/working landline telephone (TEL), 2) a functional/working cellular telephone (MPHO), 3) radio (RAD), 4) color tv or black and white tv (TV), 5) refrigerator (REF), 6) washing machine (WM), 7) microcomputer, and access to internet via microcomputer (COMPINT), 8) car (CAR)	1=yes, 2=no	Material situation	Household	MIF or MAF
Household assets (COMPINT)	Does the household own a computer and an internet for the computer in the dwelling?	1=yes, both computer and internet, 2=computer but no internet for computer, 3=no computer	Material situation	Household	MIF or MAF

		Categorical (composed)			
Education (EDU)	What is the highest level of education that… has successfully ***completed***?	0=illiterate, 1=literate but no formal schooling, 2= literacy for young and adults or literacy class[451], 3=nursery, 4=maternal or kindergarten, 5=primary, 6=middle school/secondary (1 ciclo or 2 ciclo), 7= Regular elementary school or 1st grade (1 grau), 8=high school regular or 2nd grade (2 grau), 9= Youth and adult education or supplementary elementary school or 1grau, 10= Youth and adult education or supplementary high school or 2 grau, 11=University (undergraduate, Master, or PhD) Categorical (transformed)	Educational attainment	Individual	MIF
Education_Head (EDUH)	Educational attainment of the head of the household	Same codes as EDU (see above)	Household composition	Household	MAF
Economic activity status (ES1)	Current economic activity status (based on the week of reference[452])	1=economically active, 2=not economically active Categorical (composed)	Economic activity	Individual	MIF

[451] Alfabetização de jovens e adultos or Classe de alfabetização – CA (source: questionnaire).

[452] 18-24 September 2011. For being considered as an economically active, the person should fulfill at least one of the following 4 conditions: (i) working on the reference week; (ii) temporarily being absent; (iii) performing tasks in cultivation, fishing or animal husbandry, intended for the own feeding of people living at home during the reference week; (iv) performing tasks in construction of building room, well or other construction works, intended for the own use of people living at home in the reference week.

Economic activity status_Head (ES1H)	Current economic activity status of the head of the household	Same codes as ES1 (see above)	Household composition	Household	MAF
Multiple jobs (JOB)[453]	Having more than one job (based on the week of reference)	0=not economically active, 1=only one job, 2=two jobs, 3=more than two jobs Categorical (composed)	Economic activity	Individual	MIF
Occupational group (OG)[454]	Occupational group (type of job/description of job) during the reference week	0=not economically active, 1=Managers, 2=Science and art professionals, 3=Mid-level technicians, 4=administrative workers, 5=service workers, 6=trade vendors and service providers, 7=agricultural workers, 8=goods and service production and repair and maintenance workers, 9=armed and auxiliary members Categorical (composed)	Economic activity	Individual	MIF
Occupational group_Head (OGH)	Type of job of the head of the household	Same codes as OG (see above)	Household composition	Household	MAF
Industry (IND)	Main activity/industry group during the reference week	0=not economically active, 1=agriculture, 2=other industrial activities, 3=manufacturing, 4=construction, 5=trade and repair, 6=accommodation and food, 7=transport, storage and communication, 8=public administration, 9=education, health	Economic activity	Individual	MIF

[453] The following economic activity related variables concern only the main job.
[454] The following economic activity related questions concern only people working in non-agricultural sector.

		and social services, 10=domestic services, 11=other collective, social and personal services, 12=other activities Categorical (composed)			
Employment status (ES3)	Employment status in the main job during the reference week	0=not economically active, 1=employee with signed work card, 2=military, 3=statutory civil servant, 4=other employee without signed work card, 5= domestic worker with signed work card, 6=domestic worker without signed work card, 7=own account worker, 8=employer, 9=worker in production for own consumption, 10= construction worker for own use, 11= unpaid Categorical (composed)	Economic activity	Individual	MIF
Working hours (WH)	Total weekly usually working hours in her/his main job during the reference week[455]	0=not economically active, 1= up to 14 hours, 2= 15 to 39 hours, 3= 40 to 44 hours, 4= 45 to 48 hours, 5= 49 hours or more Categorical (composed)	Economic activity	Individual	MIF
Social security (SC)- Instituto de Previdencia[456]	Did the person contribute to the social security due to her/his	0=not economically active, 1=yes, 2=no	Economic activity	Individual	MAF

[455] Total number of hours usually worked in main job (the regularly paid extra working hours of employees are included). Here, we use directly the variable composed and made available in the data by the IBGE. In previous case studies such as Turkey, we created the variable based on the International Labour Organization (ILO) practice in the computation of categories of the weekly working hours for economically active population during the reference week from the national labour force surveys.
[456] Please refer to the link below for more information on the Institute (accessed 4 August 2019):

	main job during the reference week?	Categorical (composed)			
Benefits (BEN_1, BEN_2, BEN_3, BEN_4, BEN_5)[457]	In her/his main job during the reference week, does the person working as an employee or a domestic worker, receive the following benefits? 1) housing assistance/help (BEN_1), 2) food assistance/help (BEN_2), 3) transport assistance/help (BEN_3), 4) education or nursery assistance/help (BEN_4), 5) health or rehabilitation assistance/help (BEN_5)	0=not economically active, 1=yes, 2=no, 3=economically active (other employment status, ES3=7, 8, 9, 10, 11) Categorical (composed)	Economic activity	Individual	MAF
Unpaid worker (UPW)	Concerning own account workers and employers, do you have at least one unpaid worker in this job?	0=not economically active, 1=yes, 2=no, 3= economically active (other employment status, ES3=1,2,3,4,5,6,9,10, 11)	Economic activity	Individual	MACMIC F
National Register of Legal Entities (CNPJ)[458]	Concerning employees in private sector, own account workers and employers (except domestic	0=not economically active, 1=yes, 2=no, 3=other employment status	Economic activity	Individual	MAF

https://www.ibdp.org.br/

[457] Only for employees and domestic workers (as a main job).

[458] For more information, please refer to the link below (accessed 4 August 2019): http://receita.economia.gov.br/orientacao/tributaria/cadastros/cadastro-nacional-de-pessoas-juridicas-cnpj

	workers), the enterprise where you work or own is registered at CNPJ?				
Invoice (INV)	Concerning employees in private sector (except domestic workers), the enterprise where you work provide invoices to the clients?	0=not economically active, 1=yes, 2=no, 3=other employment status	Economic activity	Individual	MAF
Paycheck (PAYC)	Concerning employees in private sector (except domestic workers), the enterprise where you work provide paychecks to the employees?	0=not economically active, 1=yes, 2=no, 3=other employment status	Economic activity	Individual	MAF

4.3 Steps to follow in the analysis

In our four case studies, to measure multidimensional poverty we apply the following four steps described below:

- **Step 1 (step1 analysis): Detecting *multidimensional poverty patterns* based on the overall deprivation index.** To obtain these patterns, *first* we will create the indicators to be used for measuring poverty as a function of the questions available in the surveys. *Second*, we will group the indicators by dimension as a function of the content of indicator. There are many dimensions of life and at least one indicator in each dimension. Each indicator has the same weight. *Third*, we will count, for each single individual in the data, his or her deprivation through all indicators. To identify whether the individual is deprived or not regarding the indicator, a deprivation threshold will be determined for each indicator. A deprivation will be counted as 1 and a non-deprivation will be counted as 0. In other words, each single individual in the data will have an overall deprivation score between zero and the total number of indicators. For instance, if there are 30 indicators the maximum overall deprivation score is 30 and the minimum overall deprivation score is 0. *Fourth*, we will create multiple poverty groups on the basis of the distribution of the overall deprivation scores given that each single individual in the data will be part of a poverty group. There are three main groups, namely zero-poverty, in-between poverty, total-poverty. The second group, in-between poverty, includes several groups within. The zero-poverty group represents the most prosperous group, which do not have deprivation regarding all indicators. And, the total-poverty group is the most disadvantaged group, which experiences deprivation in all indicators. *Fifth*, we will effectuate classification trees to obtain multidimensional

poverty patterns of poverty groups given that these patterns reflect capabilities of individuals. In each dimension the person will have zero, or one, or more than one deprivation, which will correspond to the capability of the person regarding the dimension and the indicators of the dimension. More non-deprivations more real liberty to lead the life the person values. The overall deprivation score shows the magnitude of poverty and is already proposed by Alkire and Foster in 2007 (2007, rev. 2009b, 2011a). But, the originality of our work is to discover both the magnitude and the nature of poverty by finding out the combinations of deprivations and non-deprivations. As there are many combinations in case of many indicators, classification trees method will help us to extract the most relevant combinations. And these combinations are called the patterns according to data mining terminology. The indicators used in the analysis are mentioned in case studies.

- **Step 2 (Step2 analysis): Detecting *dimensional poverty profiles* of poverty groups.** At this stage, we will proceed to find out dimensional poverty profiles within each poverty group on the basis of the indicators used in the overall deprivation index, i.e., every single individual in the data has a profile regarding each dimension of the overall deprivation index. For instance, if the dimension of education has 2 indicators, then the possible profiles are: {deprived, deprived}, {deprived, non-deprived}, {non-deprived, deprived}, {non-deprived, non-deprived}, given that first profile represents total educational poverty, the second and third ones correspond to in-between educational poverty, and the last one designates zero educational poverty. Or, the profiles can be coded by 0 and 1 such as 11 for {deprived, deprived}, 10 for {deprived, non-deprived}, 01 for {non-deprived, deprived}, and 00 for {non-deprived, non-deprived}.

- **Step 3 (step3 analysis): Detecting *various profiles* of poverty groups.** The goal is to reveal various profiles of poverty groups on the basis of indicators that are not used in the overall deprivation index; for instance, profile in terms of geographical situation (e.g., setting/urban vs. rural, province), individual, household characteristics (e.g., age groups, gender, household size), economic activity (e.g., employment status, formal vs. informal job), social grants (e.g., old age grant, disability grant), and others depending on the structure of questionnaires. Here, the profiles are not expressed in terms of deprivation and non-deprivation. The profiles are composed of modalities/categories of the indicators. Deprivation thresholds are not determined for the indicators that are not incorporated in the overall deprivation index.

- **Step 4 (step4 analysis): Sub-group/decomposition analysis-detecting *multidimensional poverty profiles* of specific groups.** The aim is to focus specifically on some groups such as elderly population, low-paid workers, retired, long-term unemployed, disabled people, older workers, welfare recipients, chronic ill people, and others depending on the structure of questionnaires.

The profile analysis is already proposed by Fattore et al. (2011b, 2012), Fattore and Maggino (2014), Fattore (2016). The methodology proposed by the authors (POSET onwards) is based on the partial order theory and analyzes poverty (multidimensional ordinal deprivations) situation of individuals via the evaluation process, which takes the poverty line as the benchmark and produces a poverty degree to each single profile in the data. Beycan and Suter (2017) have applied this methodology (via the R package PARSEC- see Fattore and Arcagni

2014) to examine multidimensional poverty in Switzerland. Beycan and Suter (2017) have showed that POSET is very useful to better understand the complexity and the extent of poverty, notably regarding deprivation situations of people above the poverty line, and combinations of deprivations and non-deprivations. Compared to the popular Alkire and Foster (2007, rev. 2009b, 2011a) counting approach that is used to produce the global multidimensional poverty index (MPI), the POSET gives full information on the available profiles in the data. The AF methodology takes only into account people below the poverty line and censor all other information, which does not allow to see the other poverty situations in the data. But also, the AF methodology does not show the profiles of the poor. However, the POSET methodology is not very suitable in case of many variables because it becomes very difficult to interpret a large variety of profiles, and poverty degrees may change when the poverty line changes. In our study, we employ classification trees to find out the most relevant combinations of deprivations and non-deprivations as many profiles exist due to many indicators of poverty. We extract profiles in steps 2, 3, and 4 of the methodology- where the number of indicators is significantly less compared to step 1. And we don't determine poverty line; instead, we search for different configurations of poverty existing in the society concerning from the most prosperous to the least prosperous people. At this point, the classification trees method shows the most influential combinations of deprivations and non-deprivations distinguishing different poverty groups, which also indicates the inequality between poverty groups.

5. South Africa: an exemplary case study

In this chapter, we will focus on the measurement of multidimensional poverty in South Africa by employing the methodology we mentioned previously. We will show all the details of computations including R codes. However, the complete STATA do-file is available upon request, and some descriptive statistics tables are available in Appendix 9.9 Moreover, we remind that due to the length of thesis, the other three case studies, Turkey, Mexico, and Brazil are given in the Appendix 9.4, 9.5, 9.6, respectively.

5.1 Introduction

In case of South Africa, we will use the General Household Survey (GHS) of 2011. According to our methodology, we will carry out several analyses. The list of variables is given in section 4.2. The variables provided in the list won't be used all at the same time; depending on the types of analysis, we will select indicators. The short names of variables (see the codes in parenthesis in first column of the table of indicators) will be used in the outputs including graphics. Before describing the content of models in classification tress (first step of the methodology) such as regressors and other parameters, and presenting the results, we will provide the quality report of the variables.

Quality report of variables

First of all, we merged the household and individual data together and created unique household (hid) and personal identifiers (id). In merged data, there are 45 persons (from 10 households) having no household information. Due to that reason, these individuals were dropped from the merged data. The merged data contains 25,086 households and 93,389 persons in total. We created all indicators defined in the table of indicators (see section 4.2), and they are later saved as the derived data. We took the following 49 variables as the foundation/basis for analysis (they will be used to construct the overall deprivation index), which means that all persons not having full information of these variables were dropped from the derived data. The main variables are, with their short names, the following: DWE, ROF, WAL, ROFC, WALC, RO, KIT, SAN, SAN_SH, SAN_LOC, WATD, WATDIST, SW, CW, GTW, FFBSW, WAT_T, WATI_D, ELCCUT, ENG_C, ENG_H, ENG_L, ENVP1, RUB, LIT, WP, AP, LD, NOI, POST, TEL, INT_1, INT_2, INT_3, INT_4, INT_5, INT_6, INT_7, TV, RAD, REF, ADN, CHN, NU1, NU2, NU3, NU4, EDU, MC. After having removed missing values, 74,489 individuals remained in the data corresponding to 79.76 per cent of the initial data. Among them, there are 47,117 individuals aged 18 or over. At the initial data, there were 58,806 individuals aged 18 or over; so that, we've lost 20 per cent of the initial data for this target group. Additional to 49 variables, we created several other variables for third and fourth steps of the analysis. We obtained the quality report of the variables in statistical software R (R onwards). We note that the full quality report is given in Appendix 9.9.

Via the package Hmisc[459] in R, we obtained main descriptive statistics based on the non-weig hted data[460]. The short name of the variables corresponds to their codes mentioned in section 4

[459] Please refer to the link below for more information on this package in R (accessed 14 September 2019): https://cran.r-project.org/web/packages/Hmisc/index.html

[460] To obtain estimated numbers at national level typically for the national reports, the obtained figures can be multiplied by the appropriate weight provided in the data by the Statistics South Africa. But, in our analysis, we use the national sample, which is also an appropriate representative of the population in the country.

.2. For instance, for the variable hid, there are 20,041 unique (distinct) values in the data. In fa ct, the data contains 47,117 individuals through 20,041 households. There are no missing valu es for the hid. This variable is a string variable composed of the particular PSU number is asse ssed to each household as a result of the sampling design. Each household should have her ow n household id number and the data should not contain any duplicates in terms of hid (except a ll household members from the same household will have the same hid number). The data is s orted by the variable hid. We use hid as well as id and serial numbers for verifying no duplicat es in the data, or for creating groups and/or variables by household.

In the descriptive statistics outputs (see Appendix 9.9), 'n' means the total number of observat ions (i.e., the total number of persons interviewed), 'missing' means the number of missing va lues for the variable (i.e., the number of persons not having a value for the concerned variable) , and 'distinct' means the number of distinct/unique values the variable has in the data. The va riable 'serial' is a string variable taking values between 1 and 21 in the data. Every person in t he data should have a serial number. For instance, if in a household there are 5 persons, so that every household member should be affiliated with a number between 1 and 5 and everyone sh ould have a different number. In general, the head of the household takes the number 1. This v ariable is particularly important for creating unique personal identifiers (id numbers). As we c an see from the outputs, everyone in the data has his or her own serial number which varies fr om 1 to 21.

For qualitative variables (except hid, id), the output includes the 'value' (i.e., the values that th e variables takes), 'frequency' (i.e., how many individuals in the data take the value), and 'pro portion' (i.e., the ratio of frequency of the value by the total number of observations in the dat a). For quantitative variables, the output provides mean, Gini's difference (Gmd), quantiles, m inimum and maximum values.

5.2 First step (step1 analysis): detecting multidimensional poverty patterns

Here, the goal is to take into account simultaneously various functionings to reflect the actual capabilities of the person at a given time. Instead of focusing on overall poverty rate, we assert that patterns of capabilities as a combination of functionings yield a more comprehensive picture of living conditions the person has in his or her life. People deal with multiple aspects of life and face various conditions to continue their life. Their daily encounters including macro level functionings surrounding them define their look at life, the conception of things in their mind, their behaviors and decision mechanism including how to process for achieving what they value and have right to value. Based on the plurality of different living conditions, poverty manifests intrinsically in different configurations. A person may experience disadvantages and deprivations or advantages and achievements, or more realistically combinations of advantages, disadvantages, deprivations and achievements, which may vary over time. It is extremely important to give priority to people suffering greater deprivations and disadvantages; however, to understand the magnitude and the nature of such situations a global view on living conditions should be taken into account in society.

The General Household Survey of South Africa provides a large variety of questions, which allows to build multiple indicators reflecting different aspects of life. According to questionnaire, some questions are asked to every interviewee and some others are defined for more specific purposes such as addressing children, pregnant women, or people attending educational institution. Sub-group/decomposition analysis in terms of gender, province,

settings are part of usual additional analysis for enriching what hides behind the overall statistics. Other specific analysis focusing on a particular target group such as children can be also realized. Nevertheless, it is important to start with a broad analysis and then get into detail with decomposition analysis. We have decided to examine separately children's situation in last stage and to start with an *overall* deprivation analysis targeting people aged 18 or over. For such analysis, questions concerning every person aged 18 or over should be taken in consideration. But the questions concerning only a specific group such as economic activity related questions (i.e., which are asked only to individuals performing an economic activity during the reference period mentioned in the questionnaire) or unemployment related questions (i.e., which are asked only to individuals not actually performing an economic activity but seeking a job during the reference period) are not included in the overall analysis. The questions addressing specific target groups (and not addressing the whole 18+ population) are used in decomposition/sub-group analysis. In this line, our first analysis includes the following dimensions: dwelling situation where the household lives, material situation of the household, educational attainment of household members, medical coverage/health situation of the household members, environmental conditions where the household lives, and nutritional conditions of the household members. Except educational attainment and medical coverage, other questions were asked at household level in the survey. The household should have the same records for each member of the household concerning the household level questions such as dwelling standards. The 49 variables mentioned previously are the foundation/basis of the overall deprivation analysis (see section 4.2 for the content of variables). By transforming to string character these 49 variables, we obtained the actual capabilities profile of each person in the target group. The variable of actual capabilities profile is called PATH (see STATA do file-available upon request). We simply combined the scores of variables in the following order[461]:

```
PATH=DWE+ROF+WAL+ROFC+WALC+RO_D+KIT+SAN+SAN_SH+SAN_LOC+WAT
D+WATDIST+SW+CW+GTW+FFBSW+WAT_T+WATI_D+ELCCUT+ENG_C+ENG_H+
ENG_L+ENVP1+RUB+LIT+WP+AP+LD+NOI+POST+TEL+INT_1+INT_2+INT_3+INT_
4+INT_5+INT_6+INT_7+TV+RAD+REF+ADN+CHN+NU1+NU2+NU3+NU4+EDU+MC
```

[461] In technical aspect it is important to mention that: (i) the variable RO_D is used instead of RO. The variable RO is numeric reflecting the number of rooms per person in the household. This variable takes rational numbers, which is not suitable for obtaining a string variable. And, for obtaining a profile variable taking simultaneously the answers of all variables together, we need string variables. For that reason, we created RO_D that takes the value 1 if the person lives in a dwelling with a score of RO inferior than 1 (i.e., the number of rooms per person in the household is less than 1), 0 otherwise. The variable RO_D has no missing values. 58.14 per cent of 47,117 individuals do live in a dwelling with a ratio of available room per person less than 1; (ii) the variables DWE, WAL, EDU have both one- and two-digit figures; for instance, the answer options for DWE are from 1 to 11 (see section 4.2). For that reason, we transformed them to string variables, and we added "0" in front of the figure if she has one digit (e.g., "05" if the answer is "5"). When we combine string variables together (e.g., hid, id, PATH-see do file, which is available upon request), the scores of variables are just combined one another, and they are not summed (because they are not numeric). The combination of the two numeric variables V1 with the value 5, V2 with the value 2 will give the numeric score 7. Whilst the combination of the two string variables V1 with the category "5" and V2 with the category "2" will give the result "52", which does not correspond to the numeric variable fifty-two but it is a result of two string variables' scores coming side by side.

```
PATH
    n     missing    distinct
47117       0         35761
```

As we can see from the output above, there are no missing values and there are 35,761 different capabilities profile among 47,117 individuals. The length of variables in terms of character is very long due to the fact that 49 variables are considered. For instance, a person in the data has this PATH score: 0101012200122025111133234121222212222222222211361111032. This means that this person lives in a dwelling/house or brick/concrete block structure on a separate stand or yard or on farm (the category 1 of DWE, the category 1 of ROF, the category 1 of WAL) with weak conditions both of roof and wall (the category 2 for ROFC and WALC) and so on. Based on such profile, we can get an entire picture of the capabilities of the person reflecting his or her advantages, disadvantages, deprivations, and achievements through the 49 items. However, if we think that there are 35,761 different ways of actual capabilities experienced by the individuals in the data it becomes clearly very complex and not efficient for extracting the useful information. For that reason, our goal is to find out the most informative capabilities patterns in the data, which will be also a practical way/option for policy makers to define priorities to improve the quality of life of people. As previously mentioned, discovering patterns will be realized by employing classification trees method.

We carry out our first classification trees-based capabilities analysis via the overall deprivation index. First of all, we defined a deprivation line/threshold for each of the 49 variables. Based on that line, a person suffering deprivation regarding the variable takes the score 1, and 0 otherwise. Taking in consideration all variables, we counted the number of deprivations the person suffers to obtain his or her overall deprivation score. In this regard, we can say that we followed the way used in building Mexico's social deprivation index which corresponds to a simple *additive* index "calculated by means of a linear combination of the variables that measure deprivations, and where all of these variables bear the same relative importance. The resulting index reflects the number of unfulfilled rights for any person" (CONEVAL 2010, 12). Then, as a function of the distribution of the overall deprivation scores emerging in the data, we created several poverty groups. At last stage, we found the most influential capabilities patterns in the data distinguishing (poverty) groups. It is important to note that in fact the overall deprivation index contains 33 variables and not 49 because we created new variables composed of two or three already existing variables in the list of 49. To make clear deprivation lines and composition of new variables, we will now provide their description in the table below. The table contains 4 columns: first column gives the code name of the variable given that each variable represents a functioning, second column mentions deprivation line regarding the variable, third column indicates the utility of the functioning for being capable of living a decent life according to universal human rights chart, and last column contains remarks about the variable. Referring to universal human rights chart, the reason behind the selection of the variable for measuring capabilities emerges. In other words, this shows the utility of the functioning for being capable to live a decent life. According to universal human rights chart, many functionings should be taken in consideration to evaluate capabilities of the person; however, data structure and survey questionnaire are the two crucial points that define availability of variables to be used in the analysis. If the survey is rich enough in terms of questions asked to interviewees, possibility of having a large range of functionings to measure capabilities will be higher. Greater functionings greater aspects of life and greater living

conditions will be detected to find out various configurations of poverty. Regarding that, it is important that each functioning reflects a peculiar element of life and that multiple functionings do not drive to conceptual overlap.

Generally, studies using composite indices for measuring multidimensional poverty (such as MPI) (i) test also robustness of the index score (i.e., headcount ratio of poverty) in terms of whether it is sensitive to variations in arbitrarily defined poverty thresholds; (ii) produces adjusted index score by taking in consideration intensity of poverty; (iii) computes severity of poverty by taking in consideration inequality; (iv) computes poverty gap-the distance to poverty threshold for people having an index score under the poverty line; and (v) checks the quality of the index in terms of correlation coefficients between the index and its indicators. Some studies use structural equation modelling such as principal component analysis to build an index based on select variables. In our study, it is not necessary to process with these additional operations because first of all, we don't use an arbitrarily poverty line assuming that not only a specific part of the society suffers a certain type of poverty given that poverty is a dynamic phenomenon taking different shapes in terms of time and living conditions-poverty configuration is in fact appear in each stratum of society; second, our goal is not to focus on head count ratio and its adjusted derivatives but instead focus on content/shape of poverty; third, we use overall deprivation index as an intermediate tool for distinguishing different poverty groups in society to find out their different poverty configurations-so that, the aim is not to reveal poverty statistics directly based on the overall index; fourth, by employing classification tree method we find out how the poverty groups differ between them in terms of combinations of poverty indicators, which intrinsically gives the information on relevant inequality existing between groups.

Table 28. Deprivation lines and composition of variables for step 1 of the methodology: overall deprivation index and multidimensional poverty patterns, data: South Africa GHS 2011

Variable (each variable represents a functioning)	**Deprivation line[462]**	**The utility of functionings for having the capability of living a decent life according to Universal Human Rights Chart[463]**	**Remark**
DWE_D	Deprived (DWE_D=1) if DWE is 2,7,8,9,10,11.	Article 25: functioning of living with an adequate housing.	The official report of the survey mentions traditional and informal types of location vs. the formal locations. We took into account this criterion. In addition to traditional and informal types of locations, people living in backyards are also considered as being deprived of good quality

[462] For the codes and categories, see the table provided in section 4.2.
[463] Please refer to the link below for further information (accessed 12 February 2020) : https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf

			dwelling (i.e., living in not adequate dwelling).
ROF_D	Deprived (ROF_D=1) if ROF is 3,4,5,6,7,8,10,11,12.	Article 25: functioning of living with an adequate housing.	It is a complementary variable to DWE. Living in a dwelling having a roof made of corrugated iron/zinc, wood, plastic, cardboard, mixture of mud and cement, wattle and daub, mud, thatching, asbestos are considered being deprived of living in a dwelling with good quality roof, compared to bricks, cement block/concrete, tile and asbestos. Notably, asbestos is quite considered as a toxic material. Concerning corrugated iron/zinc, wood, plastic, cardboard and thatching it is very important its quality. We don't have full information on their quality, but we can consider them as a risk zone. The roof condition variable (ROFC) is a good complement for that.
WAL_D	Deprived (WAL_D=1) if WAL is 3,4,5,6,7,8,10,11,12.	Article 25: functioning of living with an adequate housing.	Same comment as ROF_D. The wall condition variable (WALC) is a good complement for that.
ROFC_D	Deprived (ROFC_D=1) if ROFC is 1,2,3.	Article 25: functioning of living with an adequate housing.	
WALC_D	Deprived (WALC_D=1) if WALC is 1,2,3.	Article 25: functioning of living with an adequate housing.	
RO_D	Deprived (RO_D=1) if the number of rooms per person in the household is inferior than 1.	Article 25: functioning of living with an adequate housing.	Excluding kitchen, bathroom, garden, and entranceway, in normal conditions every single person should have at least 1 room for her/his privacy.
KIT_D	Deprived (KIT_D=1) if KIT is 2.	Article 25: functioning of living	

		with an adequate housing.	
SAN_D	Deprived (SAN_D=1) if there is no toilet in the dwelling, or if toilet is not inside the dwelling, or if toilet is shared with other dwellings, or if toilet is chemical, pit latrine/toilet with ventilation pipe, pit latrine/toilet without ventilation pipe, or bucket toilet.[464]	Article 25: functioning of living with an adequate housing.	This variable is composed of three variables SAN, SAN_SH, SAN_LOC[465].
WATD_D	Deprived (WATD_D=1) if the main source of drinking water is not from the piped (tap) water in dwelling, or if the water source is not in the dwelling, or if the water is not safe to drink, or if the water is not clear, or if the water is not good in taste, or if the water is not free from bad smells, or if the water is treated by the household members before using it.[466]	Article 25: functioning of living with an adequate housing.	This variable is composed of 7 variables WATD, WATDIST, SW, CW, GTW, FFBSW, WAT_T[467]
ELC_D	Deprived (ELC_D=1) if the household has no connection to the mains electricity supply.	Article 25: functioning of living with an adequate housing.	
ENG_C_D	Deprived (ENG_C_D=1) if ENG_C is 0,4,5,6,7,8.	Article 25: functioning of living with an adequate housing.	
ENG_H_D	Deprived (ENG_H_D=1) if ENG_H is 0,4,5,6,7,8.	Article 25: functioning of living with an adequate housing.	

[464] The release of Statistics South Africa (2011, 47) has mentioned that "hygienic toilet facility refers to flush toilet, chemical toilet or pit latrine with ventilation pipe". But still, we can say that flush toilet is the best option. Thus, we considered chemical toilets, pit latrine with ventilation pipe as being deprived of adequate sanitation.

[465] See the STATA do file (available upon request) for programming.

[466] The release of Statistics South Africa (2011, 47) has mentioned that "Piped water in dwelling or on site is piped water inside the household's own dwelling or in their yard. It excludes water from a neighbor's tap or a public tap that is not on site". But we considered piped water on-site or in yard as being deprived of adequate access to drinking water because people should be able to have access to the main source of water inside their wedding and they should not be obligated to go outside during the night, or hard weather conditions, or even in the day.

[467] See the STATA do file (available upon request) in the Appendix for programming.

ENG_L_D	Deprived (ENG_H_D=1) if ENG_L is 0,3,4,5,6,7,8.	Article 25: functioning of living with an adequate housing.	
RUB_D	Deprived (RUB_D=1) if RUB is 2,4,6,7,8,9, or if ENVP1 is 1.	Article 25: functioning of living with an adequate housing.	This variable is composed of 2 variables RUB and ENVP1[468]
LIT_D	Deprived (LIT_D=1) if LIT is 1.	Article 25: functioning of living with an adequate environment.	
WP_D	Deprived (WP_D=1) if WP is 1.	Article 25: functioning of living with an adequate environment.	
AP_D	Deprived (AP_D=1) if AP is 1.	Article 25: functioning of living with an adequate environment.	
LD_D	Deprived (LD_D=1) if LDP is 1.	Article 25: functioning of living with an adequate environment.	
NOI_D	Deprived (NOI_D=1) if NOI is 1.	Article 25: functioning of living with an adequate environment.	
POST_D	Deprived (POST_D=1) if POST is 3,4,5,6,7,8.	Article 25: functioning of living with an adequate environment and social service. It is not necessary that post service is provided by the State, but at least State should make sure that every single citizen has access to such service.	
TEL_D[469]	Deprived (TEL_D=1) if TEL is 4,5,6,7.	Article 25: functioning of living	Due to an increasing number (tendency) of usage

[468] See the STATA do file (available upon request) for programming.

[469] We also note that for material deprivation (tel, internet, mobile phone, tv, and so on), there are 2 answer types determined in the questionnaire: yes, household have it; no, household does not have it. We don't have information about the reason of not having the item in the household. Many reasons are

		with an adequate housing in terms of material and social service regarding availability of landline network in the neighborhood where the person lives. It is not necessary that telephone, internet networks are provided by the State, but at least State should make sure that every single citizen has access to such services.	of mobile phone, the use of a landline phone at home may decrease or people may find unnecessary the use of a landline phone at home. However, we keep this variable as a one of the (main) tools of communication because still some parts of the country such as rural areas may lack access to mobile phone service, and we don't have information whether household choose mobile phone rather than landline phone and if yes, why.
INT_D	Deprived (INT_D=1) if INT_1, INT_2, INT_3 INT_4, INT_5, INT_6, and INT_7 is 2.	Article 25: functioning of living with an adequate housing in terms of material and social service regarding availability of internet service in the neighborhood where the person lives.	This variable is composed of 7 variables INT_1, INT_2, INT_3, INT_4, INT_5, INT_6 and INT_7. Each of them represents a way of reaching the internet. For being deprived, we considered that the person should have no access to each of them simultaneously.
TV_D	Deprived (TV_D=1) if TV is 2.	Article 25: functioning of living with an adequate housing in terms of material.	
RAD_D	Deprived (RAD_D=1) if RAD is 2.	Article 25: functioning of living with an adequate housing in terms of material.	
REF_D	Deprived (REF_D=1) if REF is 2.	Article 25: functioning of living with an adequate	

possible such as personal willing to not have it, or a family member does not allow other member(s) to have it, or household head decides for all members of the household, or there is no accessible point/service to buy it, and so on. Because of that, we consider the second answer option as an indicator of deprivation regarding the item. If we had full information whether the person has in fact the full liberty to have it if he or she wants it but simply he or she does not prefer to have the item; then, we could not consider such possibility as an indicator of deprivation.

		housing in terms of material.	
AND_D	Deprived (AND_D=1) if AND is 2,3,4,5.	Article 25: functioning of living with an adequate nutrition (food).	Households with no adults are directly considered no deprived.
CHN_D	Deprived (CHN_D=1) if AND is 2,3,4,5.	Article 25: functioning of living with an adequate nutrition (food).	Households with no children are directly considered no deprived.
NU1_D	Deprived (NU1_D=1) if NU1 is 1.	Article 25: functioning of living with an adequate nutrition (food).	
NU2_D	Deprived (NU2_D=1) if NU2 is 1.	Article 25: functioning of living with an adequate nutrition (food).	
NU3_D	Deprived (NU3_D=1) if NU3 is 1.	Article 25: functioning of living with an adequate nutrition (food).	
NU4_D	Deprived (NU4_D=1) if NU4 is 1.	Article 25: functioning of living with an adequate nutrition (food).	
EDU_D	Deprived (EDU_D=1) if EDU is 0,1,2,3,4,5,6,9,13	Article 26: functioning of having an educational attainment[470].	If the person is illiterate, or literate but no schooling, or if no grade 12, or if NTC1 or NTC2 (vocational schooling level 1 and 2- max. is 6), then the person is considered being deprived of sufficient educational attainment. Concerning vocational schooling, there are 2 diplomas: at the end of NTC3 and another is at the end of NTC6. NTC3 can be considered as an equivalent of grade 12.

[470] Article 26 mentions that "Everyone has the right to education. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical and professional education shall be made generally available and higher education shall be equally accessible to all on the basis of merit". However, we argue that at least a general high school diploma or its equivalent in terms of vocational technical high school is a fundamental necessity for people to obtain opportunities for higher professional education and specific job skills.

MC_D	Deprived (MC_D=1) if MC is 2	Article 22: functioning of benefiting social security. Article 25: functioning of the right to medical care for a healthy life.	

It is important to note that we did not integrate the variables WATI_D (i.e., water supply interruption longer than 2 days) and ELCCUT (i.e., electricity cutoff) because: (i) concerning WATI_D, not every household has water supply by municipality and the related question was asked only to those receive water from municipality; (ii) concerning ELCCUT, the related question was not asked to households having a pre-paid card system. Thus, we cannot know whether the electricity was cut or not for those having a pre-paid card system. Instead, we directly employed the main electricity indicator (ELC) which is available for every household (having electricity, no electricity, or free basic electricity). However, we will use these two variables later for another purpose. Regarding the overall deprivation index 'DEP' we built based on 33 variables mentioned above, 2.18 per cent of the target population (18+) experience no deprivation regarding all items whilst 0.01 percent of the population suffer deprivation regarding 32 items. The most frequent group is those experience deprivations regarding 10 items, which represents 5.83 per cent of the population. From 5 to 13 deprivations, the percentage varies between 5.08 and 5.37. The summary statistics of DEP as follows:

```
SA2011M1$DEP
      n  missing distinct     Info     Mean      Gmd      .05      .10      .25      .50      .75      .90      .95
  47117        0       33    0.998    10.51    7.183        1        2        6       10       15       19       22

lowest:  0  1  2  3  4, highest: 28 29 30 31 32
```

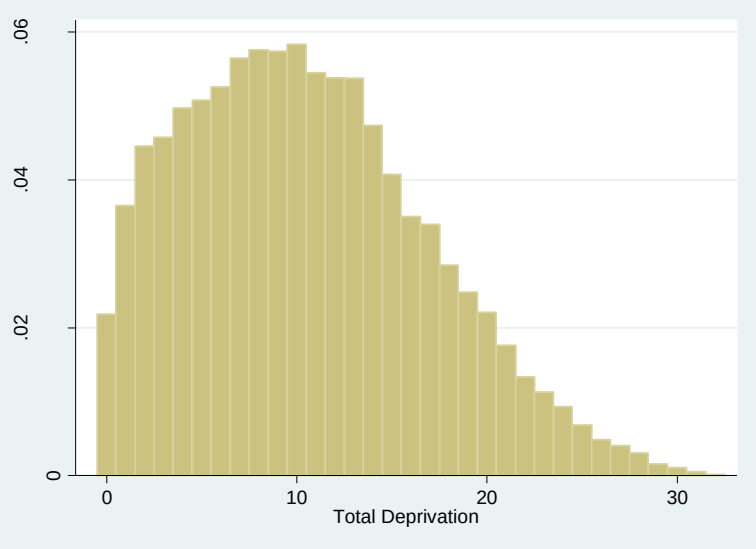


Figure 10. Histogram of the overall deprivation index based on 33 variables, data: South Africa GHS 2011

According to the quantiles of DEP, we created the deprivation variable called DEP_GR composed of 8 different poverty groups: (1: "DEP_GR==1") 0-1 deprivations, (2: "DEP_GR=="2") 2-6 deprivations, (3: "DEP_GR==3") 7-10 deprivations, (4: "DEP_GR=="4") 11-15 deprivations, (5: "DEP_GR==5") 16-19 deprivations, (6: "DEP_GR==6") 20-22 deprivations, (7: "DEP_GR==7") 23-27 deprivations, (8: "DEP_GR==8") 28-32 deprivations. First group represents the most prosperous situation, and in opposite the last group corresponds to the most deprived/poorest situation. The distribution of DEP_GR is:

SA2011M1$DEP_GR

n	missing	distinct
47117	0	8

Value	1	2	3	4	5	6	7	8
Frequency	2750	11470	10831	11787	5766	2501	1715	297
Proportion	0.058	0.243	0.230	0.250	0.122	0.053	0.036	0.006

When we take into account simultaneously 33 variables to find out the capabilities profiles of 47,117 individuals in the data (the variable named PATH33), the results show that there are 17,509 different patterns-which is a substantially high amount. Diversity of profiles is such voluminous that the most frequent five profiles correspond to only 3.97 per cent of the population. When we select the patterns having more than 100 frequencies, we only reach 7.49 per cent of the data (by data, we mean 'the 18+ population' which is the target group of the analysis).

PATH33=DWE_D+ROF_D+WAL_D+ROFC_D+WALC_D+RO_D+KIT_D+SAN_D+WATD_D+ELC_D+ENG_C_D+ENG_H_D+ENG_L_D+RUB_D+LIT_D+WP_D+AP_D+LD_D+NOI_D+POST_D+TEL_D+INT_D+TV_D+RAD_D+REF_D+ADN_D+CHN_D+NU1_D+NU2_D+NU3_D+NU4_D+EDU_D+MC_D

groups PATH33, select (5) order(h)

PATH33	Freq.	Percent	%<=
00000000000000000000000000000000	1029	2.18	2.18
00000000000000000000000000000001	264	0.56	2.74
01000000000000000000000000000000	219	0.46	3.21
00000000100000000000000000000000	193	0.41	3.62
00000000000000000000100000000000	165	0.35	3.97

groups PATH33, select(freq>=100) order(h)

PATH33	Freq.	Percent	%<=
00000000000000000000000000000000	1029	2.18	2.18
00000000000000000000000000000001	264	0.56	2.74
01000000000000000000000000000000	219	0.46	3.21
00000000100000000000000000000000	193	0.41	3.62
00000000000000000000100000000000	165	0.35	3.97
00000000001000000000000000000000	157	0.33	4.30
00000000000000000000000000000010	151	0.32	4.62
00000100000000000000000000000000	151	0.32	4.94
00000000000000000000000000000011	142	0.30	5.24
00000100000000000000000000000001	142	0.30	5.55
01000000000000000000100000000011	142	0.30	5.85
01000100000100000000100000000011	115	0.24	6.09
01000100000000000000100000000011	114	0.24	6.33
01000101101101000001010000000011	114	0.24	6.58

010000000000000000000000000000001	112	0.24	6.81
010001000000000000000000000000011	110	0.23	7.05
010001011001000000000100000000011	105	0.22	7.27
010111011001000000000100000000011	103	0.22	7.49

For instance, 2.18 per cent of the population experience no deprivation regarding all items (the most prosperous), 0.56 per cent experience deprivation only regarding medical coverage, 0.46 per cent experience deprivation only regarding the quality of dwelling roof material, 0.41 per cent experience deprivation only regarding access to main source of water, 0.30 percent suffer deprivations regarding the number of room per person in the household and medical coverage, 0.23 per cent experience deprivation regarding the quality of dwelling roof material, the number of room per person in the household, educational attainment and medical coverage, and so on.

If we look at inside of each poverty group, it is still complicated to have a complete picture of existing capabilities patterns, except the first group of poverty (PATH33_GR1) which contains 26 distinct profiles among 2,750 persons. With the most frequent 11 profiles, we reach 90.80 per cent of the population.

PATH33_GR1	Freq.	Percent	%<=
000000000000000000000000000000000	1029	37.42	37.42
000000000000000000000000000000001	264	9.60	47.02
010000000000000000000000000000000	219	7.96	54.98
000000001000000000000000000000000	193	7.02	62.00
000000000000000000000100000000000	165	6.00	68.00
000000000001000000000000000000000	157	5.71	73.71
000000000000000000000000000000010	151	5.49	79.20
000001000000000000000000000000000	151	5.49	84.69
000000000000001000000000000000000	57	2.07	86.76
000000000000010000000000000000000	56	2.04	88.80
000000000000000001000000000000000	55	2.00	90.80

These top 11 profiles correspond to (from highest to lowest in terms of frequency): (1) no deprivation regarding all items; (2) not having medical coverage; (3) living in a dwelling lacking good/adequate roof material; (4) not having access to main source of drinking water; (5) not having access to internet; (6) living in a dwelling lacking energy source for heating; (7) not having sufficient educational attainment; (8) living in a household where the number of room per person is inferior than 1; (9) experiencing the bad littering situation in the neighborhood of the dwelling; (10) living in a dwelling lacking adequate rubbish removal system; and (11) experiencing land degradation. These 11 topics are the most frequent situations among the members of the most prosperous group in the data.

From second group, complexity of patterns starts clearly raising. There are 2,994 distinct profiles among 11,470 individuals suffering deprivation regarding between two and six items (PATH33_GR2). With the most frequent 27 profiles, there is only information for 20.44 per cent of the population. The most frequent 7 patterns reflect only 7.65 per cent of the population, and we can see that from 5th most frequent pattern the percentage of frequency decreases under 1 per cent (not cumulated frequency!).

PATH33_GR2	Freq.	Percent	%<=
00000000000000000000000000000011	142	1.24	1.24
00000100000000000000000000000001	142	1.24	2.48
01000000000000000000100000000011	142	1.24	3.71
01000100000100000000100000000011	115	1.00	4.72
01000100000000000000100000000011	114	0.99	5.71
01000000000000000000000000000001	112	0.98	6.69
01000100000000000000000000000011	110	0.96	7.65
01000100000000000000000000000001	99	0.86	8.51
00000000000000000000100000000011	95	0.83	9.34
00000100000000000000000000000011	90	0.78	10.12

The most frequented profiles of the top 10 per cent of the second poverty group are: (1) not having sufficient educational attainment and not having medical coverage; (2) living in a household where the number of room per person is inferior than 1 and not having medical coverage; (3) living in a dwelling lacking good/adequate roof material, not having access to internet, not having sufficient educational attainment and not having medical coverage; (4) living in a dwelling lacking good/adequate roof material, living in a household where the number of room per person is inferior than 1, living in a dwelling lacking energy source for heating, not having access to internet, not having sufficient educational attainment and not having medical coverage; (5) living in a dwelling lacking good/adequate roof material, living in a household where the number of room per person is inferior than 1, not having access to internet, not having sufficient educational attainment and not having medical coverage; (6) living in a dwelling lacking good/adequate roof material and not having medical coverage; (7) living in a dwelling lacking good/adequate roof material, living in a household where the number of rooms per person is inferior than 1, not having sufficient educational attainment and not having medical coverage; (8) living in a dwelling lacking good/adequate roof material, living in a household where the number of room per person is inferior than 1 and not having medical coverage; (9) not having access to internet, not having sufficient educational attainment and not having medical coverage; and (10) living in a household where the number of room per person is inferior than 1, not having sufficient educational attainment and not having medical coverage.

Inside the third group of poverty (PATH33_GR3) composed of 10,831 persons, there are 4,480 unique patterns. From the most frequent pattern, the frequency of each single pattern in the data remains under 1 per cent.

PATH33_GR3	Freq.	Percent	%<=
01000101100100000000100000000011	105	0.97	0.97
01011101100100000000100000000011	103	0.95	1.92
01011100000100000000100000000011	86	0.79	2.71
01000101100101000001010000000011	83	0.77	3.48
01000101100000000000100000000011	81	0.75	4.23
01000001101101000001010000000011	77	0.71	4.94
01000001100001000001010000000011	64	0.59	5.53
01000101100001000001010000000011	51	0.47	6.00

Regarding the fourth group of poverty (PATH33_GR4) composed of 11,787 individuals, 5,347 unique patterns emerge. The most frequent three patterns concern only 1.94 per cent of the population.

PATH33_GR4	Freq.	Percent	%<=
01000101101101000001010000000011	114	0.97	0.97
01011101101101000001010000000011	65	0.55	1.52
11100101100101000001010000000011	50	0.42	1.94

The fifth group of poverty (PATH33_GR5) includes 5,766 individuals manifesting in 2,640 distinct profiles. The profile, 11100101111111000001011010000011, is the most frequent one with only 37 frequencies, which is equivalent of 0.64 per cent of the population. In sixth group (PATH33_GR6), there are 2,501 individuals with 1,133 different profiles. The most frequent profile, 11111101111111000101011010000011, has only 32 observations-equivalent of 1.28 per cent of the population. In seventh group (PATH33_GR7), there are 763 distinct profiles among 1,715 individuals. The most frequent three profiles cover 4.08 per cent of the population.

PATH33_GR7	Freq.	Percent	%<=
11111101111111111101011010000011	25	1.46	1.46
11111101111111000001011111111111	23	1.34	2.80
11111101111111000101011111111111	22	1.28	4.08

And the last group (PATH33_GR8) comprises of 297 individuals with 126 distinct profiles. The first three profiles correspond to 12.12 per cent of the population. These 3 profiles are: (1) not experiencing the bad littering situation in the neighborhood of the dwelling, not experiencing water pollution, not experiencing air pollution, not experiencing noise pollution, having access to telephone, but experiencing deprivation regarding other items; (2) having kitchen in the dwelling, not experiencing water pollution, not experiencing air pollution, not experiencing land degradation, having access to telephone, but experiencing deprivation regarding other items; (3) having kitchen in the dwelling, not experiencing water pollution, not experiencing land degradation, having access to telephone, but experiencing deprivation regarding other items.

PATH33_GR8	Freq.	Percent	%<=
11111111111111000101011111111111	14	4.71	4.71
11111101111111100101011111111111	11	3.70	8.42
11111101111111110101011111111111	11	3.70	12.12

From the statistics shown in the table below, we can see that diversity of profiles is significantly less for the first group but also it is relatively less for the second group notably. And the ratio of profiles remains around 0.4 for the rest, which shows a greater complexity.

Table 29. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GR and 33 variables), data: South Africa GHS 2011

Group	Unique capabilities profiles	Number of persons in the group	Ratio (Unique profiles/Number of persons in the group)
DEP_GR1	26	2750	0.009
DEP_GR2	2994	11470	0.261
DEP_GR3	4480	10831	0.414
DEP_GR4	5347	11787	0.454
DEP_GR5	2640	5766	0.458

DEP_GR6	1133	2501	0.453
DEP_GR7	763	1715	0.445
DEP_GR8	126	297	0.424
Total (whole sample)	**17509**	**47117**	**0.372**

If we look at capabilities profiles per poverty group based on the essential 49 variables mentioned in the beginning of the section, the plurality of distinct profiles increases even more. The ratio of unique profiles remains between 0.6 and 0.75 through 8 poverty groups.

Table 30. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GR and 49 variables), data: South Africa GHS 2011

Group	Unique capabilities profiles	Number of persons in the group	Ratio (Unique profiles/Number of persons in the group)
DEP_GR1	1699	2750	0.617
DEP_GR2	8220	11470	0.716
DEP_GR3	8318	10831	0.767
DEP_GR4	9368	11787	0.794
DEP_GR5	4594	5766	0.796
DEP_GR6	1970	2501	0.787
DEP_GR7	1368	1715	0.797
DEP_GR8	224	297	0.754
Total (whole sample)	**35761**	**47117**	**0.758**

For instance, for the first group of poverty, the most frequent capabilities profile is '0109015501121015111133611121222221312212211111622221 01' representing only 0.76 per cent of the group (living in a dwelling/house or brick/concrete block structure on a separate stand or yard or on farm & roof material is tile & wall material is brick & very good roof and wall conditions, the ratio of rooms per person in the dwelling is higher than 1, there is kitchen, etc.). In second poverty group, the per centage of each of the most frequent three patterns correspond to 0.08 per cent of the population. This number is substantially low and even other per centages for other profiles are less than this figure.

It is important to emphasize that for an accurate classification analysis; it is necessary to have a balanced data. In our case of 8 different poverty groups, the data emerges quite unbalanced especially for the 8th group but also for 1st, 5th, 6th, and 7th groups. Due to this fact, we created another variable called DEP_GRS regrouping the population by 20th, 40th, 60th, and 80th percentiles. The table below shows the number of persons in the group, unique capabilities profiles based on the 33 variables and the ratio of unique profiles by the number of persons in the group. We also note that we removed people with no deprivation (DEP==0) as they form

the most prosperous group in the data. It will be more convenient to examine separately this particular group which represents 2 per cent of the data.

Table 31. Ratio of unique profiles to the number of persons in the poverty group (based on the variable DEP_GRS and 33 variables), data: South Africa GHS 2011

Group	Unique capabilities profiles	Number of persons in the group	Ratio (Unique profiles/Number of persons in the group)
DEP_GRS1 (1-4 deprivations)	1213	8321	0.145
DEP_GRS2 (5-7 deprivations)	2886	7533	0.383
DEP_GRS3 (8-11 deprivations)	4536	10735	0.422
DEP_GRS4 (12-15 deprivations)	4211	9220	0.456
DEP_GRS5 (16-32 deprivations)	4662	10279	0.434
Total (whole sample)	**17508**	**46088**	**0.379**

As we can see from the data there is clearly a large variety of living conditions people face in their life. For instance, within second group, the most frequent 6 patterns represent only 6.88 per cent of the group, or within fourth group the most frequent 2 patterns reach only 1.25 per cent of the group. The analysis of capabilities, therefore, becomes quite complicated due to this plurality. In next paragraphs, we will show how to deal with such complexity by employing classification trees method. The goal is to find out the most informative capabilities patterns distinguishing the 5 poverty groups. More specifically, we will do in total 10 classification tree analysis predicting how the poverty groups distinguish by 33 variables: {DEP_GRS1 vs. DEP_GRS2, DEP_GRS1 vs. DEP_GRS3, DEP_GRS1 vs. DEP_GRS4, DEP_GRS1 vs. DEP_GRS5, DEP_GRS2 vs. DEP_GRS3, DEP_GRS2 vs. DEP_GRS4, DEP_GRS2 vs. DEP_GRS5, DEP_GRS3 vs. DEP_GRS4, DEP_GRS3 vs. DEP_GRS5, DEP_GRS4 vs. DEP_GRS5}.

Table 32. Per centage of groups in the samples of target groups, data: South Africa GHS 2011

Target group in classification analysis	Per centage of groups in the sample
DEP_GRS1 vs. DEP_GRS2	{52.48, 47.52)
DEP_GRS1 vs. DEP_GRS3	{43.66, 56.34}
DEP_GRS1 vs. DEP_GRS4	{47.43, 52.57}

DEP_GRS1 vs. DEP_GRS5	{44.73, 55.27}
DEP_GRS2 vs. DEP_GRS3	{41.23, 58.77}
DEP_GRS2 vs. DEP_GRS4	{44.96, 55.04}
DEP_GRS2 vs. DEP_GRS5	{42.29, 57.71}
DEP_GRS3 vs. DEP_GRS4	{53.79, 46.21}
DEP_GRS3 vs. DEP_GRS5	{51.08, 48.92}
DEP_GRS4 vs. DEP_GRS5	{47.28, 52.72}

As the table above shows we have a quite balanced data but still we don't have a 100 per cent equal distribution between two groups (e.g., DEP_GRS1: 50 per cent vs. DEP_GRS2: 50 per cent).

We note that in the following analysis concerning classification trees by applying the package rpart in R, we will give only the essential statistical outputs.

CAP111[471]: GR1 vs. GR2 analysis

We start off the analysis with the 1st target group DEP_GRS1 vs. DEP_GRS2 (for facilitating writing, we will employ GR1 vs. GR2 further). We use the data containing only people from the poverty group 1 and the poverty group 2. We recall that we will use the rpart package in R environment. The outcome (called *Endpoint* also in the outputs of the package rpart) is a dichotomous variable, namely CAP111, taking the value 1 for the group 1 and the value 2 for the group 2 (so that people identified in the group 1 take the value 1 and people identified in the group 2 take the value 2 according to CAP111).

The functions in the rpart library/package produce binary classification trees. According to the functions of the rpart library, when the outcome is categorical as in our case, the "class" method should be used. In addition, for a continuous variable the method "anova", for a poisson process/count the method "poisson", for a survival process the method "exp" are available in this package (see Atkinson and Therneau 2000, 3). Concerning regressors, they can be both continuous and categorical (factor) variables. To make clear how the rpart produces binary classification trees, we will provide all necessary theoretical explanations as well as the R codes in following paragraphs for the CAP111 analysis. For further examinations (including other country analysis), theoretical information and R codes won't be given again. It is important to note that we already gave the theoretical insights of the decision, and more specifically classification tress in the method section (see 4.1). Here, the goal is to explain the functions of the rpart package we will use in the analysis. Our explanations will be mainly based on the official rpart documents written by Atkinson and Therneau from Mayo Foundation, which are the three technical reports called "An introduction to Recursive Partitioning Using the RPART Routines" and published in 1997, 2000, and 2015. Moreover, some other sources will be also cited for complementary information.

The language of the rpart is the Standard Splus modeling. The typical coding of fitting the model is (the coding is done in the script of R):

[471] The meaning of CAP111 is: Capabilities analysis (**CAP**) of 1st step (**1**) for 1st target group (**1**).

```
model=rpart(outcome code~regressor 1 + regressor 2 +…regressor k
codes, data= data name, method=method name)
```

where 'model' is the name given by the researcher to model, 'outcome code' is the name of the outcome variable mentioned in the data (e.g., CAP111), 'regressor code' is the name of the regressor existing in the data, '*data name*' is the name of the data determined by the researcher, '*method name*' is the method will be used in rpart such as "class" or "anova" or "poisson" or "exp". Other terms 'rpart', 'data, 'method' are the names defined in the package which should be kept in that way in the coding. The sign ' +' is used in case of more than one regressors. If the data contains only the outcome and the all regressors, the sign ' . ' can be directly employed for simplifying the coding (outcome code ~ .). For fitting the model, the package offers a range of options. We will mention the options we will use and provide the information written by Atkinson and Therneau in their article of 2000 (4-8).

1) parms: a list of some specific parameters available according to the features of the method. For classification trees, there are three parameters: 'prior', 'loss', and 'split'.

- prior: the vector of prior probabilities which should be positive and sum to 1. The researcher can determine in priori the probabilities that the categories of the outcome would be taking in the analysis (e.g., the researcher may set in priori that the two categories of the outcome in the model, X and Y, should have the probabilities 0.4 and 0.6 respectively at each node). The default is the prior probabilities that are proportional to the observed data frequencies.
- loss: the loss matrix with zeros on the diagonal and positive off the diagonal. With this parameter, the researcher may penalize the misclassification of the categories such as weighting two times more the X observations classified as Y in prediction. The default is all misclassifications are considered having equal costs.

Suppose that the outcome (WH) has two categories coded 1 and 2. 1 is used for working long hours and 2 is used for the opposite. The confusion matrix as a result of prediction is:

Table 33. Confusion matrix for showing the utility of the loss parameter in rpart

	Predicted values	
Actual values	WH=1	WH=2
WH=1	True positive (TP)	False positive (FP)
WH=2	False negative (FN)	True negative (TN)

True positive and true negative mean that both WH=1 and WH=2 are correctly classified in prediction, and this is why there is no reason to penalize these cases and in the loss matrix the diagonal where TP and TN are placed should take the value 0. On the other hand: (i) FN means that some people not working long hours in the actual data are however predicted as working long hours; (ii) FP means that some people working long hours in the actual data are however predicted as not working long hours. The researcher may prefer to weight more the cost of a false positive regarding her framework. Suppose she wants to cost three times higher FPs, so that she can define a loss matrix in her fitting model as follows: (0, 3, 1,0). The coding of the loss matrix in the rpart is as follows: loss=matrix(c(0,x,y,0), nrow=2, ncol=2, byrow=F) where x is the cost of misclassifying a positive example as a negative and y is the

cost of misclassifying a negative example as a positive (see also evaluation measures and confusion matrix in section 4.1).

- `split:` the splitting criteria can be found based on the gini index or information gain. The first case is called 'gini' and the second case is called 'information' (or known as deviance) in the rpart. The coding is as follows: `split= "gini"` or `split= "information"` in the rpart. The default is gini. Maindonald and Braun (2010, 359) mention that "for the two-class problem, the Gini index and the deviance will almost always choose the same split as the deviance or entropy".

By taking in consideration the parms, a classification fitting model can be written in R as follows:

```
lossm=matrix(c(0,3,1,0), nrow=2, ncol=2, byrow=F)

fit=rpart(WH ~ age + gender + region + sector + education, data=labour,
method= "class", parms=list(split= "gini", loss=lossm))
```

where age, gender, region, sector, education are the regressors in the model, data is called *labour*, and the fitting model is called *fit* in the coding above.

`2) na.action:` the action for missing values. The default action in rpart is `na.rpart` which removes the rows for which either the outcome or all of the predictors are missing in the data.

`3) control:` this is a list of control parameters. For integrating the control parameters selected by the researcher, the `rpart.control` function should be used in the coding. The elements of the list are (definitions below are taken from Atkinson and Therneau (2000, 5-6)):

- `minsplit:` the minimum number of observations in a node for which the routine will even try to compute a split. The default is 20.
- `minbucket:` the minimum number of observations in a terminal node. This defaults to minsplit/3.
- `maxcompete:` this parameter controls the number that will be printed. The default is 5.
- `xval:` the number of cross-validations to be done. The default is 10.

Cross-validation process includes splitting the data into *k* subsets. In rpart, as we mentioned above the default value is 10, in other words the data is normally split into 10 subsets except if the user sets for the `xval` a value different than 10. Suppose that the data is randomly split into X_1, X_2,…,X_{10}. Initially, the first partition X_1 is left out, the model is fitted based on the remaining data which includes 9 subsets from X_2 to X_{10}, then the classification rules obtained from the fitting are applied to the subset that has been left out, thus we obtain the predicted results for the data that has been left out. In second turn, the second partition X_2 is left out, the model is fitted based on the remaining data which includes 9 subsets X_3, X_4, X_5,X_6,X_7,X_8,X_9, X_1, and then the classification rules from this fitting are used to predict the classification for the subset X_2 which was left out. In this mechanism, the data which is left out has in fact the function of testing the results obtained on the remaining data. For that reason, the data left out is called the *test* data and the remaining data is called the *training* data. The fact that the model is fitted through the remaining data can be seen as a training process of prediction of classification which will be tested later on the test data. In the literature, it is known that at the *k*th fold the *k*th subset is defined as the test data whilst the remaining data is considered as the

training. For instance, at the 2nd fold, the 2nd subset has the role of testing data. Cross-validation process is used for obtaining a reliable error rate for the classification model in a way that at each fold an average error rate is computed. The average error rate corresponds to the division of the total number of observations misclassified by the total number of observations. After it is done for each fold, at the end the average error is computed by summing all error rates and diving this sum by the total number of observations in the data.

Suppose that a teacher of a high school would like to discover the patterns of the students of terminal classes, 60 students in total, in terms of their preference for sport activities. In this line, the target response variable takes 3 categories representing 3 sport activities offered at school: tennis, swimming, and baseball. The teacher determines 6 regressors: gender (male vs. female), average math score (between 1 and 4), average literature score (between 1 and 4), average natural science score (between 1 and 4), average music score (between 1 and 4), average painting score (between 1 and 4), and the interest in sport (1=very interested, 2=average interest, 3=few interested, 4=not at all interested). Average scores for 5 classes are due to student's last semester achievement. After collected information for these 6 items for each student, the teacher constructs data with 6 columns and 60 rows. The model is fitting by employing 3-fold cross-validation. The data is randomly divided into 3 subsets covering each of them 20 students. At 1st fold, through the ensemble of the training data composed of the 2nd and 3rd subsets (in total 40 students), it is found 4 patterns in the tree (with 4 terminal nodes, gender variable at the root node, and natural science score as the second variable appearing at the 2nd level of the tree) : female students with a natural science score higher than 3.5 prefer tennis, female students with a natural science score lower than 3.5 prefer swimming, male students with a natural science score lower than 3 prefer baseball, and male students with a natural science score higher than 3 prefer tennis. Then, these 4 classification rules are applied on the test set which is 1st subset (20 students). The actual target response variable's values for each student in the test set are compared with the predicted results if these 4 rules are applied. It is found 3 persons misclassified regarding first pattern, 5 persons misclassified regarding second pattern, 0 misclassification regarding third pattern and 2 persons misclassified regarding fourth pattern. For instance, 2 male students with a natural science score higher than 3 has said to prefer swimming whereas they are predicted as preferring tennis regarding the fourth pattern extracted from the training data classification. Thus, the 1st fold yields a misclassification rate of: (3+5+0+2)/60 = 0.166 or 16.6 per cent. The 2nd fold (1st and 3rd subsets form the training data where classification rules are predicted, and 2nd subset is used as the test data for testing the predicted results) produces 14 of 40 students misclassified. The 3rd fold (1st and 2nd subsets form the training data and the 3rd subset is the test data) generates 8 of 40 students misclassified. To obtain an average rate of cross-validated error rate through 3 folds, the sum of the three folds' misclassification scores (10+14+8=32) is divided by the total number of students in the entire data (60) arriving at 53.3 per cent cross-validated error rate. The package rpart produces a cross-validated error rate per the size of tree (or in other words, by the number of terminal nodes). With a larger size of tree, we will have a lower cross-validated error rate.

- `maxsurrogate`: the maximum number of surrogate variables to retain at each node. The default is 5. Surrogate variables are used for missing values based on the question "which other splits would classify the same subjects in the same way".
- `usesurrogate`: If the value is 0, then a subject who is missing the primary split variable does not progress further down the tree. If the value is 1, then all surrogate

variables except "go with the majority" are used to send a case further down the tree. The default is 2, which splits subjects according to the way described in `maxsurrogate` above.

- `cp`: the threshold complexity parameter. Greater cp means small size of tree with a higher cross-validated error rate. On the one hand, at maximum level of tree (the most saturated) the cp will be equal to 0; on the other hand, a value of cp=1 provides a tree with no splits. The default is 0.01. It is recommended to set a smaller cp such as the default value in rpart to see at which level of the tree the cross-validation error rates reaches a minimum. After that, the saturated tree can be pruned by fitting the model with a higher level of cp to obtain a more optimal interpretable tree (see the application below for pruning phase and the selection of cp).

In CAP111 analysis (1st poverty group vs. 2nd poverty group), there is in total 15,854 individuals and 33 deprivation variables. We will divide the data into training and test samples[472]. The decision tree and thereby the classification rules will be built on the training sample (which means that the classes will be predicted in the training sample based on the cross-validation process). Then, the discovered classification rules will be applied in prediction the test data. Comparison between the actual and predicted classes in the test data will inform about the accuracy of the classification model. The training sample is set 70 per cent and the test sample is set 30 per cent of the initial complete data. For that, we employed 3 different ways/formulas (see commands in the rpart coding box below). The first one gives 5,812 persons of GR1 and 5,285 persons of GR2 in the training sample vs. 2,509 persons of GR1 and 2,248 persons of GR2 in the test sample. The second gives 5,547 persons of GR1 and 5,022 persons of GR2 in the training sample vs. 2,774 persons of GR1 and 2,511 persons of GR2 in the test sample. The third gives 5,864 persons of GR1 and 5,250 persons of GR2 vs. 2,457 persons of GR1 and 2,283 persons of GR2. According to the initial data composed of 15,854 persons there are 8,312 persons of GR1 and 7,533 persons of GR2, which leads to the ratio of GR1 to GR2: 1.103. The first method yields a ratio of 1.099 for the training sample and 1.116 for the test sample. The second method yields a ratio of 1.104 for both the training and test samples. The third method yields a ratio of 1.116 for the training sample and 1.07 for the test sample. Based on that, we select the second method which produces a more consistent result. However, with such sample results from the second method, the training sample equals to 67 per cent and the test sample equals to 33 per cent of the data (we see that the proportion of the samples is not {70%, 30%) but the difference is not substantial as well)[473].

Based on the second technique of sampling mentioned in previous paragraph, we have in total:

Table 34. Training vs. test samples in CAP111 analysis, data: South Africa GHS 2011

	GR1	GR2
Training sample	5547	5022
Test sample	2774	2511

[472] We note again that we use the terms 'sample', 'subset', and 'data' interchangeably in case of training vs. test sets.

[473] With the first method, the training sample equals to 70 per cent and the test sample equals to 30 per cent of the data. With the second method, the training sample equals to 70.10 per cent and the test sample equals to 29.90 per cent of the data.

We start off the analysis with a first attempt of classification trees by integrating a loss matrix and specific values for the parameters of minbucket, minsplit and cp in the equation of the fitting the model. The R codes are as follows:

```
#SA GHS 2011 Data Analysis for PhD thesis
#Author: A. Tugce Beycan (tugce.beycan@unine.ch), Department of Sociology,
University of Neuchatel
#Updated: 17 September 2019


library(foreign)

SA2011CAP111=read.dta("E:\\PhD_Research\\Datas_May2017\\SA_GHS_2011\\Stata\\CAP111_
SA_GHS2011_11Sept2019_TB.dta")

#Data concerns only people from poverty group 1 and poverty group 2 and there are
no missing values.

pkgs=c("rpart","rpart.plot","party","random Forest","e1071")

install.packages(pkgs,depend=TRUE)

dim(SA2011CAP111)

attributes(SA2011CAP111)

SA2011CAP111[1:5,]#first five observations in the data

df=SA2011CAP111[-1]

df$CAP111=factor(df$CAP111, levels=c(1,2),labels=c("GR1","GR2"))

#the response variable is changed to a factor for classification analysis.

set.seed(1234)

train=sample(nrow(df),0.7*nrow(df))

#The data is randomly divided into a training sample (70%) and a test sample (30%).

df.train=df[train,]

df.test=df[-train,]

table(df.train$CAP111)

# category 1: 5812 and category 2: 5285 in the training sample

table(df.test$CAP111)

#category 1: 2509 and category 2: 2248 in the test sample

#Training sample is for creating classification rules by employing decision trees
method and the test sample is to evaluate the effectiveness of these obtained
rules.

#The outcome is the variable CAP111 which is already a factor variable, 1=poverty
group 1 vs. 2=poverty group 2

#Regressors are the 33 deprivation variables, which of them is a binary factor
variable.

#Kabacoff (2015, 393) defines decision trees as "they involve creating a set of
binary splits on the predictor variables in order to create a tree that can be used
to classify new observations into one of two groups.

#Or another way of creating training vs. test samples

library(foreign)
```

```
SA2011CAP111=read.dta("E:\\PhD_Research\\Datas_May2017\\SA_GHS_2011\\Stata\\CAP111_
SA_GHS2011_11Sept2019_TB.dta")

SA2011CAP111$CAP111=factor(SA2011CAP111$CAP111,
levels=c(1,2),labels=c("GR1","GR2"))

set.seed(1234)

GR1.sub=c(1:nrow(SA2011CAP111))[SA2011CAP111$CAP111=="GR1"]

GR2.sub=c(1:nrow(SA2011CAP111))[SA2011CAP111$CAP111=="GR2"]

#2/3 for training, 1/3 for test

train.GR1=sample(GR1.sub,floor(length(GR1.sub)*2/3))

train.GR2=sample(GR2.sub,floor(length(GR2.sub)*2/3))

train=c(train.GR1,train.GR2)

train.set=SA2011CAP111[train,]

table(train.set$CAP111)

#gr1: 5547 gr2: 5022

test.set=SA2011CAP111[-train,]

table(test.set$CAP111)

#gr1: 2774 gr2: 2511

#or another way:

library(foreign)

SA2011CAP111=read.dta("E:\\PhD_Research\\Datas_May2017\\SA_GHS_2011\\Stata\\CAP111_
SA_GHS2011_11Sept2019_TB.dta")

SA2011CAP111$CAP111=factor(SA2011CAP111$CAP111,
levels=c(1,2),labels=c("GR1","GR2"))

set.seed(1234)

part=sample(2,nrow(SA2011CAP111),replace=TRUE,prob=c(0.7,0.3))

CAP111.train=SA2011CAP111[part==1,]

table(CAP111.train$CAP111)

#gr1:5864 gr2: 5250

CAP111.test=SA2011CAP111[part==2,]

table(CAP111.test$CAP111)

#gr1: 2457 gr2: 2283

#The data has 8312 for GR1 and 7533 for GR2 initially. According to that, the ratio
of (GR1|GR2) is 1,103.

#The 2nd method of sampling gives the most near number to 1.103, with both
train(GR1/GR2) and test(GR1/GR2) is equal to 1.104.

#We will employ the train.set and test.set for following classification analysis.

############################Using Rpart package for classification
trees#####################################

library(rpart)#Recursive PARTitioning (Therneau & Atkinson, Mayo Clinic) attach the
library for the functions
```

```
#We've already downloaded the package. Or, use require (rpart). We attach the
library for having the necessary functions of the package.

set.seed(123456)#set the seed value to make sure that through the cross-validation
process we obtain the same trees.

#parms: optional parameters. prior (the vector of prior probabilities), loss (loss
matrix), split (splitting index)

lossm=matrix(c(0,3,4,0))#penalizes more being actual GR2 and classified as GR1.

temp=rpart.control(xval=10, minbucket=2, minsplit=4, cp=0)

#Fitting the model with the parameters:

dtreeCAP111=rpart(CAP111~.,method="class",data=train.set,control=temp,
parms=list(loss=lossm,split="gini"))

#Fitting the model without the parameters, by mentioning only Gini index:

dtreeCAP111S=rpart(CAP111~.,method="class",data=train.set,parms=list(split="gini"))

print(dtreeCAP111S)#print a text version of the tree.

summary(dtreeCAP111S)#print a summary which examines each node in depth.

plot(dtreeCAP111S, uniform=T, branch=0.1,main="CAP111S saturated tree")

text(dtreeCAP111S, pretty=1, use.n=T, cex=0.6) #Plot the standard version of the
tree

dtreeCAP111S$cptable#complexity parameter (cp) penalizes larger trees

#For obtaining a pruned tree, we look at the cp table which shows the prediction
error for different tree sizes.

plotcp(dtreeCAP111S) #This graph illustrates the cross-validated error rate against
the complexity parameter.

#Plot the errors

opar=par(mfrow=c(1,2))

with(dtreeCAP111S,plot(cptable[,3],xlab="Tree Number", ylab="Resubstitution Error
(R)", type="b"))

with(dtreeCAP111S,plot(cptable[,4],xlab="Tree Number", ylab="Cross-Validated Error
(R(cv))", type="b"))

par(opar)

par(pty="s")

plotcp(dtreeCAP111S)

with(dtreeCAP111S,{

  lines(cptable[,2]+1,cptable[,3],type="b",col="red")

  legend(locator(1),c("Resub. Error", "CV Error", "min(CV Error+1SE"),

         lty=c(1,1,2),col=c("red","black","black"),bty="n")

})
```

The table below provides information on complexity parameter, number of splits, misclassification error rate (resubstitution estimate), cross validated (or cross-validation) error rate, and standard error of the cross- validated error rate. As a result of fitting the model, the saturated tree emerges with 636 splits. As it is very long the information for 636 splits, we just

provide the information for the first 22 splits. The "rel error" (resubstitution error rate) designates the error rate for a tree of a given size in the training sample. The cross validated error (xerror) is based on the 10-fold cross validation by using the training sample. And the final column "xstd" is the standard error of the cross-validation error. For obtaining a pruned tree, we look at the cp table which shows the prediction error for different tree sizes. We note that the tree size is defined by "nsplit" (the number of branch splits), and the largest tree has always a lowest resubstitution error rate.

Table 35. cp table- summary information on complexity parameter, split, and error rates of the saturated CAP111 tree, data: South Africa GHS 2011

	CP	nsplit	rel error	xerror	xstd
1	2.235497e-01	0	1.00000000	1.3333333	0.013630523
2	1.410461e-01	1	0.77645029	0.7414709	0.010943907
3	1.991239e-02	2	0.63540422	0.7068897	0.011332459
4	1.871764e-02	3	0.61549184	0.6705828	0.011014078
5	1.679278e-02	4	0.59677419	0.6440993	0.010814421
6	1.458029e-02	5	0.57998142	0.6262445	0.010732660
7	1.148281e-02	9	0.51008894	0.6111775	0.010918118
8	1.015532e-02	11	0.48712332	0.5823709	0.010735906
9	8.993761e-03	12	0.47696801	0.5535643	0.010485279
10	7.964954e-03	14	0.45898049	0.5185185	0.010176295
11	7.765830e-03	16	0.44305058	0.4987389	0.009956180
12	7.301208e-03	17	0.43528475	0.4859950	0.009837215
13	7.035710e-03	18	0.42798354	0.4733838	0.009712333
14	5.575468e-03	19	0.42094783	0.4663481	0.009646791
15	4.911722e-03	21	0.40979689	0.4512147	0.009536144
16	4.530068e-03	22	0.40488517	0.4455064	0.009500025

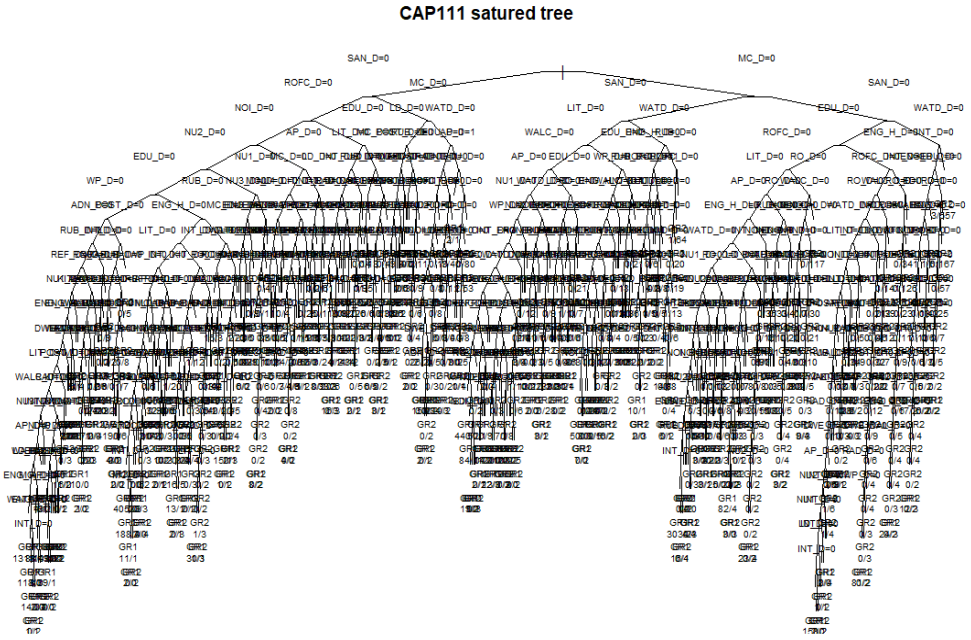


Figure 11. The saturated tree obtained by fitting the model CAP111, data: South Africa GHS 2011

From the figure above we can see that the initial saturated tree is substantially big, and it is almost impossible to examine the patterns. For that reason, we will prune the tree by focusing on the cross-validated error rate (or called cross-validated estimate of error, or cross-validation

error rate) and the standard error for that error mentioned in the cp table above. By plotting the cross-validated error rate by the cp (complexity parameter) and size of tree (the package rpart reflects the size of tree by the number of splits in the tree-see below), we can also observe how the error rate varies. For instance, with a size of tree around 30, the cp is around 0.0042 and the cross validated error rate is between 0.4 and 0.6. The cross-validated error rate reaches its minimum level when the tree becomes the largest. To define the optimal pruned tree, the best way is to select the smallest tree whose cross-validated error rate is within one standard error of the minimum cross-validated error value. For that, we focus on figure 12 which shows the evolution of the cross-validated error rate and a dotted line which is the upper limit of the one standard error rule. This plot suggests selecting the tree with the leftmost complexity parameter value below the line. In this case, the complexity parameter around 0.00042 seems the most convenient. Under this condition, we can select the tree with 208 splits (so that 209 terminal nodes) having a complexity parameter of 0.0004314, a cross-validated error rate of 0.2513607 and a standard error of 0.007432793. Nevertheless, such pruned tree will be still remaining a quite big for interpretation. As a result, we will change the parameters of the model by implementing a higher minimum number of observations in a terminal node and a higher minimum number of observations in a node. In other words, we will fit the model with the default values included in rpart. We call this model CAP111S however the target response variable is still called CAP111- just the name of the model must be changed in R coding.

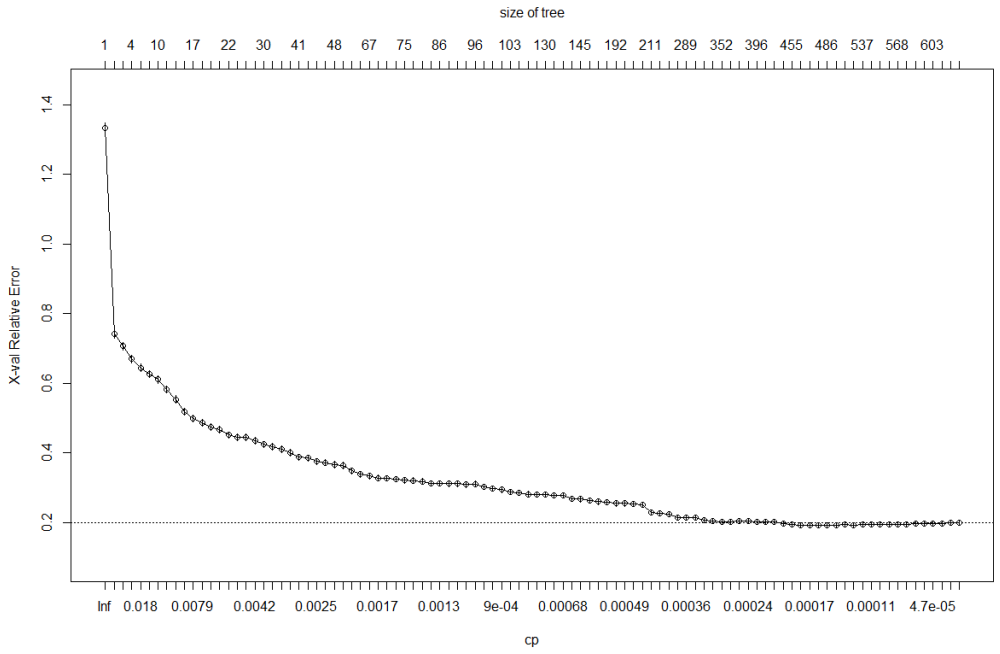


Figure 12. The CAP111 saturated tree's cross-validated error rate by the cp and size of tree, data: South Africa GHS 2011

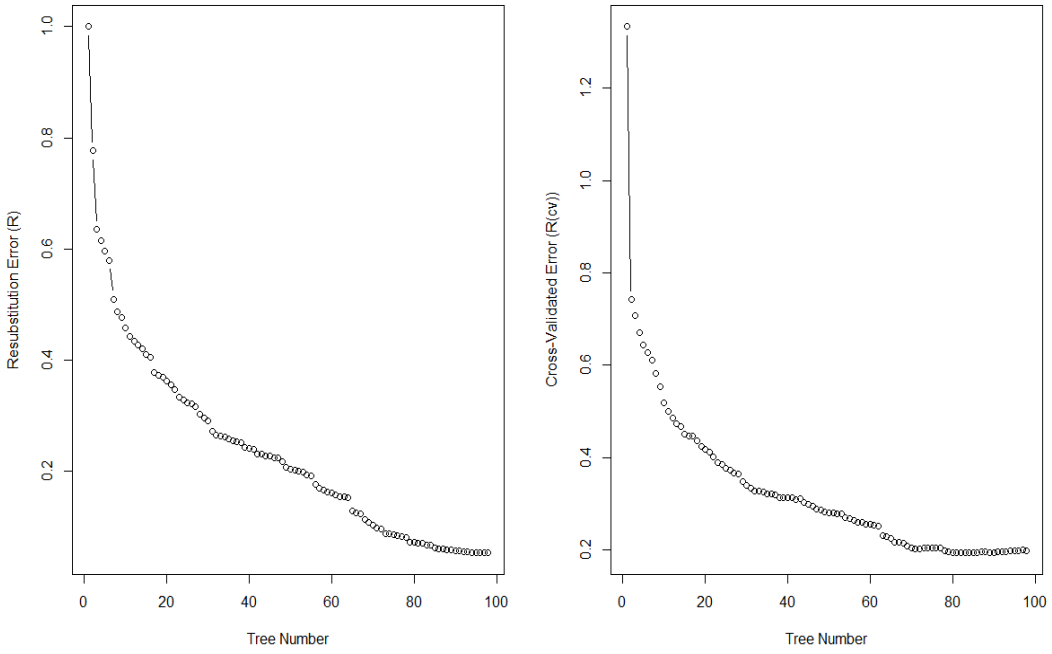


Figure 13. The CAP111 saturated tree's resubstitution and cross-validated error rates by the tree number, data: South Africa GHS 2011

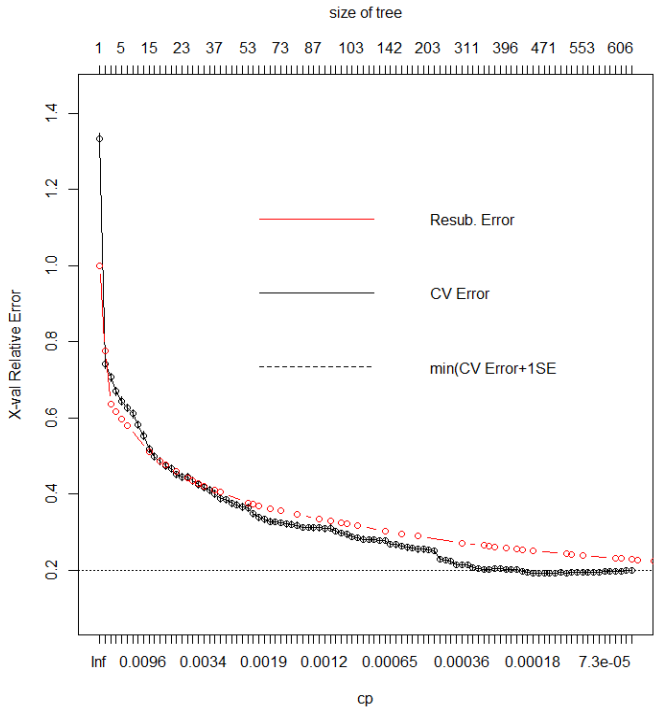


Figure 14. The CAP111 saturated tree's resubstitution, cross-validated error rates and min(CV Error + 1SE Rule), data: South Africa GHS 2011

By using the default values in the rpart (the model CAP111S), we obtain a smaller tree with the following results printed in the box below containing: (i) the formula used in fitting the model; (ii) the total number of observations at the root node (n=10569); (iii) complexity parameter (cp); number of split (nsplit), resusbtitution error rate (rel error), cross-validation error rate

(xerror) and standard error of the cross-validation error rate per split in the tree; (iv) the most important variables; (v) detailed information for each node in terms of the total number of observations with the class counts[474] for each category of the target response variable (e.g., 5,547 persons from GR1 and 5,022 persons from GR2 at the first node), the value of complexity parameter, the predicted class (GR1 or GR2), the probabilities for each class (by majority vote, the class with a higher probability defines the classification of the node), the expected loss (the information coming from the class which is not selected at the node), how the node is split into left and right branches (the number of observations going to left branch, the number of observations going to right branch, the split variables (including the split variable used at the node, the list of other potential split variables and the list of potential surrogate variables in case of missing values). The node information is given for all 25 nodes appeared at the saturated tree however the information on error rates and variable importance are provided only for the first 6 splits for making shorter the output as it is quite long due to the size of the saturated tree.

```
rpart(formula = CAP111 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"))
  n= 10569
          CP nsplit rel error    xerror        xstd
1 0.34448427      0 1.0000000 1.0000000 0.010222892
2 0.08024691      1 0.6555157 0.6555157 0.009480090
3 0.03106332      2 0.5752688 0.5752688 0.009123484
4 0.02309837      3 0.5442055 0.5591398 0.009041984
5 0.01911589      4 0.5211071 0.5396256 0.008938706
6 0.01075269      5 0.5019912 0.4932298 0.008671561
7 0.01000000      6 0.4912386 0.4830745 0.008608827
Variable importance
 ROF_D   MC_D  SAN_D ROFC_D WATD_D WALC_D  LIT_D  INT_D   RO_D   AP_D
    39     24     11      7      4      4      3      3      2      2
Node number 1: 10569 observations,    complexity param=0.3444843
  predicted class=GR1  expected loss=0.4751632  P(node) =1
    class counts:  5547  5022
   probabilities: 0.525 0.475
  left son=2 (4971 obs) right son=3 (5598 obs)
  Primary splits:
      ROF_D splits as  LR, improve=765.7484, (0 missing)
      MC_D  splits as  LR, improve=729.8675, (0 missing)
      SAN_D splits as  LR, improve=597.3361, (0 missing)
      INT_D splits as  LR, improve=473.4160, (0 missing)
      EDU_D splits as  LR, improve=401.4701, (0 missing)
  Surrogate splits:
      MC_D  splits as  LR, agree=0.598, adj=0.145, (0 split)
      INT_D splits as  LR, agree=0.567, adj=0.080, (0 split)
      RO_D  splits as  LR, agree=0.552, adj=0.047, (0 split)
      AP_D  splits as  RL, agree=0.549, adj=0.042, (0 split)
      SAN_D splits as  LR, agree=0.549, adj=0.041, (0 split)
Node number 2: 4971 observations,    complexity param=0.03106332
  predicted class=GR1  expected loss=0.2731845  P(node) =0.4703378
    class counts:  3613  1358
   probabilities: 0.727 0.273
  left son=4 (4457 obs) right son=5 (514 obs)
  Primary splits:
      SAN_D  splits as  LR, improve=164.3155, (0 missing)
      MC_D   splits as  LR, improve=160.2793, (0 missing)
      RUB_D  splits as  LR, improve=154.9652, (0 missing)
      ROFC_D splits as  LR, improve=136.2970, (0 missing)
      EDU_D  splits as  LR, improve=122.5767, (0 missing)
```

[474] The term *class* designates the category of the target response variable selected by the majority vote as the predicted class at the node. Based on the classification rules extracted in the training data, an instance should be at the end classified as one of the categories of the target response variable. However, in some papers, the terms *class* and *category* are used interchangeably.

```
Node number 3: 5598 observations,    complexity param=0.08024691
  predicted class=GR2  expected loss=0.3454805  P(node) =0.5296622
    class counts:  1934  3664
   probabilities: 0.345 0.655
  left son=6 (1637 obs) right son=7 (3961 obs)
  Primary splits:
      MC_D  splits as  LR, improve=356.5976, (0 missing)
      SAN_D splits as  LR, improve=231.8651, (0 missing)
      RO_D  splits as  LR, improve=198.3278, (0 missing)
      INT_D splits as  LR, improve=191.3404, (0 missing)
      EDU_D splits as  LR, improve=185.3776, (0 missing)
  Surrogate splits:
      ELC_D splits as  RL, agree=0.708, adj=0.001, (0 split)
Node number 4: 4457 observations,    complexity param=0.01911589
  predicted class=GR1  expected loss=0.2295266  P(node) =0.421705
    class counts:  3434  1023
   probabilities: 0.770 0.230
  left son=8 (4033 obs) right son=9 (424 obs)
  Primary splits:
      ROFC_D splits as  LR, improve=137.9592, (0 missing)
      LD_D   splits as  LR, improve=112.2650, (0 missing)
      LIT_D  splits as  LR, improve=107.2265, (0 missing)
      EDU_D  splits as  LR, improve=101.3934, (0 missing)
      NOI_D  splits as  LR, improve=101.0520, (0 missing)
  Surrogate splits:
      WALC_D splits as  LR, agree=0.959, adj=0.571, (0 split)
Node number 5: 514 observations
  predicted class=GR2  expected loss=0.348249  P(node) =0.04863279
    class counts:   179   335
   probabilities: 0.348 0.652
Node number 6: 1637 observations,    complexity param=0.02309837
  predicted class=GR1  expected loss=0.376909  P(node) =0.1548869
    class counts:  1020   617
   probabilities: 0.623 0.377
  left son=12 (1163 obs) right son=13 (474 obs)
  Primary splits:
      WATD_D splits as  LR, improve=80.39280, (0 missing)
      ROFC_D splits as  LR, improve=79.34536, (0 missing)
      SAN_D  splits as  LR, improve=77.91089, (0 missing)
      LIT_D  splits as  LR, improve=66.75541, (0 missing)
      WALC_D splits as  LR, improve=63.78232, (0 missing)
  Surrogate splits:
      SAN_D   splits as  LR, agree=0.755, adj=0.154, (0 split)
      POST_D  splits as  LR, agree=0.719, adj=0.030, (0 split)
      ENG_C_D splits as  LR, agree=0.713, adj=0.008, (0 split)
Node number 7: 3961 observations
  predicted class=GR2  expected loss=0.2307498  P(node) =0.3747753
    class counts:   914  3047
   probabilities: 0.231 0.769
Node number 8: 4033 observations
  predicted class=GR1  expected loss=0.1891892  P(node) =0.3815877
    class counts:  3270   763
   probabilities: 0.811 0.189
Node number 9: 424 observations
  predicted class=GR2  expected loss=0.3867925  P(node) =0.04011732
    class counts:   164   260
   probabilities: 0.387 0.613
Node number 12: 1163 observations,    complexity param=0.01075269
  predicted class=GR1  expected loss=0.2768702  P(node) =0.1100388
    class counts:   841   322
   probabilities: 0.723 0.277
  left son=24 (935 obs) right son=25 (228 obs)
  Primary splits:
      LIT_D  splits as  LR, improve=66.16736, (0 missing)
      ROFC_D splits as  LR, improve=57.19611, (0 missing)
      WALC_D splits as  LR, improve=49.48065, (0 missing)
      NOI_D  splits as  LR, improve=42.42247, (0 missing)
      RO_D   splits as  LR, improve=27.10818, (0 missing)
```

```
  Surrogate splits:
      NOI_D splits as  LR, agree=0.821, adj=0.088, (0 split)
      WP_D  splits as  LR, agree=0.810, adj=0.031, (0 split)
Node number 13: 474 observations
  predicted class=GR2  expected loss=0.3776371  P(node) =0.04484814
    class counts:   179   295
   probabilities: 0.378 0.622
Node number 24: 935 observations
  predicted class=GR1  expected loss=0.1935829  P(node) =0.08846627
    class counts:   754   181
   probabilities: 0.806 0.194
Node number 25: 228 observations
  predicted class=GR2  expected loss=0.3815789  P(node) =0.02157252
    class counts:    87   141
   probabilities: 0.382 0.618
```

The figure below shows the evolution of the cross-validation error rate as a function of the tree size. From this plot, we can see that when the size of tree is 6, the cross-validation error rate decreases to around 0.5. To decide based on which complexity parameter we have to prune the tree in an optimized way, we will check the minimum cross-validation error rate which here equals to 0.4830745 (see summary information above) with a standard error of 0.008608827. Based on these figures, we select the smallest tree with a cross-validation error within 0.4830745+/- 0.008608827 (between 0.4916 and 0.4744): the tree with 6 splits (7 terminal nodes). Moreover, the dotted horizontal line in figure 15 (complexity parameter vs. cross-validated error) is the upper limit of the one standard error rule: 0.4830745+1*0.08735095=0.4916. The plot suggests selecting the tree with the leftmost complexity parameter (cp) value below the line: this should be a cp number between 0.014 and 0.012.

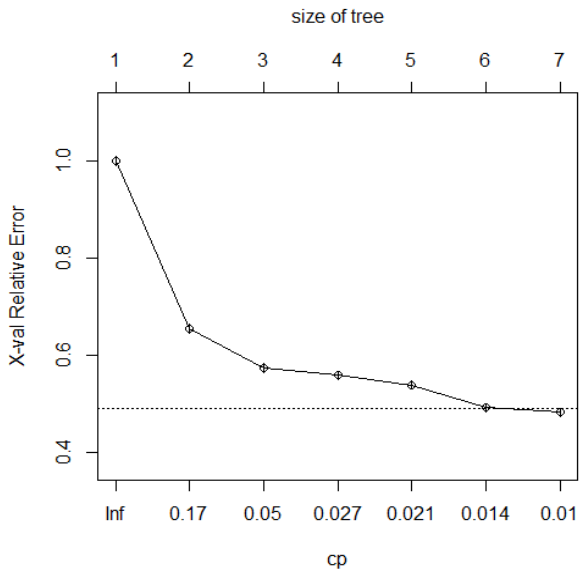


Figure 15. Cross-validated error rate by the size of tree and complexity parameter (cp), based on the model with the default parameters (CAP111S), data: South Africa GHS 2011

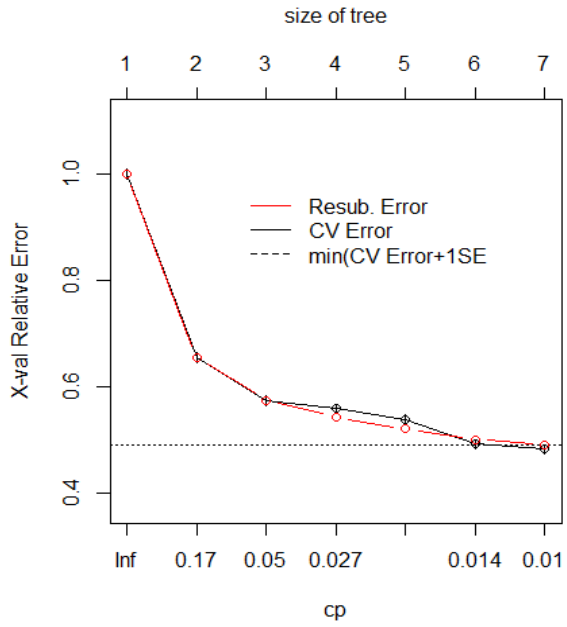


Figure 16. Cross-validated error rate by the size of tree, complexity parameter (cp) and min (CV Error+1SE) rule based on the model with default parameters (CAP111S), data: South Africa GHS 2011

We prune the tree with a cp equal to 0.012:

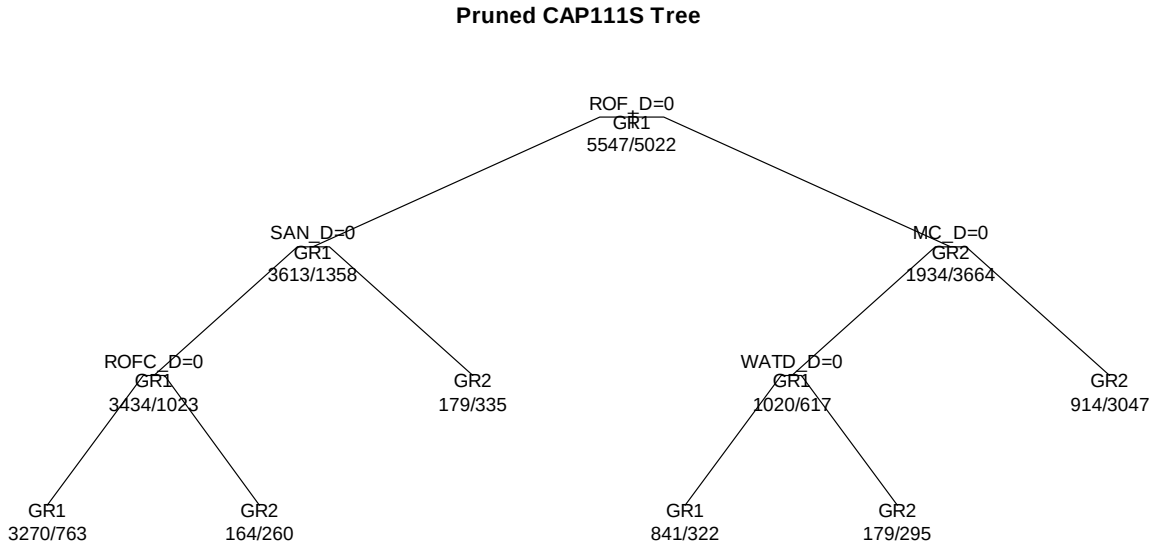


Figure 17. Pruned classification tree of the model CAP111S (with default parameters), data: South Africa GHS 2011

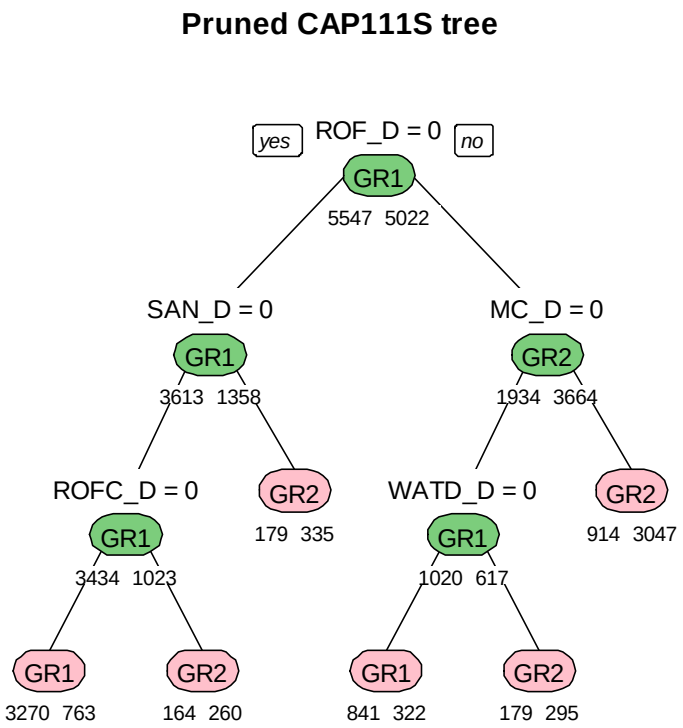


Figure 18. Pruned classification tree of the model CAP111S (with default parameters), by employing the package rpart.plot, data: South Africa GHS 2011

As a reminder, the tree should be read by starting from the root node where the split condition ROF_D=0 (being not deprived regarding the type of main material used for the roof of the dwelling) appears (there is a split condition at each internal node except terminal ones). When the split criteria at the node is fulfilled, the reader should follow the left side of the node; otherwise the right side. For instance, the left side after the root node indicates being not deprived regarding the type of main material used for the roof of the dwelling whilst the right side defines being deprived regarding the type of main material of the roof of the dwelling. The numbers appearing at each node specify the number of persons from GR1 and GR2 respectively. The root node contains the entire training set composed of 5,547 individuals from the 1st poverty group and 5,022 individuals from the 2nd poverty group. Then, this set is spread through 6 patterns. For instance, 3,613 of 5,547 persons from GR1 take the category 0 for the variable ROF_D (left side of the root node) and the rest 1,934 persons take the category 1 (right side of the root node). Each terminal node as the end of the pattern is classified by a category of the target response variable (CAP111): GR1 or GR2 (1st poverty group or 2nd poverty group). In figure 18, terminal nodes are colored in pink.

As we can see from figures 17 and 18, the 6 different patterns are discovered: (Pattern 1) being deprived regarding the type of main material used for the roof of the dwelling (ROF), and medical coverage (MC); (Pattern 2) deprived regarding the type of main material used for the roof of the dwelling (ROF), not deprived regarding medical coverage (MC) and deprived regarding the type of main source of drinking water (WATD); (Pattern 3) deprived regarding the type of main material used for the roof of the dwelling (ROF) and not deprived regarding both medical coverage (MC) and the type of main source of drinking water (WATD); (Pattern 4) not deprived regarding the type of main material used for the roof of the dwelling (ROF) and deprived regarding the type of sanitation facilities the household uses (SAN); (Pattern 5) not

deprived regarding both the type of main material used for the roof of the dwelling (ROF) and the sanitation type the household uses (SAN) but deprived regarding the roof condition of the dwelling (ROFC); (Pattern 6) not deprived regarding the type of main material used for the roof of the dwelling (ROF), the sanitation type the household uses (SAN) and the roof condition of the dwelling (ROFC). The patterns 1, 2, 4 and 5 end up with GR2 whilst the patterns 3 and 6 are classified as GR1. In each terminal node, some individuals are misclassified; for instance, if the person comes from GR1 but classified as GR2, then this person is considered misclassified. At the end of the first pattern, 914 individuals from GR1 are misclassified because by the majority vote process this pattern is classified as GR2. In total 3,937 individuals are correctly classified, which means that our model fits with an overall accuracy of 76.14 per cent. Among GR2, 3,047 persons from the pattern 1, 295 individuals from the pattern 2, 335 persons from the pattern 4 and 260 persons from the pattern 5 are detected and correctly classified as GR2. On the other hand, among GR1, 841 persons from the pattern 3 and 3,270 persons from the pattern 6 are captured and correctly classified as GR1. In total 78 per cent of GR2 and 74 per cent of GR1 are correctly classified.

According to the results, there are 5 variables emerging as the most influential ones among 33 and the capabilities patterns of GR2 are more informative than those of GR1. Notably, the patterns 3 and 6 are less reflective of disadvantages and/or non-achievements. Among the correctly classified 4,111 individuals of GR1, 79.5 per cent of them are likely to be not deprived regarding the variables ROF, SAN, and ROFC. On the other hand, the rest 20.45 per cent of them are likely to be deprived regarding the variable ROF and to not be deprived regarding the variables MC and WATD. Concerning the correctly classified 3,937 individuals of GR2, 77 per cent of them are likely to be deprived regarding both the variables ROF and MC. We can say that the type of the type of main material used for the roof of the dwelling is a deprivation issue for both GR1 and GR2, and the non-medical coverage is the most substantial deprivation item in the actual capabilities set of GR2. By moving to other poverty groups which have less advantages and achievements in terms of functionings, we will be able to see whether medical coverage, dwelling roof material and other 3 variables (ROFC, SAN, WATD) are still part of the most influential capabilities patterns.

We reach also the same results with figure 19 displayed below. However, the information is given in a different way. Root and internal nodes are shown with a circle whilst terminal nodes are shown with a vertical bar. Inside circles, the code of the split variables appears (such as ROF_D at the root node). Including terminal, there are in total 11 nodes. Root and other terminal nodes are denoted by a number upon the circle at the node. Terminal nodes are denoted from node 4 to node 11 upon the vertical bar at the end of the tree. The value/modality that the split variable takes is mentioned upon the lines connecting the nodes. For instance, the line connecting the left side of the root node ROF_D with the split variable SAN_D is denoted by 0. Such case means that the variable ROF_D taking the category 0 is followed by the variable SAN_D whilst the variable ROF_D taking the category 1 is followed by the variable MC_D. The vertical bars at the terminal nodes inform about the per centage share of each category of the response variable, given that the greatest share determines the class of the node (i.e., majority vote). At node 4, there are 4033 persons in total aggrouping under GR1 (81 per cent) and GR2 (19 per cent). We can see that GR1 emerges as the majority at the nodes 4 and 9 vs. GR2 emerges as the majority at the nodes 5,6,10 and 11.

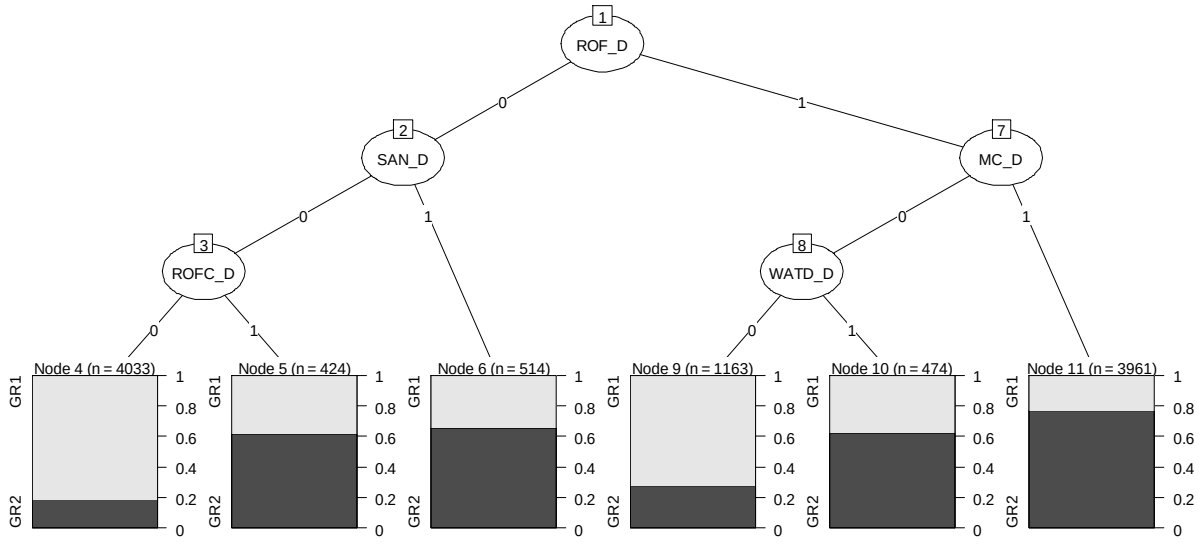


Figure 19. Pruned classification tree of the model CAP111S (with default parameters), by employing the package partykit, data: South Africa GHS 2011

From the pruning process to plotting the pruned classification tree, we used the following R codes:

```
#Prune tree:

dtreeCAP111S.pruned=prune(dtreeCAP111S, cp=0.012)

printcp(dtreeCAP111S.pruned)

summary(dtreeCAP111S.pruned)

#Plot in different ways the pruned tree:

plot(dtreeCAP111S.pruned, uniform=T,branch=.1,margin=.1,main="Pruned CAP111S Tree")

text(dtreeCAP111S.pruned, use.n=T, all=T)

plot(dtreeCAP111S.pruned, branch=.4, uniform=T, compress=T, main="Pruned CAP111S
Tree")

text(dtreeCAP111S.pruned,all=T,use.n=T)

plot(dtreeCAP111S.pruned, branch=.1, uniform=T,margin=0.1,main="Pruned CAP111S
Tree")

text(dtreeCAP111S.pruned,pretty=1,all=T,use.n=T)

#For a better plot of the tree

library(rpart.plot)

#type 2 option: the split label below each node.

#extra 104 option: probabilities for each class, and the % of observations in each
node.

#fallen leaves option: terminal nodes at the bottom of the graph.

prp(dtreeCAP111S.pruned, type = 0, extra = 1, main="Pruned CAP111S tree",
box.col=c("pink", "palegreen3") [dtreeCAP111S.pruned$frame$yval])

#0r
```

```
prp(dtreeCAP111S.pruned, type = 1, extra = 1, under=TRUE, split.font=1, varlen=-10,
main="Pruned CAP111 tree", box.col=ifelse(dtreeCAP111S.pruned$frame$var=="<leaf>",
'pink', 'palegreen3'))

plot.rpart.obj <- function(rpart.obj, font.size = 0.8) {

  ## plot decision tree

  plot(rpart.obj,

       uniform   = T,    # if 'TRUE', uniform vertical spacing of the nodes is used

       branch    = 1,    # controls the shape of the branches from parent to child
node

       compress  = F,    # if 'FALSE', the leaf nodes will be at the horizontal
plot

       nspace    = 0.1,

       margin    = 0.1, # an extra fraction of white space to leave around the
borders

       minbranch = 0.3)  # set the minimum length for a branch

  ## Add text

  text(x      = rpart.obj,   #

       splits = T,           # If tree are labeled with the criterion for the split

       all    = T,           # If 'TRUE', all nodes are labeled, otherwise just
terminal nodes

       use.n  = T,           # Use numbers to annotate

       cex    = font.size)   # Font size

}

plot.rpart.obj(dtreeCAP111S.pruned)

#Or, use the package partykit

library(partykit)

plot(as.party(dtreeCAP111S.pruned))

as.party(dtreeCAP111S.pruned)
```

As a final step we evaluated the predicted results with the actual values on the test data[475]. The confusion matrix is:

Table 36. Confusion matrix of the CAP111's predicted results on the test data (based on the fitting model with default parameters), data: South Africa GHS 2011

	Predicted	
Actual	GR1	GR2
GR1	2077 **(TP)**	697 **(FN)**
GR2	530 **(FP)**	1981 **(TN)**

[475] In the literature, the test data is also called as the validation data. We use here the terms 'test set, test sample, and test data' interchangeably, and same for the training data as well.

The accuracy of prediction is 76.78 per cent: 21 per cent of GR2 is predicted as GR1 and 25 per cent of GR2 is predicted as GR1.

```
   Accuracy : 0.7678
                 95% CI : (0.7562, 0.7792)
    No Information Rate : 0.5249
    P-Value [Acc > NIR] : < 2.2e-16
                  Kappa : 0.536
 Mcnemar's Test P-Value : 2.148e-06
            Sensitivity : 0.7487
            Specificity : 0.7889
         Pos Pred Value : 0.7967
         Neg Pred Value : 0.7397
             Prevalence : 0.5249
         Detection Rate : 0.3930
   Detection Prevalence : 0.4933
      Balanced Accuracy : 0.7688
       'Positive' Class : GR1
```

For the prediction process, we used the following R codes:

```
#Prediction

dtreeCAP111S.pred=predict(dtreeCAP111S.pruned, test.set, type="class")

dtreeCAP111S.perf=table(test.set$CAP111, dtreeCAP111S.pred, dnn=c("Actual",
"Predicted"))

dtreeCAP111S.perf

#Or

library(caret)

confusionMatrix(dtreeCAP111S.pred, test.set$CAP111)

#Or by the following codes:

dtreeCAP111S.pred=predict(dtreeCAP111S.pruned, test.set)

plabels.CAP111S=colnames(dtreeCAP111S.pred)[apply(dtreeCAP111S.pred,1,which.max)]

dtreeCAP111S.pred

classification.summary=function(plabels.CAP111S, tlabels){

  TP=sum((plabels.CAP111S=="GR1")*(tlabels=="GR1"))

  FP=sum((plabels.CAP111S=="GR1")*(tlabels=="GR2"))

  TN=sum((plabels.CAP111S=="GR2")*(tlabels=="GR2"))

  FN=sum((plabels.CAP111S=="GR2")*(tlabels=="GR1"))

  A=(TP+TN)/(TP+TN+FP+FN)

  sens=TP/(TP+FN)

  spec=TN/(TN+FP)

  prec=TP/(TP+FN)

  C=matrix(c(TP,FP,FN,TN),2,2)

  colnames(C)=c("predicted GR1", "predicted GR2")

  rownames(C)=c("actual GR1", "actual GR2")
```

```
  return(list(A=A,TP=TP,FP=FP,TN=TN,FN=FN,C=C,sens=sens,spec=spec))

}

s=classification.summary(plabels.CAP111S,test.set$CAP111)

print(s)
```

In order to increase the accuracy of the tree, we change the cp parameter in fitting the model, and we set a smaller cp=0.0025 given that other parameters remain in their default values. We call this model CAP111S2 (but the target response variable is still called CAP111!).

The results of this fitting show that (see below) we are able to obtain a slightly lower cross-validated error rate (xerror) compared to the previous tree. However, it is important to note that the first tree is quite saturated with a substantial number of splits (see the tree CAP111S2 below). We therefore need to prune the tree in an optimal way by using an adequate complexity parameter (cp). To achieve that, we will plot cross-validated, resubstitution error rates, cp and check the error rates regarding the 1-SE (one standard error) Rule.

```
Call:
rpart(formula = CAP111 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = tempS2)
  n= 10569
           CP nsplit rel error    xerror        xstd
1  0.344484269      0 1.0000000 1.0000000 0.010222892
2  0.080246914      1 0.6555157 0.6555157 0.009480090
3  0.031063321      2 0.5752688 0.5752688 0.009123484
4  0.023098367      3 0.5442055 0.5573477 0.009032715
5  0.019115890      4 0.5211071 0.5410195 0.008946256
6  0.010752688      5 0.5019912 0.4970131 0.008694530
7  0.009889818      6 0.4912386 0.4878534 0.008638546
8  0.008960573      9 0.4615691 0.4830745 0.008608827
9  0.007566706     10 0.4526085 0.4641577 0.008487668
10 0.006969335     11 0.4450418 0.4412585 0.008333213
11 0.006819992     12 0.4380725 0.4328953 0.008274590
12 0.005575468     19 0.3767423 0.4054162 0.008073207
13 0.004480287     20 0.3711669 0.3853047 0.007916858
14 0.003982477     23 0.3562326 0.3799283 0.007873712
15 0.003916102     24 0.3522501 0.3747511 0.007831610
16 0.003783353     27 0.3405018 0.3699721 0.007792256
17 0.003584229     28 0.3367184 0.3675826 0.007772400
18 0.003285544     29 0.3331342 0.3643967 0.007745737
19 0.003185982     31 0.3265631 0.3586221 0.007696858
20 0.003136201     32 0.3233771 0.3580247 0.007691760
21 0.002986858     38 0.3044604 0.3562326 0.007676421
22 0.002887296     39 0.3014735 0.3492632 0.007616094
23 0.002588610     41 0.2956989 0.3387097 0.007522658
24 0.002500000     47 0.2801673 0.3283552 0.007428465
```

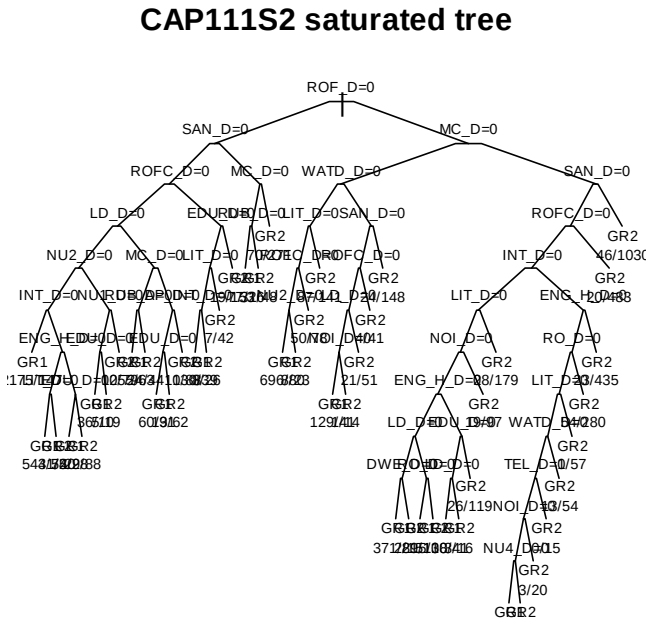


Figure 20. CAP111S2 saturated tree. We add S2 next to CAP111 as this is the second attempt to obtain a better classification tree for an optimal interpretation, data: South Africa GHS 2011

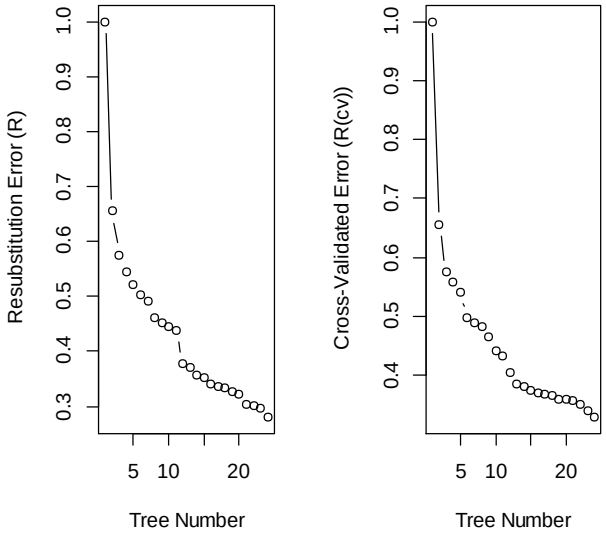


Figure 21. Resubstitution and cross-validated error rates for the CAP111S2 classification tree, data: South Africa GHS 2011

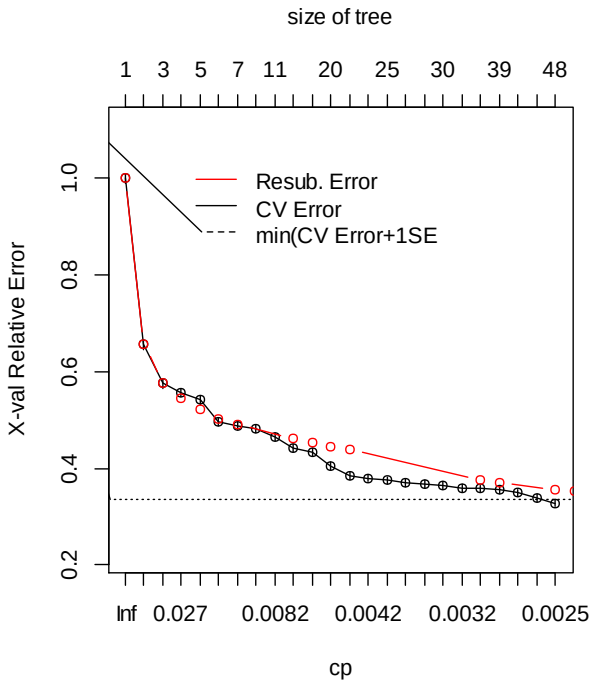


Figure 22. Resubstitution, cross-validated error rates and the 1-SE Rule (min (CV Error + 1 SE)) to find out the optimum cp for pruning process, data: South Africa GHS 2011

If we decide based on the 1-SE Rule[476], it is clear from Figure 22 that for pruning the tree in an optimal accuracy, the cp around 0.0020 should be used. However, such cp provides at the same time a quite complicated tree with around 40 nodes. Interpretation of patterns with such tree will be a difficult task and even not possible given that the visualization of saturated tree won't allow to see properly each pattern separately. This is why, we have to decide on a certain cp level yielding a higher accuracy compared to CAP111S tree, but it might be still with some misclassification. Interpretation of between 10 and 12 patterns seems feasible, which will provide us a sufficiently rich content of capabilities emerging from the data. Regarding to these conditions, we set cp at 0.006 for pruning process. We remind that as a result of prediction in the training data, each terminal node is assigned to a class which corresponds in fact to one of the categories of the target response variable of the model (here, there are two possible classes at terminal nodes: GR1 or GR2 as they are the two possible outputs of CAP111 target response variable). Our data is composed of individuals. Each individual is classified as GR1 or GR2 after the prediction in the training data. As the pruned results show below, the cp of exact 0.006 gives in fact 20 patterns, a higher number than expected (i.e., corresponding to a tree with 19 splits and 20 terminal nodes/leaves[477], and with a cross-validation error rate of 0.40542).

```
> dtreeCAP111S2.pruned=prune(dtreeCAP111S2, cp=0.006)
> printcp(dtreeCAP111S2.pruned)

Classification tree:
rpart(formula = CAP111 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = tempS2)
```

[476] In the graphics we obtain from rpart (e.g., Figure 22 above), 1-SE Rule is meant by min(CV Error+1SE) where CV is the cross-validation error rate and SE is the standard error of the CV.

[477] By definition, the number of leaves/terminal nodes is always one plus than the number of splits in the tree. A tree with *k* splits has *k+1* terminal nodes.

```
Variables actually used in tree construction:
 [1] EDU_D   ENG_H_D INT_D   LD_D    LIT_D   MC_D    NOI_D   NU2_D   ROF_D
[10] ROFC_D  SAN_D   WATD_D
Root node error: 5022/10569 = 0.47516
n= 10569

          CP nsplit rel error  xerror      xstd
1  0.3444843      0   1.00000 1.00000 0.0102229
2  0.0802469      1   0.65552 0.65552 0.0094801
3  0.0310633      2   0.57527 0.57527 0.0091235
4  0.0230984      3   0.54421 0.55735 0.0090327
5  0.0191159      4   0.52111 0.54102 0.0089463
6  0.0107527      5   0.50199 0.49701 0.0086945
7  0.0098898      6   0.49124 0.48785 0.0086385
8  0.0089606      9   0.46157 0.48307 0.0086088
9  0.0075667     10   0.45261 0.46416 0.0084877
10 0.0069693     11   0.44504 0.44126 0.0083332
11 0.0068200     12   0.43807 0.43290 0.0082746
12 0.0060000     19   0.37674 0.40542 0.0080732
```

The confusion matrix we get by comparing the predicted and actual values on the test data provides information on the accuracy of our model in prediction[478]. There are 2,314 persons from GR1 correctly predicted as GR1 on the test data (called *true positives*) and 2,010 persons from GR2 correctly predicted as GR2 on the test data (called *true negatives*), given that GR1 is considered positive class and GR2 is considered negative class. However, 501 persons from GR2 are predicted as GR1 that are called *false positives* (FP) and 460 persons from GR1 are predicted as GR2 that are called *false negatives* (FN). The FP and FN lowers the accuracy of the prediction, producing a misclassification rate of ((460 + 501) / (2314+501+460+2010) =0.18) 18 per cent. In other words, 82 per cent of the test data can be correctly classified by using the classification patterns extracted on the training data. We achieved here to increase the accuracy by 5 per cent by considering more patterns and employing a lower value of cp in pruning. It is important to note that the prediction does not yield a correct-classification around 90 per cent or plus; however, given the complexity of the data that we mentioned earlier in terms of available patterns of capabilities, in such data, obtaining a correct-classification around 80 per cent or more can be also considered a relatively good result. The improvement in terms of reducing more the complexity parameter will be resulting in increasing the number of patterns to interpret which will limit our capability to yield a synthetized information. The accuracy of the model depends also on how the regressors are able to distinguish correctly the categories of the target response variable among the observations/cases of the data (i.e., individuals here). Another way of obtaining a model with a high accuracy is playing with the model by modifying the regressors in the model. We preferred to keep the 33 regressors and instead changing the cp for a better prediction.

The measures of predictive accuracy such as specificity and sensitivity measures become especially very important when the data is unbalanced. In our case, as we have showed at the beginning of the analysis, we have a quite balanced data. Regarding Kabacoff (2015, 405-406), the sensitivity is "the probability of getting a positive classification when the true outcome is

[478] For information we note that generally the rows in confusion matrix are the actual values and the columns are the predictive values but, in the output, obtained by using the package caret in R the confusion matrix is presented with prediction values at row and the actual (reference) values at column. However, this is just a matter of illustration and it has no effect on the results.

positive (also called true positive rate or recall)", the specificity is "the probability of getting a negative classification when the true outcome is negative (also called true negative rate)". Moreover, he defines the positive predictive value as "the probability that an observation with a positive classification is correctly identified as positive (also called precision), the negative predictive value as "the probability that an observation with a negative classification is correctly identified as negative" and the accuracy as "the proportion of observations correctly identified (also called ACC)". These three statistics correspond to: 82.20 (= ((2314)/ (2314+501)) *100), 81.38(= ((2010)/ (460+2010)) *100), and 81.82 per cent, respectively. The sensitivity value of our prediction is equal 83.42 per cent (= ((2314)/ (2314+460)) *100) and the specificity value is equal 80 per cent (= ((2010)/ (2010+501)) *100) which are quite close to accuracy rate. For instance, suppose that there is a confusion matrix as follows given that X is the positive class[479]:

Table 37. An example of confusion matrix for explaining accuracy

	Predicted	
Actual	X	Y
X	5	15
Y	10	130

In such case, the overall accuracy is 84.37 per cent. The item X is correctly classified by 25 per cent and the item Y is correctly classified by 93 per cent. It is important to note that such data is very unbalanced with 20 X and 140 Y corresponding to 12.5 and 87.5 per cent of the data respectively. The positive predictive value (precision) equals to (5/ (5+10)) *100=33 per cent and the negative predictive value equals to (130/ (130+15)) *100=89.65 per cent. On the other hand, the specificity corresponds to (130/ (130+10)) *100= 93 per cent, and the sensitivity corresponds to (5/ (5+15)) *100=25 per cent. Thus, the specificity is in fact the true negative rate and the sensitivity is the true positive rate. As we can see from the results, in such unbalanced data, the overall accuracy rate is not an adequate tool for informing about the efficiency of the prediction.

If we return to our analysis, from the confusion matrix statistics below, we can see that there is a no-information rate of 52.49 per cent. This rate designates the per centage of the largest class in the data. In our case, GR1 has the largest class. So that, this rate can be obtained via the division of the number of persons having GR1 as the reference group (actual values) by the total number of persons in the data: (2314+460) / (2314+460+501+2010). It is important that the accuracy rate is higher than the no information rate. The p-value test shows that the accuracy is significantly higher than the no information rate. The prevalence is the rate of the positive class (GR1): (2314+460) / (2314+460+501+2010) =0.5249. The detection rate is the rate of individuals from positive class predicted also as being positive class among the total population of the test set: (2314/ (2314+501+460+2010)) =43.78. The detection prevalence is the rate of individuals predicted as the positive class among the total population of the test set: (2314+501) / (2314+501+460+2010) =53.26. The McNemar's test is statistical test to detect whether the row and column marginal frequencies are equal, in a 2*2 contingency table. Here, the result of

[479] As mentioned in the description of the package, in the case of two categories/modalities for a categorical variable, the first category/modality will be used as the positive class (please see the link below-accessed 29 September 2019):
https://cran.r-project.org/web/packages/caret/caret.pdf

McNemar test is not significant so that the null hypothesis (i.e., the two marginal probabilities for each category of the target response variable are the same) is not rejected. The Kappa statistic shows the similarity between the two types of misclassification (false positive and false negative). We obtain a value of 0.6351 which can be seen as a sign of a moderate level of agreement in terms of interrater reliability.

```
Confusion Matrix and Statistics
          Reference
Prediction  GR1  GR2
       GR1 2314  501
       GR2  460 2010
               Accuracy : 0.8182
                 95% CI : (0.8075, 0.8285)
    No Information Rate : 0.5249
    P-Value [Acc > NIR] : <2e-16
                  Kappa : 0.6351
 Mcnemar's Test P-Value : 0.1969
            Sensitivity : 0.8342
            Specificity : 0.8005
         Pos Pred Value : 0.8220
         Neg Pred Value : 0.8138
             Prevalence : 0.5249
         Detection Rate : 0.4378
   Detection Prevalence : 0.5326
      Balanced Accuracy : 0.8173
       'Positive' Class : GR1
```

The description of the confusion matrix statistics from caret package mentioned above is given at http://127.0.0.1:28726/library/caret/html/confusionMatrix.html (accessed 29 September 2019) as follows[480]: "Suppose a 2x2 table with notation

	Reference	
Predicted	Event	No Event
Event	A	B
No Event	C	D

The formulas used here are:

```
Sensitivity = A/(A+C)

Specificity = D/(B+D)

Prevalence = (A+C)/(A+B+C+D)

PPV = (sensitivity * prevalence)/((sensitivity*prevalence) + ((1-specificity)*(1-
prevalence)))

NPV    =    (specificity    *    (1-prevalence))/(((1-sensitivity)*prevalence)    +
((specificity)*(1-prevalence)))

Detection Rate = A/(A+B+C+D)
```

[480] See also Kuhn (2008) at http://www.jstatsoft.org/article/view/v028i05/v28i05.pdf (accessed 29 September 2019).

```
Detection Prevalence = (A+B)/(A+B+C+D)

Balanced Accuracy = (sensitivity+specificity)/2

Precision = A/(A+B)

Recall = A/(A+C)

F1 = (1+beta^2)*precision*recall/((beta^2 * precision)+recall)

where beta = 1 for this function.
```

For more than two classes, these results are calculated comparing each factor level to the remaining levels (i.e., a "one versus all" approach). The overall accuracy and unweighted Kappa statistic are calculated. A p-value from McNemar's test is also computed using mcnemar.test (which can produce NA values with sparse tables). The overall accuracy rate is computed along with a 95 percent confidence interval for this rate (using binom.test) and a one-sided test to see if the accuracy is better than the "no information rate," which is taken to be the largest class percentage in the data".

Based on the evaluation results we obtained in terms of predictive accuracy, we can say that the pruned tree of our classification model provides relatively good results however it is important to accept that they are not perfect. Good prediction demands a high overall accuracy but also low level of misclassification regarding all categories/modalities of the target response variable. For that reason, to check in detail misclassifications, we computed additional accuracy statistics such as specificity and sensitivity. Misclassification depends on several conditions: (i) the initial data should be balanced (i.e., for instance if the target response variable is a categorical variable with 2 modalities then the best case is 50 per cent of the observations/cases having the modality 1 and the rest 50 per cent having the modality 2. But the distribution of 60 vs. 40 per cent can be also seen as an acceptable threshold); (ii) the optimal pruned tree should be selected based on the cross-validation error rate in case of a saturated tree. In a saturated tree, it is very complicated and even impossible in very large cases the interpretation of patterns. The full information is at maximum level and hence the misclassification rate is at minimum level when the tree reaches the highest size. It is normal that when we prune the tree, we lose some patterns which means loss of information and this can in turn generate higher misclassification rate of one category/modality to another. For instance, it can happen that the terminal nodes classified as category 1 (in other words predicted as category 1) may be pruned/cut in a higher proportion compared to those classified as category 2. Such case risks to generate a higher misclassification rate for the category 1. When the misclassification rates of the categories are not equal or significantly far from equal, then accuracy of prediction will decrease as well; (iii) because of complexity of the data, a substantial amount of profiles may emerge (as in our case) resulting in need of many patterns for being able to distinguish the categories of the target response variable. Unbalanced data problem can be solved by employing some sampling techniques recommended in the statistical literature. For making optimal the pruning process, it is crucial to produce cross-validation error rate based on 10^{th} fold cross-validation at least, and not be contented with resubstitution error rate. Based on the cross-validation error rate, then the decision on pruning the tree is done with complexity parameter. So that, it is also important to know the role of complexity parameter (cp) and not set a high level of cp at the beginning of the analysis. After the first tree, as a function of results in complexity parameter table where

cross-validation error rates appear the parameters for pruning process should be decided. Moreover, loss matrix or prior probabilities can be added in fitting the model for weighting more a particular misclassification however it is a quite arbitrary and delicate process and it probably requires a good level of knowledge on both theoretical and application aspects of the study. By playing with the regressors in the sense that by removing and adding different regressors, the quality of prediction can also be checked. But there are no specific assumptions to be respected for implementing the analysis such as normality assumption or a certain degree of correlation between variables.

More advanced ensemble techniques or random forest method can also be applied for testing the accuracy of the model for further examination of prediction. Confusion matrix statistics based on the test data as we did in our analysis is a plus. It is important to distinguish cross-validation process done on the training data and the division of the initial data into training vs. test data (or called sample/subset/set). The 10^{th} fold cross -validation process as applied in rpart package for the classification divides randomly the training data into 10 subsets and the cross-validation error rate is independently produced at each fold. However, before going into that process, we also divided at the beginning the initial data into training (70 per cent) and test (30 per cent) sets. This initial division is an attempt to measure later in detail the accuracy of prediction via confusion matrix, but it is not obligatory. In our analysis, training vs. test data sets partition provides a higher accuracy compared to the prediction based on the full data even if the cross-validation is applied. When the data is separated into training vs. test samples, the distribution of the categories of the target response variable should be in balance and respect the same or similar shares as in the full data. For instance, if the ratio of category 1 to category 2 is 0.9 in the full data, then the training data should also respect this ratio which is quite balanced and in the test data as well. The partition of training vs. test data may be also useful when the full data is complex including a significantly high number of profiles reflecting combinations of the regressors. The initial classification tree risks to be higher with the full data compared to the training data. With less observations, complexity of patterns can be dealt in an easier way (if we assume that more observations produce more variety of patterns or if the data has a lot of patterns due to plurality of individuals' profiles). At least with 70 per cent of the initial data, the classification rules can be extracted with less complexity and then the prediction of these rules can be tested on the remaining test data for evaluating the accuracy. This last point is just a proposition as we did not perform any statistical test on that and in some cases the data may be classified easily with one or two variables even if there is a substantial amount of observations in the data. The complexity is due to the profiles in other words to the behavior of each observation through the regressors that we cannot necessarily know in advance (e.g., national living conditions surveys). In dealing with the complex data, it is important to not fall into trap of reduction, instead a synthetic picture of data should be captured with efficient information and relatively good prediction by reflecting at the same time a certain variety of patterns.

From the figure of the pruned CAP111S2 tree below, we can observe that there are in total 7 patterns of GR1 and 13 patterns of GR2. At the root node, there are 5,547 persons from GR1 and 5,022 persons from GR2 forming together the entire population of the training data. The population of both GR1 and GR2 are then spread across the branches of the tree under the different patterns ending with terminal nodes. As the overall accuracy statistics show we reach to gather 84 per cent of information for GR1 and 80.4 per cent of information for GR2 through the discovered patterns. In other words, 84 per cent of people from GR1 are likely to be detected

through the seven GR1 patterns and on the other hand, 80.4 per cent of people from GR2 are likely to be captured through the thirteen GR2 patterns with our model in the pruned tree. The patterns of GR1 are less informative than those of GR2 in terms of non-achievement and/or disadvantage.

The discovered 20 patterns are as follows (starting from the left side of the tree): (1) not deprived regarding the type of main material of the roof of the dwelling (ROF), sanitation (toilet facility) type the household uses in the dwelling (SAN), roof condition of the dwelling (ROFC), land degradation (LD) and nutrition (i.e., cutting the size of meals because there was not enough food in the house) (NU2); (2) not deprived regarding ROF, SAN, ROFC, LD but deprived regarding NU2; (3) not deprived regarding ROF, SAN, ROFC, medical coverage (MC) but deprived regarding LD; (4) not deprived regarding ROF, SAN, ROFC but deprived regarding LD and MC; (5) not deprived regarding ROF, SAN, educational attainment (EDU), littering situation where the dwelling is located (LIT) but deprived regarding ROFC; (6) not deprived regarding ROF, SAN, EDU but deprived regarding ROFC and LIT; (7) not deprived regarding ROF, SAN but deprived regarding ROFC and EDU; (8) not deprived regarding ROF and MC but deprived regarding SAN; (9) not deprived regarding ROF but deprived regarding SAN and MC; (10) not deprived regarding MC, access to drinking water (WATD) and LIT but deprived regarding ROF; (11) not deprived regarding MC and WATD but deprived regarding ROF and LIT; (12) not deprived regarding MC but deprived regarding ROF and WATD; (13) deprived regarding ROF, MC and SAN; (14) not deprived regarding SAN but deprived regarding ROF, MC and ROFC; (15) not deprived regarding SAN and ROFC but deprived regarding ROF, MC and access to internet (INT); (16) not deprived regarding SAN, ROFC and INT but deprived regarding ROF, MC and LIT; (17) not deprived regarding SAN, ROFC, INT and LIT but deprived regarding ROF, MC, and excessive noise (noise pollution) where the dwelling is located (NOI); (18) not deprived regarding SAN, ROFC, INT, LIT, NOI and the dwelling's main source of energy for heating (ENG_H) but deprived regarding ROF and MC; (19) not deprived regarding SAN, ROFC, INT, LIT, NOI but deprived regarding ROF, MC and ENG_H; (20) not deprived regarding SAN, ROFC, INT, LIT, NOI but deprived regarding ROF, MC, ENG_H and EDU.

The patterns of GR1 are 1^{st}, 3^{rd}, 5^{th}, 8^{th}, 10^{th} and 19^{th} whilst the others are of GR2. The 1^{st} pattern of GR1 does not yield any information on disadvantages and non-achievements, which concerns 51.4 per cent of the population of GR1. The 10^{th} pattern covers 13.6 per cent of the population of GR1, then it is followed by 18^{th}, 3^{rd}, 5^{th}, 8^{th},19^{th} patterns with 7.3, 5, 2.5, 2, and 1.85 per centages. From the most frequent to the least frequent patterns of GR1 are mentioned in the table below: first column indicates the number of pattern, second column informs how many per cent of people classified as GR1 are detected by the given pattern, third column is the cumulative frequency of the percentages given in second column, fourth column informs how many per cent of people of GR1 are captured by the given pattern, and last column is the cumulative frequency of the percentages given in fourth column. In second column, the denominator is 4,641 whilst in fourth column the denominator is 5,547 for GR1. As a recall, 5,547 is the total number of people of GR1 in the training sample. The discovered 7 patterns do not include 906 people of GR1. These people are either part of other patterns which do exist in the saturated tree but do not exist in the pruned tree or they are misclassified in the pruned tree. The second column statistics is only based on the people of GR1 who are correctly classified as GR1 by the patterns in the pruned tree. On the other hand, the fourth column statistics includes not only people from GR1 who are correctly classified as GR1 by the pruned tree

patterns but also (i) people from GR1 who are misclassified (classified as GR2) by the pruned tree patterns; and (ii) people from GR1 who are not part of the pruned tree. In other words, the denominator of the fourth column statistics includes the entire population of GR1.

Table 38. Coverage of GR1's patterns among the population classified as GR1 (in per centage), among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP111S2

Pattern no	Coverage of pattern (in %) among the population classified as GR1 (A)	Cumulative coverage (in %) of (A)	Coverage of pattern (in %) among the total population of GR1 (B)	Cumulative coverage (in %) of (B)
1	61.4	61.4	51.4	51.4
10	16.2	77.6	13.6	65
18	8.8	86.4	7.3	72.3
3	6	92.4	5	77.3
5	3	95.4	2.5	79.8
8	2.3	97.7	2	81.8
19	2.2	**100**	1.85	83.65 ~ **84**

We can also say that through these 7 patterns, deprivation regarding ROF, MC, LD, ROFC, SAN and ENG_H concern 22.8, 9.2, 5, 2.5, 2, and 1.85 per cent of the entire population of GR1 (among 5547 persons) respectively.

Concerning GR2, the 13th pattern covers 20.5 per cent of the population of GR2, then it is followed by 15th 14th 12th, 9th, 16th, 4th (16th and 4th are equal), 7th, 11th, 20th, 2th, 17th, and 6th patterns with 18.6, 9.6, 6, 5.4, 3.6, 3.6, 3, 2.8, 2.4, 2.15, 2 and 0.83 per centages respectively. From the most frequent to the least frequent patterns of GR2 are mentioned in the table below. In second column, the denominator is 4,036 whilst in fourth column the denominator is 5,022 for GR2. As a recall, 5,022 is the total number of people of GR2 in the training sample. The discovered 13 patterns do not include 986 people from GR2. These people are either part of other patterns which do exist in the saturated tree but do not exist in the pruned tree or they are misclassified in the pruned tree (see also the explanations given for GR1).

Table 39. Coverage of GR2's patterns among the population classified as GR2 (in per centage), among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP111S2

Pattern no	**Coverage of pattern (in %) among the population classified as GR2 (C)**	**Cumulative coverage (in %) of (C)**	**Coverage of pattern (in %) among the total population of GR2 (D)**	**Cumulative coverage (in %) of (D)**
13	25.5	25.5	20.5	20.5
15	23.2	48.7	18.6	39.1
14	12	60.7	9.6	48.7
12	7.3	68	6	54.7
9	6.7	74.7	5.4	60.1
16	4.4	79.1	3.6	63.7
4	4.4	83.5	3.6	67.3
7	3.8	87.3	3	70.3
11	3.5	90.8	2.8	73.1
20	3	93.8	2.4	75.5
2	2.7	96.5	2.15	77.65
17	2.4	98.9	2	79.65
6	1	**100**	0.83	80.48~**80.5**

Pruned CAP111S2 tree

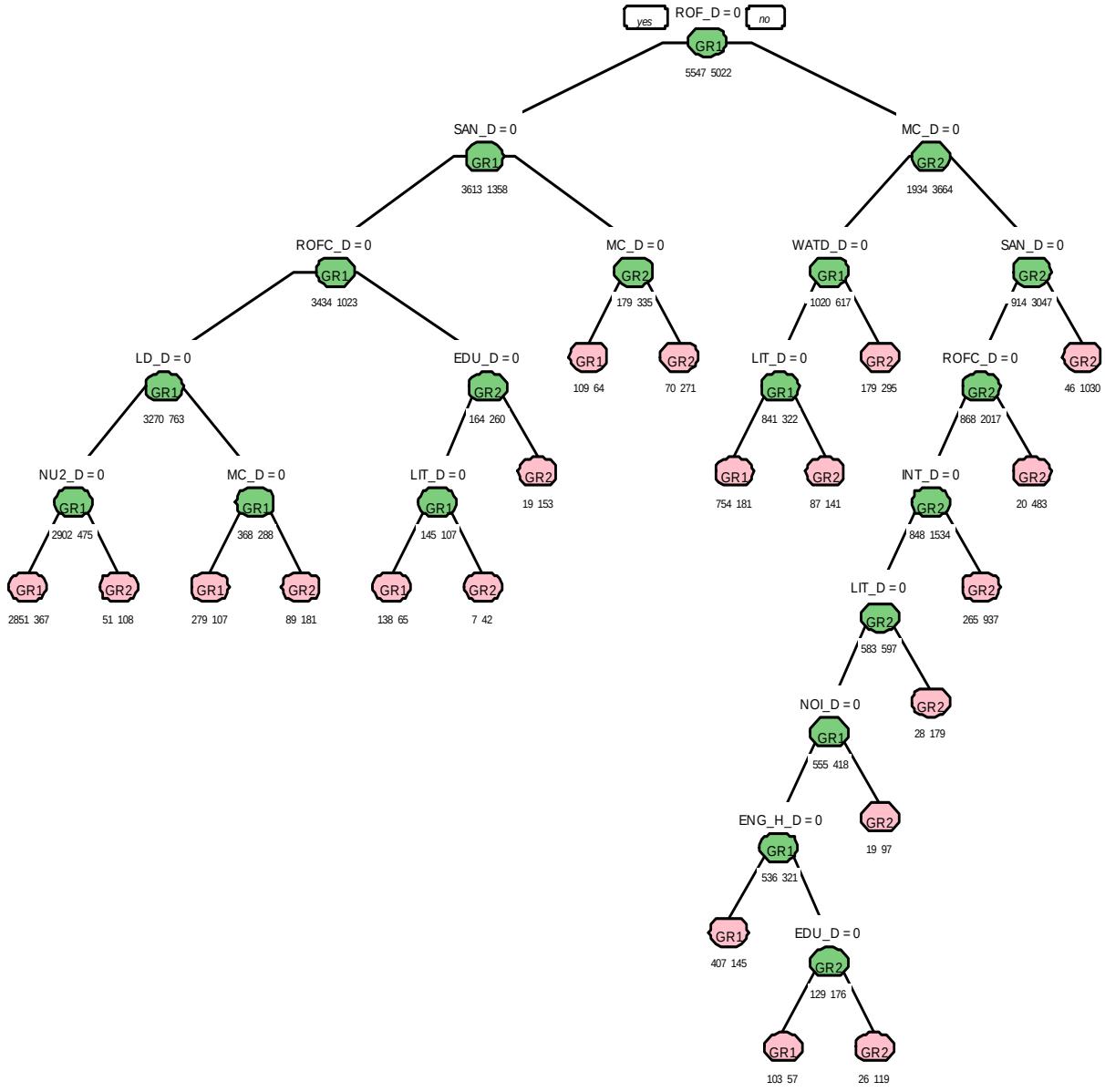


Figure 23. Pruned classification tree CAP111S2, data: South Africa GHS 2011

For GR2, through the 13 patterns, deprivation regarding MC, ROF, SAN, INT, ROFC, LIT, WATD, EDU, LD, ENG_H, NOI, and NU2 concern 65.7, 65.5, 26, 18.6, 15.4, 7.2, 6, 5.4, 3.6, 2.4, 2.15, and 2 per cent of the entire population of GR2 (among 5,022 persons) respectively.

In total, 12 variables emerge as the most influential ones among 33 to distinguish GR1 and GR2. Deprivation regarding ROF, MC, LD, ROFC, SAN, and ENG_H are the common for both GR1 and GR2 but in a different magnitude.

The pruned tree CAP111S2 provides 20 patterns in total as we can see from the explanations above. However, such number of discovered patterns can be still making harder getting a synthesized information notably for comparisons among the groups. For that reason, we will prune the tree CAP111S with a cp equal to 0.006969335 (11 splits with 12 terminal nodes). We call this model Pruned2 CAP111S2. This process gives the following results:

```
Classification tree:
rpart(formula = CAP111 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = tempS2)

Variables actually used in tree construction:
[1] EDU_D  LD_D   LIT_D  MC_D   NU2_D  ROF_D  ROFC_D SAN_D  WATD_D

Root node error: 5022/10569 = 0.47516

n= 10569

          CP nsplit rel error  xerror      xstd
1  0.3444843      0   1.00000 1.00000 0.0102229
2  0.0802469      1   0.65552 0.65552 0.0094801
3  0.0310633      2   0.57527 0.57527 0.0091235
4  0.0230984      3   0.54421 0.55735 0.0090327
5  0.0191159      4   0.52111 0.54102 0.0089463
6  0.0107527      5   0.50199 0.49701 0.0086945
7  0.0098898      6   0.49124 0.48785 0.0086385
8  0.0089606      9   0.46157 0.48307 0.0086088
9  0.0075667     10   0.45261 0.46416 0.0084877
10 0.0069693     11   0.44504 0.44126 0.0083332
```

Compared to CAP111S2 pruned tree, we obtain here a tree with 11 splits and 12 patterns in total. For GR1, we don't have any more the 5th, 6th, 18th, and 9th patterns, and we have a new modified version of the 5th pattern. This new pattern is: no deprivation regarding ROF, SAN, and EDU but deprivation regarding ROFC covering 145 persons of GR1. On the other hand, for GR2, the 13th, 14th, 15th, 16th, 17th, 18th, 19th, and 20th patterns do not exist anymore but there is a new modified version of the 12th pattern. This new pattern is: deprivation regarding ROF, and MC covering 3047 persons of GR2. We lose in fact 9 per cent of information for GR1 and we win 3 per cent of information for GR2. The term information is employed in terms of population coverage in per centage; in other words, the patterns are able to capture how per cent of population. If we look at in terms of content, by pruning the tree we have less combinations of variables. In terms of overall accuracy, 3 per cent of information is lost and sensitivity is increased by 6 per cent and specificity is decreased by 3 per cent. We therefore reach to capture more people of GR2 with less patterns compared to the results of CAP111S2 pruned tree. Certainly, 20 patterns include more combinations of variables, but it makes also more complicated the comparison among the groups.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR2
       GR1 2080  424
       GR2  694 2087

               Accuracy : 0.7885
                 95% CI : (0.7772, 0.7994)
    No Information Rate : 0.5249
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.578
```

```
Mcnemar's Test P-Value : 8.618e-16

            Sensitivity : 0.7498
            Specificity : 0.8311
         Pos Pred Value : 0.8307
         Neg Pred Value : 0.7504
             Prevalence : 0.5249
         Detection Rate : 0.3936
   Detection Prevalence : 0.4738
      Balanced Accuracy : 0.7905

       'Positive' Class : GR1
```

As long as we move towards more deprived groups, we will be able to see more clearly whether the differences among the groups become more pronounced in terms of nature and magnitude. At the end of the analysis, we will provide a comparative summary of patterns distinguishing groups. Now, we continue with the CAP112 analysis focusing on the 1st poverty group vs. 3rd poverty group.

CAP112[481]: GR1 vs. GR3 analysis

In this analysis, we focus on DEP_GRS1 vs. DEP_GRS3. We use the data containing only people from the poverty group 1 and the poverty group 3. The outcome/target response variable is dichotomous, namely CAP112, taking the value 1 for the group 1 and the value 3 for the group 3 (so that people identified in the group 1 take the value 1 and people identified in the group 3 take the value 3 according to CAP112).

Table 40. Training vs. test samples in CAP112 analysis, data: South Africa GHS 2011

	GR1	GR3
Training sample	5547	7156
Test sample	2774	3579

As the table above shows, there are 12,703 individuals in the training data (corresponding to 67 per cent of the initial data) and 6,353 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 44 per cent for GR1 and 56 per cent for GR3 in both training and test samples (in the initial data, the proportion is also 44 vs. 56 per cent).

[481] The meaning of CAP112 is: Capabilities analysis (CAP) of 1st step (1) for 2nd target group (2).

Pruned2 CAP111S2 tree

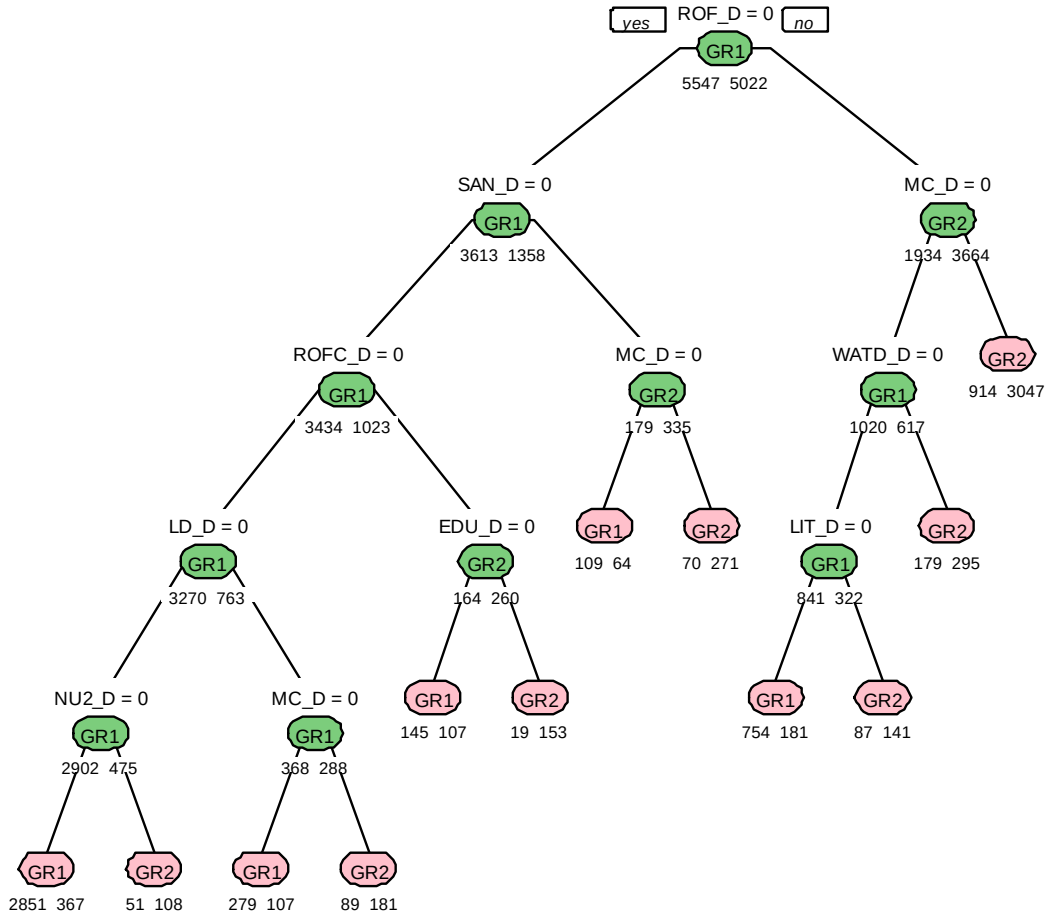


Figure 24. Pruned2 CAP111S2 tree, data: South Africa GHS 2011

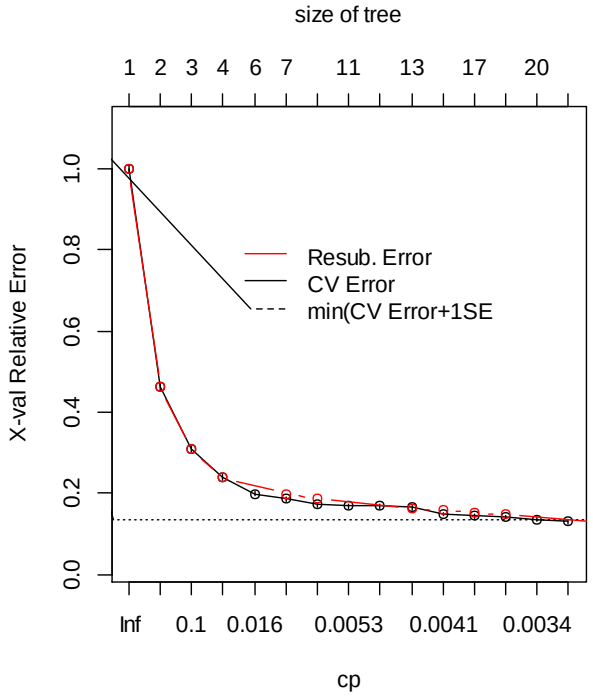


Figure 25. Resubstitution, cross-validation error rates by cp and size of tree and 1-SE Rule for CAP112 saturated tree, data: South Africa GHS 2011

```
Classification tree:
rpart(formula = CAP112 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
 [1] AP_D    EDU_D   ENG_C_D INT_D   LIT_D   MC_D    NOI_D   NU1_D   NU2_D
[10] REF_D   ROF_D   ROFC_D  SAN_D

Root node error: 5547/12703 = 0.43667

n= 12703

          CP nsplit rel error  xerror      xstd
1  0.5379484      0   1.00000 1.00000 0.0100775
2  0.1532360      1   0.46205 0.46205 0.0081542
3  0.0695872      2   0.30882 0.30882 0.0069401
4  0.0205516      3   0.23923 0.23923 0.0062147
5  0.0117180      5   0.19813 0.19813 0.0057120
6  0.0077519      6   0.18641 0.18641 0.0055560
7  0.0054083      9   0.16315 0.17235 0.0053602
8  0.0052281     10   0.15774 0.16856 0.0053057
9  0.0048675     11   0.15251 0.16820 0.0053005
10 0.0045069     12   0.14765 0.16477 0.0052505
11 0.0037858     15   0.13034 0.14855 0.0050043
12 0.0036056     16   0.12655 0.14404 0.0049330
13 0.0035154     17   0.12295 0.14152 0.0048924
14 0.0032450     19   0.11592 0.13629 0.0048071
15 0.0025000     20   0.11267 0.13070 0.0047136
```

The results of the CAP112 tree show that there is no need to continue with pruning process. Following the 1-SE Rule as shown in Figure 25 above, the tree with 20 terminal nodes remains under the minimum cross-validated error rate plus the standard error of that error. We also tested the prediction and as the results show in the confusion matrix box below, the overall

accuracy as well as the balanced accuracy including specificity and sensitivity rates are quite high around 95 per cent.

From the figure of the CAP112 tree, we can observe that there are in total 8 patterns of GR1 and 13 patterns of GR3. At the root node, there are 5,547 persons from GR1 and 7,156 persons from GR3 forming together the entire population of the training data. The population of both GR1 and GR3 are then spread across the branches of the tree under the different patterns ending with terminal nodes. As the overall accuracy statistics show we reach to gather 95 per cent of information for both GR1 and GR3 through the discovered patterns. In other words, 95 per cent of people from GR1 are more likely to be detected through the 8 patterns of GR1 and on the other hand, 95 per cent of people from GR3 are more likely to be captured through the 13 patterns of GR3 with our model. The patterns of GR1 are less informative than those of GR3 in terms of disadvantage and/or deficit. The discovered 21 patterns are as follows (starting from the left side of the tree): (1) no deprivation regarding the sanitation (toilet facility) type the household uses (SAN), roof condition of the dwelling (ROFC), nutrition (i.e., household runs out of money to buy food) (NU1), air pollution where the dwelling is located/neighborhood, community (AP), littering situation where the dwelling is located (LIT), having a refrigerator in the dwelling (REF), and the dwelling's main source of energy for cooking (ENG_C); (2) no deprivation regarding SAN, ROFC, NU1, AP, LIT, REF but deprivation regarding ENG_C; (3) no deprivation regarding SAN, ROFC, NU1, AP, LIT and deprivation regarding REF; (4) no deprivation regarding SAN, ROFC, NU1, AP, educational attainment (EDU) but deprivation regarding LIT: (5) no deprivation regarding SAN, ROFC, NU1, AP, the type of main material used for the roof of the dwelling (ROF) but deprivation regarding LIT, and EDU; (6) no deprivation regarding SAN, ROFC, NU1, AP but deprivation regarding LIT, EDU, ROF; (7) no deprivation regarding SAN, ROFC, NU1, medical coverage (MC) but deprivation regarding AP; (8) no deprivation regarding SAN, ROFC, NU1 but deprivation regarding AP, MC, and ROF; (9) no deprivation regarding SAN, ROFC, NU1, ROF, excessive noises where the dwelling is located (NOI) but deprivation regarding AP, and MC; (10) no deprivation regarding SAN, ROFC, NU1, ROF but deprivation regarding AP, MC, and NOI; (11) no deprivation regarding SAN, ROFC but deprivation regarding NU1, and MC; (12) no deprivation regarding SAN, ROFC, MC but deprivation regarding NU1 and nutrition (i.e., cut the size of meals because there was not enough food in the house) (NU2): (13) no deprivation regarding SAN, ROFC, MC, NU2 but deprivation regarding NU1; (14) no deprivation regarding SAN, MC but deprivation regarding ROFC, and NU1; (15) no deprivation regarding SAN, and MC but deprivation regarding ROFC, and access to internet (INT); (16) no deprivation regarding SAN, MC, and INT but deprivation regarding ROFC; (17) no deprivation regarding SAN but deprivation regarding ROFC and MC; (18) deprivation regarding SAN and MC; (19) deprivation regarding SAN, MC, and INT; (20) no deprivation regarding INT but deprivation regarding SAN, MC, and LIT; (21) no deprivation regarding INT, and LIT but deprivation regarding SAN, and MC.

The patterns of GR1 are 1st, 4th, 5th, 7th, 9th, 13th, 16th, and 21st whilst the others are of GR3. The 1st pattern of GR1 does not yield any information on disadvantages and non-achievements, which concerns 75.2 per cent of the population of GR1. The 4th pattern covers 7 per cent of the population of GR1, then it is followed by 7th, 19th, 21st, 13th, 9th, and 5th patterns with 3.8, 2.7, 3, 1.2, 1, and 0.9 per centages respectively. From the most frequent to the least frequent patterns of GR1 are mentioned in the table below: first column indicates the number of pattern, second column informs how many per cent of people classified as GR1 are detected by the given

pattern, third column is the cumulative frequency of the percentages given in the second column, fourth column informs how many per cent of people of GR1 are captured by the given pattern, and last column is the cumulative frequency of the percentages given in the fourth column. In second column, the denominator is 5,248 whilst in fourth column the denominator is 5,547 for GR1. As a recall, 5,547 is the total number of people of GR1 in the training sample. The discovered 8 patterns do not include 299 people from GR1. These people are misclassified in the tree. The second column statistics is only based on the people from GR1 who are correctly classified as GR1 by the patterns in the tree. On the other hand, the fourth column statistics include not only people from GR1 who are correctly classified as GR1 by the tree patterns but also include people from GR1 who are misclassified (classified as GR2) by the tree patterns. In other words, the denominator of the fourth column statistic includes the entire population of GR1.

```
Confusion Matrix and Statistics
          Reference
Prediction  GR1  GR3
       GR1 2603  166
       GR3  171 3413

               Accuracy : 0.947
                 95% CI : (0.9412, 0.9523)
    No Information Rate : 0.5634
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.8922

 Mcnemar's Test P-Value : 0.8275

            Sensitivity : 0.9384
            Specificity : 0.9536
         Pos Pred Value : 0.9401
         Neg Pred Value : 0.9523
             Prevalence : 0.4366
         Detection Rate : 0.4097
   Detection Prevalence : 0.4359
      Balanced Accuracy : 0.9460

       'Positive' Class : GR1
```

CAP112 tree

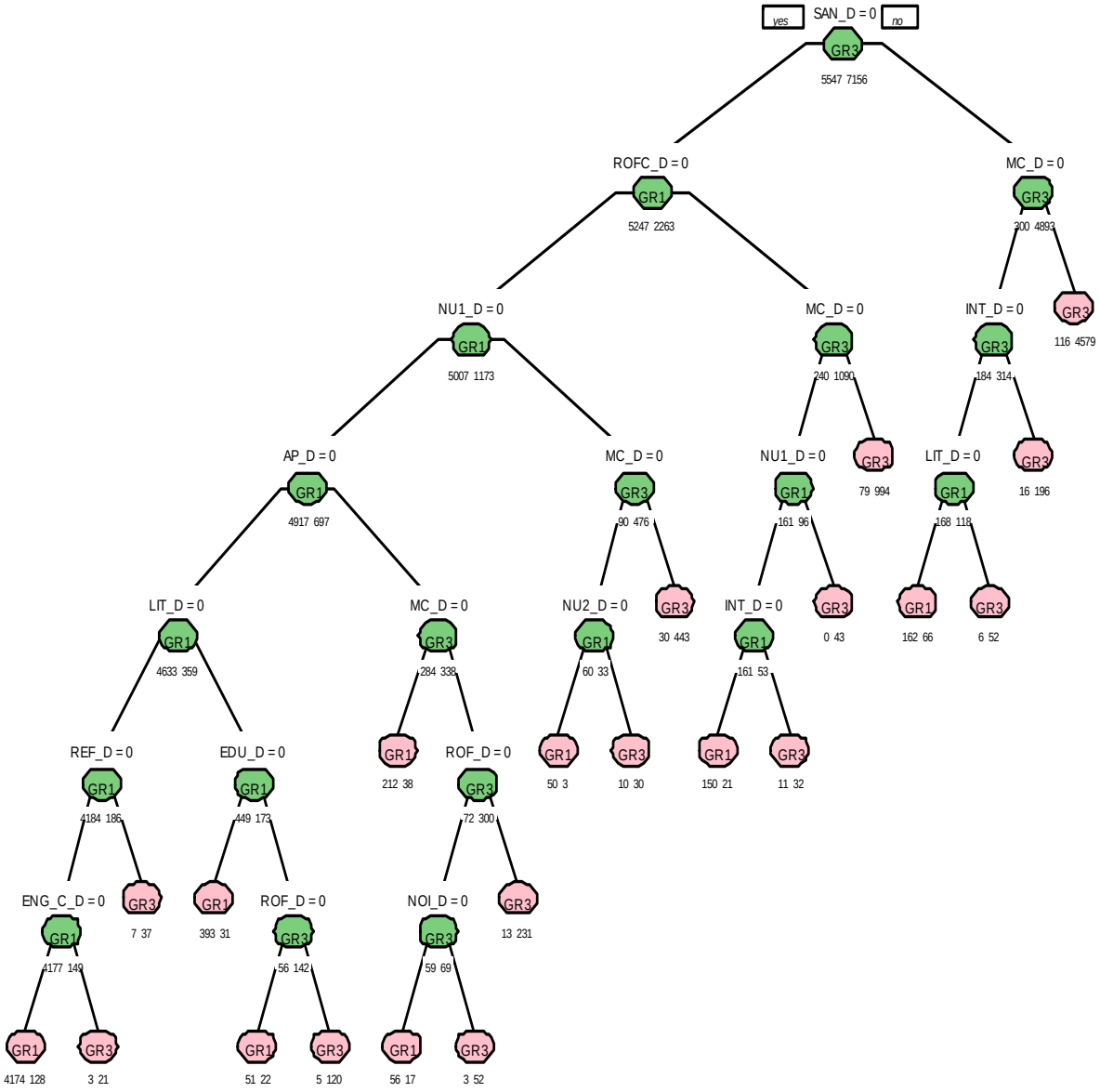


Figure 26. CAP112 tree, data: South Africa GHS 2011

Table 41. Coverage of GR1's patterns among the population classified as GR1 (in per centage), among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP112, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the population classified as GR1 (A)	Cumulative coverage (in %) of (A)	Coverage of pattern (in %) among the total population of GR1 (B)	Cumulative coverage (in %) of (B)
1	79.5	79.5	75.2	75.2
4	7.5	87	7	82.2
7	4	91	3.8	86
16	2.85	93.85	2.7	88.7
21	3	96.85	3	91.7
13	0.95	97.8	1.2	92.9
9	1	98.8	1	93.9
5	1	**99.8~100**	0.9	**94.9~95**

We can also say that through these 8 patterns, deprivation regarding LIT, AP, MC, SAN, ROFC, NU1, and EDU concern 8, 4.8, 4, 3, 3, 1.2, and 1 per cent of the entire population of GR1 (among 5,547 persons) respectively.

Concerning GR3, the 18th pattern covers 64 per cent of the population of GR3, then it is followed by 17th, 11th, 8th, 19th, 6th, 10th, 20th (10th and 20th are equal), 14th, 3rd, 12th, 15th (12th and 15th are equal) and 2nd patterns with 64, 14, 6, 3.2, 2.7, 1.7, 0.7, 0.7, 0.6, 0.5, 0.4, 0.4, and 0.3 per centages respectively. From the most frequent to the least frequent patterns of GR3 are mentioned in the table below. In second column, the denominator is 6830 whilst in fourth column the denominator is 7,156 for GR3. As a recall, 7,156 is the total number of people of GR3 in the training sample. The discovered 13 patterns do not include 326 people from GR2. These people are misclassified in the tree (see the explanations above given for GR1).

Table 42. Coverage of GR3's patterns among the population classified as GR3 (in per centage), among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP112, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the population classified as GR3 (C)	Cumulative coverage (in %) of (C)	Coverage of pattern (in %) among the total population of GR3 (D)	Cumulative coverage (in %) of (D)
18	67	67	64	64
17	14.5	81.5	14	78
11	6.5	88	6	84
8	3.4	91.4	3.2	87.2

19	2.9	94.3	2.7	89.9
6	1.75	96.05	1.7	91.6
10	0.76	96.81	0.7	92.3
20	0.76	97.57	0.7	93
14	0.63	98.2	0.6	93.6
3	0.54	98.74	0.5	94.1
15	0.46	99.2	0.4	94.5
12	0.44	99.64	0.4	94.9
2	0.3	**99.94~100**	0.3	**95.2~95**

For GR3, through the 13 patterns, deprivation regarding MC, SAN, ROFC, NU1, ROF, AP, INT, LIT, EDU, NOI, REF, NU2, and ENG_C concern 91.3, 67.4, 15, 7, 5, 4, 3, 2.4, 1.7, 0.7, 0.5, 0.4, and 0.3 per cent of the entire population of GR3 (among 7156 persons) respectively.

In total, 13 variables emerge as the most influential ones among 33 to distinguish GR1 and GR3. Deprivation regarding LIT, AP, MC, SAN, ROFC, NU1, and EDU are the common for both GR1 and GR3 but in different magnitude.

As we carried out in the analysis of CAP111, we will here also prune the tree for obtaining a more synthetic view for facilitating comparison among the groups. We saw that the initial saturated tree yields 21 patterns with around 95 per cent of accuracy. In fact, the saturated tree is quite acceptable and not very much big size compared to trees with 200 splits. However, we will prune the tree to obtain 12 patterns. For that, the cp= 0.004867496 will be used and the new model is called CAP112.pruned.

```
> dtreeCAP112.pruned=prune(dtreeCAP112, cp=0.004867496)
> printcp(dtreeCAP112.pruned)

Classification tree:
rpart(formula = CAP112 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] AP_D   EDU_D  LIT_D  MC_D   NU1_D  REF_D  ROF_D  ROFC_D SAN_D

Root node error: 5547/12703 = 0.43667

n= 12703

         CP nsplit rel error  xerror      xstd
1 0.5379484      0   1.00000 1.00000 0.0100775
2 0.1532360      1   0.46205 0.46205 0.0081542
3 0.0695872      2   0.30882 0.30882 0.0069401
4 0.0205516      3   0.23923 0.23923 0.0062147
5 0.0117180      5   0.19813 0.19813 0.0057120
6 0.0077519      6   0.18641 0.18641 0.0055560
7 0.0054083      9   0.16315 0.17235 0.0053602
8 0.0052281     10   0.15774 0.16856 0.0053057
9 0.0048675     11   0.15251 0.16820 0.0053005
```

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR3
       GR1 2483  155
       GR3  291 3424

               Accuracy : 0.9298
                 95% CI : (0.9232, 0.936)
    No Information Rate : 0.5634
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.8565

 Mcnemar's Test P-Value : 1.633e-10

            Sensitivity : 0.8951
            Specificity : 0.9567
         Pos Pred Value : 0.9412
         Neg Pred Value : 0.9217
             Prevalence : 0.4366
         Detection Rate : 0.3908
   Detection Prevalence : 0.4152
      Balanced Accuracy : 0.9259

       'Positive' Class : GR1
```

As the CAP112 pruned tree shows (see above), there are 12 terminal nodes with 11 splits in the tree. The 12 discovered patterns are: (1) no deprivation regarding the sanitation (toilet facility) type the household uses (SAN), roof condition of the dwelling (ROFC), nutrition-household runs out of money to buy food (NU1), air pollution where the dwelling is located/neighborhood, community (AP), littering situation where the dwelling is located (LIT), and having a refrigerator in the dwelling (REF); (2) no deprivation regarding SAN, ROFC, NU1, AP, LIT but deprivation regarding REF; (3) no deprivation regarding SAN, ROFC, NU1, AP, educational attainment (EDU) but deprivation regarding LIT; (4) no deprivation regarding SAN, ROFC, NU1, AP, the type of main material used for the roof of the dwelling (ROF) but deprivation regarding LIT, and EDU; (5) no deprivation regarding SAN, ROFC, NU1, AP but deprivation regarding LIT, EDU, and ROF; (6) no deprivation regarding SAN, ROFC, NU1, medical coverage (MC) but deprivation regarding AP; (7) no deprivation regarding SAN, ROFC, NU1 but deprivation regarding AP, and MC; (8) no deprivation regarding SAN, and ROFC but deprivation NU1; (9) deprivation regarding SAN; (10) no deprivation regarding SAN but deprivation regarding ROFC, and MC; (11) no deprivation regarding MC, and NU1 but deprivation regarding ROFC; (12) no deprivation regarding SAN, MC but deprivation regarding ROFC, and NU1.

CAP112 pruned tree

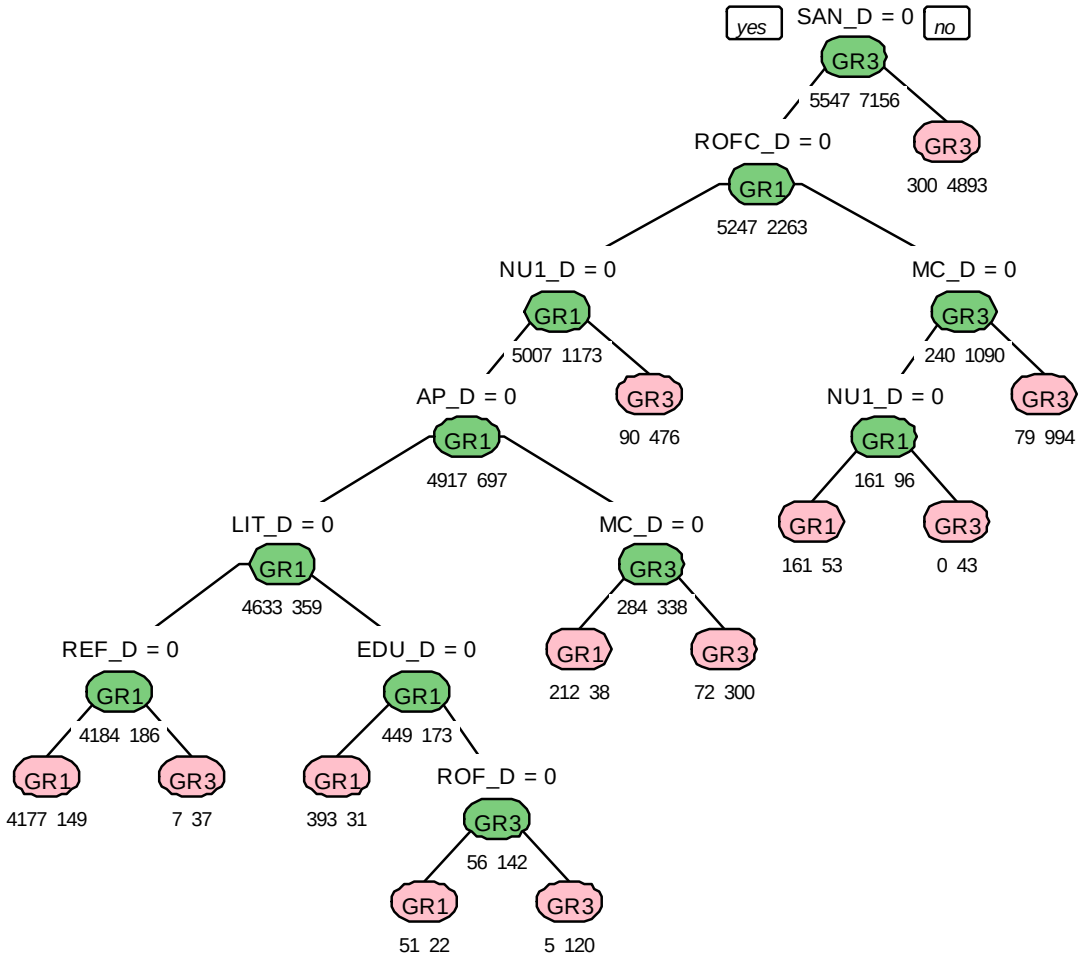


Figure 27. CAP112 pruned tree, data: South Africa GHS 2011

The patterns of GR1 are 1st, 3rd, 6th, 11th, and 4th with 75, 7, 3.8, 2.9, and 0.9 per centage of coverage respectively. For instance, with the 1st pattern, 75 per cent of people from GR1 are detected. On the other hand, the patterns of GR3 are 9th, 10th, 8th, 7th, 5th, 12th, and 2nd with 68.3, 14, 6.65, 4, 1.7, 0.6, and 0.5 per centage of coverage respectively. For instance, with the 9th pattern, 68.3 per cent of people from GR3 are captured. In the training data, by using the CAP112 pruned tree classification patterns, we lose 10.4 per cent of information for GR1 and 4.25 per cent of information for GR3. Compared to CAP112 tree, we lose more information for GR1. Concerning the confusion matrix statistics, the overall accuracy is still very good with 93 per cent and the specificity as well but the sensitivity is decreased as we lose information for GR1.

CAP113[482]: GR1 vs. GR4 analysis

In this analysis, we focus on DEP_GRS1 vs. DEP_GRS4. We use the data containing only people from the poverty group 1 and the poverty group 4. The outcome/target variable is

[482] The meaning of CAP113 is: Capabilities analysis (CAP) of 1st step (1) for 3rd target group (3).

dichotomous, namely CAP113, taking the value 1 for the group 1 and the value 4 for the group 4 (so that people identified in the group 1 take the value 1 and people identified in the group 4 take the value 4 according to CAP113).

Table 43. Training vs. test samples in CAP113 analysis, data: South Africa GHS 2011

	GR1	GR4
Training sample	5547	6146
Test sample	2774	3074

As the table above shows, there are 11,963 individuals in the training data and 5848 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 46.7 per cent for GR1 and 53.3 per cent for GR4 in both training and test samples.

There are 33 variables are the regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used (based on that cp, the tree grows up which means that the last split in the saturated tree should be associated with the cp=0.025), and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree with 14 splits (15 terminal nodes) is associated with 0.0025 complexity parameter. The tree could be even bigger than this, but we determined a cp level equal 0.0025 in the fitting model as the larger tree becomes not efficient to interpret the patterns. The standard error rate of the smallest cross-validation error rate is very few. The cross-validation error rate is also very low even for smaller trees; for instance, a tree with 11 splits has only 5.28 per cent of cross-validation error rate. To see more precisely the evolution of error rates per cp and size of tree, we plot the figure indicating us the 1-SE Rule which permits us also to select the optimal pruned tree. The optimal pruned tree is considered having a cross-validation error rate below the 1-SE Rule line in the graph. We can say that here the optimal level is around with cp=0.0031 corresponding to the tree with 14 splits. However, as the cross-validation error rates are very small for other lower splits and the information earned from the nodes remaining at the lower branches of the tree as the number of splits goes far can be quite small in terms of per centage, we can still prune the tree to around 10-11 splits to obtain a more synthetic information.

```
Call:
rpart(formula = CAP113 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 11693

            CP nsplit  rel error     xerror        xstd
1  0.776996575      0 1.00000000 1.00000000 0.009734289
2  0.111411574      1 0.22300343 0.22300343 0.005995791
3  0.011447629      2 0.11159185 0.11159185 0.004364922
4  0.010726519      4 0.08869659 0.09843158 0.004112957
5  0.007030827      6 0.06724356 0.07030827 0.003500322
6  0.004506941      7 0.06021273 0.06327745 0.003326421
7  0.003845923      8 0.05570579 0.05895078 0.003214079
8  0.003605553     11 0.04416802 0.05282134 0.003046948
9  0.003425275     12 0.04056247 0.04831440 0.002917255
10 0.002884442     13 0.03713719 0.04326663 0.002764039
11 0.002500000     14 0.03425275 0.04110330 0.002695463
```

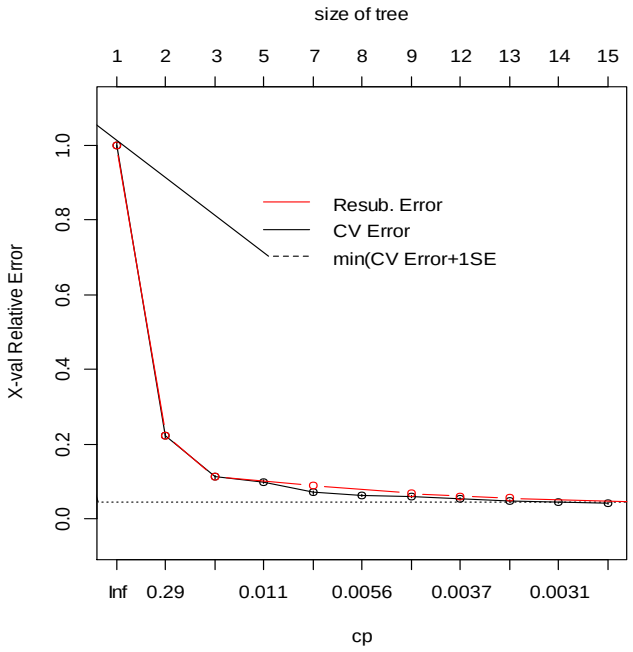


Figure 28. Resubstitution, cross-validated error rates by cp and size of tree and 1-SE Rule for CAP113 saturated tree, data: South Africa GHS 2011

The tree is pruned with a cp equal to 0.0037 to obtain a tree with around 10-12 patterns. The other parameters are kept the same. The information in the box below provides the formula used in the fitting, the variables appearing in the tree, and the classical complexity parameter table introducing the error rates.

```
dtreeCAP113.pruned=prune(dtreeCAP113, cp=0.0037)
> printcp(dtreeCAP113.pruned)

Classification tree:
rpart(formula = CAP113 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] EDU_D   ENG_C_D INT_D   LIT_D   MC_D    NU1_D   ROFC_D  SAN_D   WATD_D

Root node error: 5547/11693 = 0.47439

n= 11693

         CP nsplit rel error   xerror      xstd
1 0.7769966      0  1.000000 1.000000 0.0097343
2 0.1114116      1  0.223003 0.223003 0.0059958
3 0.0114476      2  0.111592 0.111592 0.0043649
4 0.0107265      4  0.088697 0.098432 0.0041130
5 0.0070308      6  0.067244 0.070308 0.0035003
6 0.0045069      7  0.060213 0.063277 0.0033264
7 0.0038459      8  0.055706 0.058951 0.0032141
8 0.0037000     11  0.044168 0.052821 0.0030469
```

As the CAP113 pruned tree shows (see above), there are 12 terminal nodes with 11 splits in the tree. The 12 discovered patterns are: (1) no deprivation regarding the sanitation (toilet facility) type the household uses (SAN), nutrition-household runs out of money to buy food (NU1), roof condition of the dwelling (ROFC), and the dwelling's main energy source for cooking

(ENG_C); (2) no deprivation regarding SAN, NU1, ROFC but deprivation regarding ENG_C; (3) no deprivation regarding SAN, NU1, access to internet (INT) but deprivation regarding ROFC; (4) no deprivation regarding SAN, NU1 but deprivation regarding ROFC, and INT; (5) no deprivation regarding SAN, and medical coverage (MC) but deprivation regarding NU1; (6) no deprivation regarding SAN but deprivation regarding NU1 and MC; (7) no deprivation regarding MC and littering (LIT) but deprivation regarding SAN; (8) no deprivation regarding MC but deprivation regarding SAN and LIT; (9) deprivation regarding SAN, MC, and access to drinking water (WATD); (10) no deprivation regarding WATD but deprivation regarding SAN, MC, and educational attainment (EDU); (11) no deprivation regarding WATD, EDU, ROFC but deprivation regarding SAN and MC; (12) no deprivation regarding WATD and EDU but deprivation regarding SAN, MC, and ROFC.

The patterns of GR1 are 1st, 3rd, 7th, 11th, and 5th with 88.5, 4, 3.2, 1.5, and 1 per centage of coverage respectively (see the table below[483]). For instance, with the 1st pattern, 88.5 per cent of people from GR1 are detected.

Table 44. Coverage of GR1's patterns among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP113 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR1	Cumulative coverage (in %)
1	88.5	88.5
3	4	92.5
7	3.2	95.7
11	1.5	97.2
5	1	98.2~**98**

[483] From CAP113 until the end, we will provide only the percentages according to the total population within the group. The first two columns provided in previous two trees (CAP111 and CAP112) are also informative, but they don't allow to see the misclassification rate. For saving the space, we prefer to keep only these two statistics mentioned in the table above.

CAP113 pruned tree

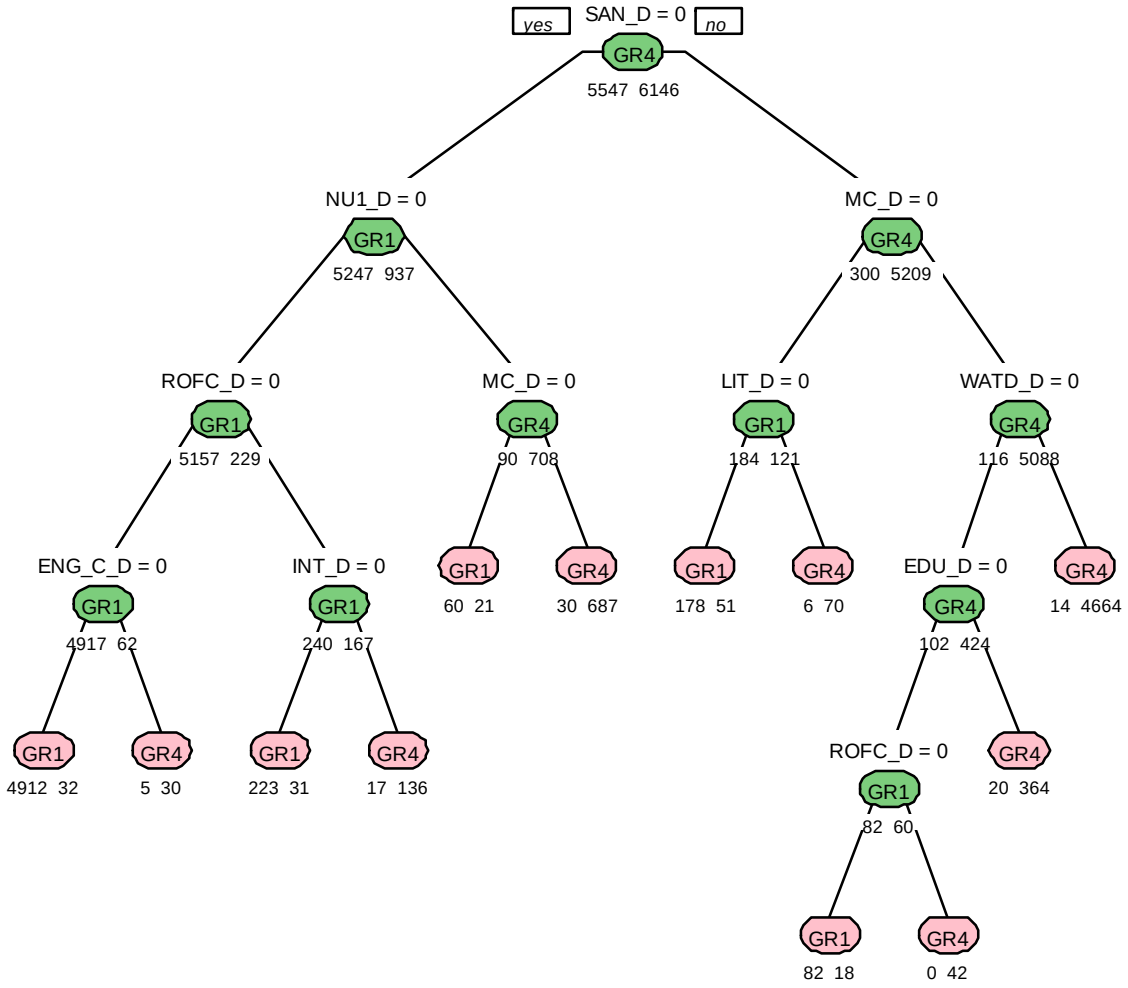


Figure 29. CAP113 pruned tree, data: South Africa GHS 2011

The patterns of GR2 are 9th, 6th, 10th, 4th, 8th, 12th, and 2nd with 76, 11.17, 6, 2.2, 1.13, 0.7 and 0.5 per centage of coverage respectively (see the table below). For instance, with the 9th pattern 76 per cent of people from GR4 are captured.

In the training data, by using the CAP113 pruned tree classification patterns, we lose only 2 per cent of information for both GR1 and GR4 (i.e., 2 per cent of GR1 and GR4 are not included in the pruned tree).

Table 45. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP113 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR4	Cumulative coverage (in %)
9	76	76
6	11.17	87.17
10	6	93.17
4	2.2	95.37
8	1.13	96.5
12	0.7	97.2
2	0.5	97.7~**98**

We can also say that through the 5 patterns of GR1, deprivation regarding SAN, ROFC, MC, and NU1 concern 4.7, 4, 1.5, and 1 per cent of the entire population of GR1 (among 5547 persons) respectively. In terms of deficits and disadvantages, we have less information for GR1 compared to GR4. On the other hand, concerning GR4, the 7 patterns show that the variables MC, SAN, WATD, NU1, EDU, ROFC, INT, LIT, and ENG_C affect 94, 84, 76, 11, 6, 3, 2, 1, 0.5 per cent of the entire population of GR4 (among 6146 persons). The 4 deprivation variables, SAN, MC, NU1, and ROFC, emerging in the patterns of GR1 are the common in terms of deficits for both GR1 and GR4; however, their magnitude are higher for GR4 compared to GR1 (e.g., 88.5 per cent of the population of GR1 are not deprived regarding SAN, and NU1).

```
 Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR4
       GR1 2714   68
       GR4   60 3006

               Accuracy : 0.9781
                 95% CI : (0.974, 0.9817)
    No Information Rate : 0.5256
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.9561

 Mcnemar's Test P-Value : 0.5361

            Sensitivity : 0.9784
            Specificity : 0.9779
         Pos Pred Value : 0.9756
         Neg Pred Value : 0.9804
             Prevalence : 0.4744
         Detection Rate : 0.4641
   Detection Prevalence : 0.4757
      Balanced Accuracy : 0.9781

       'Positive' Class : GR1
```

If we check the accuracy of the prediction via the classification rules obtained on the training data, the confusion matrix statistics mentioned in the box above yield a comprehensive picture on the results of testing. The overall accuracy rate is quite high around 98 per cent. The misclassification rate is around 2 per cent for both GR1 and GR4.

CAP114[484]: GR1 vs. GR5 analysis

In this analysis, we focus on DEP_GRS1 vs. DEP_GRS5. We use the data containing only people from the poverty group 1 and the poverty group 5. The outcome/target variable is dichotomous, namely CAP114, taking the value 1 for the group 1 and the value 5 for the group 5 (so that people identified in the group 1 take the value 1 and people identified in the group 5 take the value 5 according to CAP114).

Table 46. Training vs. test samples in CAP114 analysis, data: South Africa GHS 2011

	GR1	GR5
Training sample	5547	6852
Test sample	2774	3427

As the table above shows, there are 12399 individuals in the training data and 6201individuals in the test data. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 44.7 per cent for GR1 and 55.3 per cent for GR5 in both training and test samples. In the model, there are 33 regressors. Concerning parameters, the gini index for the split, 0.025 as the complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. Given the complexity parameter set at 0.0025, we obtain the biggest tree with 8 splits (9 terminal nodes). The error rate of the smallest cross-validation error rate is 0.0022 and the cross-validation error rate is very low from the 2nd split. We plot cross-validation error rate by cp and size of tree for choosing the optimal pruned tree. We can see that the results of the saturated tree with 8 splits is already optimal regarding 1-SE Rule. Thus, there is no reason to process with pruning.

```
Call:
rpart(formula = CAP114 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 12399

           CP nsplit  rel error     xerror        xstd
1 0.884622318      0 1.00000000 1.00000000 0.009981281
2 0.046872183      1 0.11537768 0.11537768 0.004441439
3 0.025779701      2 0.06850550 0.06850550 0.003459987
4 0.005228051      3 0.04272580 0.04272580 0.002748686
5 0.005047774      4 0.03749775 0.03839913 0.002608368
6 0.002884442      6 0.02740220 0.02776275 0.002223250
7 0.002523887      7 0.02451776 0.02920498 0.002279520
8 0.002500000      8 0.02199387 0.02758248 0.002216110
```

[484] The meaning of CAP114 is: Capabilities analysis (CAP) of 1st step (1) for 4th target group (4).

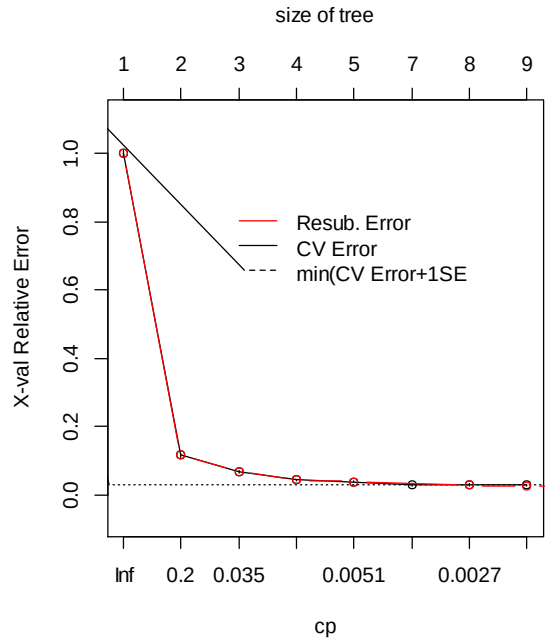


Figure 30. Resubstitution and cross-validation error rates by cp and size of tree and 1-SE Rule for CAP114 saturated tree, data: South Africa GHS 2011

As the CAP114 tree shows (see below), there are 9 terminal nodes with 8 splits in the tree. The 9 discovered patterns are: (1) no deprivation regarding sanitation (toilet facility) type the household uses (SAN), adult nutrition-any adult in the household go hungry because there was not enough food (ADN), and having a refrigerator in the dwelling (REF); (2) no deprivation regarding SAN, ADN but deprivation regarding REF; (3) no deprivation regarding SAN and the main material used for the roof of the dwelling (ROF) but deprivation regarding ADN; (4) no deprivation regarding SAN but deprivation regarding ADN, and ROF; (5) no deprivation regarding medical coverage (MC) and post/mail service (POST) but deprivation regarding SAN; (6) no deprivation regarding MC but deprivation regarding SAN and POST; (7) deprivation regarding SAN, MC, and ROF; (8) no deprivation regarding ROF but deprivation regarding SAN, MC, and access to drinking water (WATD); (9) no deprivation regarding ROF, and WATD but deprivation regarding SAN, and MC. The patterns of GR1 are 1st, 5th, 9th and 3rd with 94, 3.2, 1, and 0.45 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 94 per cent of people from GR1 are detected.

CAP114 tree

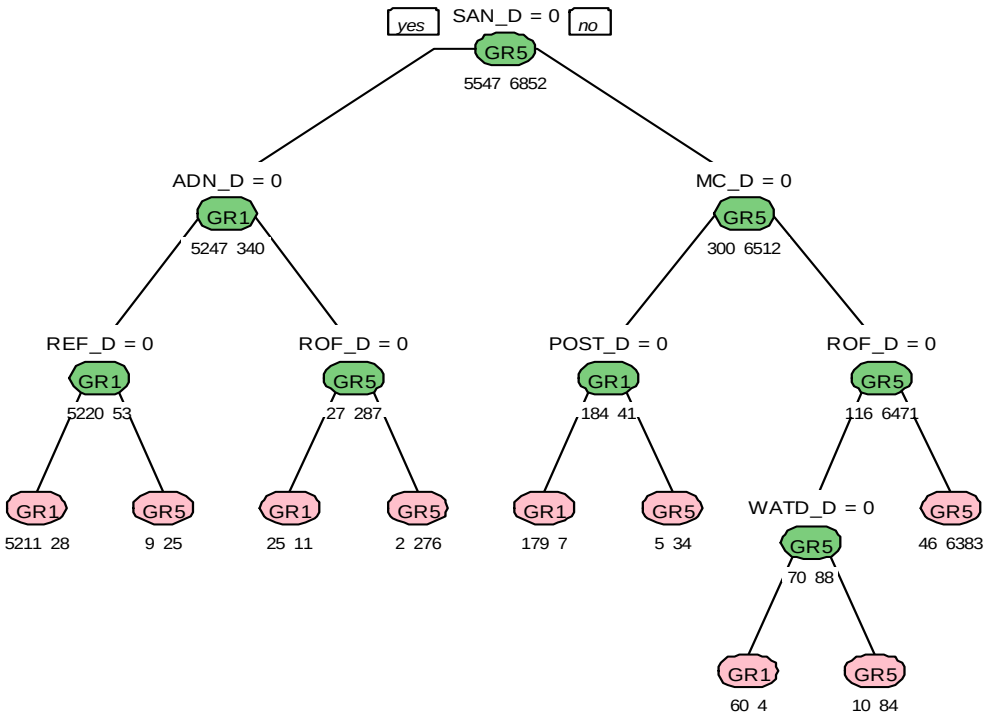


Figure 31. CAP114 tree, data: South Africa GHS 2011

Table 47. Coverage of GR1's patterns among the total population of GR1 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP114 tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR1	Cumulative coverage (in %)
1	94	94
5	3.2	97.2
9	1	98.2
3	0.45	**98.65**

The patterns of GR5 are 7th, 4th, 8th, 6th, and 2nd with 93, 4, 1.2, 0.5 and 0.36 per centage of coverage respectively (see the table below). For instance, with the 7th pattern, 93 per cent of people from GR5 are captured.

In the training data, by using the CAP114 tree classification patterns, we lose only approximatively 1 per cent of information for both GR1 and GR5 (i.e., 1 per cent of GR1 and GR5 are not included in the tree).

Table 48. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP114 tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR5	Cumulative coverage (in %)
7	93	93
4	4	97
8	1.2	98.2
6	0.5	98.7
2	0.36	**99**

We can also say that through the 4 patterns of GR1, deprivation regarding SAN, MC, and ADN concern 4.2, 1, and 0.45 per cent of the entire population of GR1 (among 5547 persons) respectively. In terms of deficits and disadvantages, we have less information for GR1 compared to GR5. On the other hand, concerning GR5, the 5 patterns show that the variables ROF, SAN, MC, AND, WATD, POST, and REF concern 97, 95, 94, 4, 1.2, 0.5, and 0.36 per cent of the entire population of GR5 (among 6852 persons) respectively. The 3 deprivation variables, SAN, MC, and ADN, emerging in the patterns of GR1 are the common in terms of deficits for both GR1 and GR5; however, their magnitude is higher for GR5 compared to GR1.

If we check the accuracy of the prediction via the classification rules obtained on the training data, the confusion matrix statistics mentioned in the box below yield a comprehensive picture on the results of testing. The overall accuracy rate is quite high around 99 per cent. The misclassification rate is around 1 per cent for both GR1 and GR5.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR5
       GR1 2746   33
       GR5   28 3394

               Accuracy : 0.9902
                 95% CI : (0.9874, 0.9925)
    No Information Rate : 0.5527
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.9801

 Mcnemar's Test P-Value : 0.6085

            Sensitivity : 0.9899
            Specificity : 0.9904
         Pos Pred Value : 0.9881
         Neg Pred Value : 0.9918
             Prevalence : 0.4473
         Detection Rate : 0.4428
   Detection Prevalence : 0.4482
      Balanced Accuracy : 0.9901

       'Positive' Class : GR1
```

CAP115[485]: GR2 vs. GR3 analysis

In this analysis, we focus on DEP_GRS2 vs. DEP_GRS3. We use the data containing only people from the poverty group 2 and the poverty group 3. The outcome/target variable is dichotomous, namely CAP115, taking the value 2 for the group 2 and the value 3 for the group 3 (so that people identified in the group 2 take the value 2 and people identified in the group 3 take the value 3 according to CAP115).

Table 49. Training vs. test samples in CAP115 analysis, data: South Africa GHS 2011

	GR2	GR3
Training sample	5022	7156
Test sample	2511	3579

As the table above shows, there are 12178 individuals in the training data and 6090 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 41.2 per cent for GR2 and 58.8 per cent for GR3 in both training and test samples. There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as the complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. Given the complexity parameter set at 0.0025, we obtain the biggest tree with 38 splits (39 terminal nodes). The first tree is obviously very saturated, so that we need to prune it. We plot cross-validation error rate by cp and size of tree for choosing the optimal pruned tree. According to the plot, a tree with around 37 splits has a cross-validated error rate under the 1-SE Rule; however, it is close to the initial saturated tree. From the cp table in the box below, we can see that the cross-validated error rate is around 0.55 for 9-10 splits. It is clear that pruning tree at such level is not the best in terms of information (because accuracy will be lower and the loss of information won't be around 5 per cent) but still we will be able to obtain patterns providing prediction with around 70 and 75 per cent of accuracy. The cross-validation error rate decreases lightly across the splits. In terms of per centage, the loss of information between a cross-validated error equal 0.55 and those equal 0.43 (the smallest possible in the actual saturated tree) is not very substantial so that for facilitating interpretation and avoiding many patterns with small percentages, the saturated tree can be pruned with a cp equal to 0.007865392 (with other parameters kept the same). The second box below shows the cp table of the pruned tree which is composed of 12 splits.

```
Call:
rpart(formula = CAP115 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 12178

           CP nsplit rel error    xerror        xstd
1  0.236160892      0 1.0000000 1.0000000 0.010817050
2  0.083632019      1 0.7638391 0.7638391 0.010207272
3  0.031859817      2 0.6802071 0.6802071 0.009871789
4  0.016129032      4 0.6164875 0.6164875 0.009568118
5  0.013142174      6 0.5842294 0.5842294 0.009397132
6  0.012345679      7 0.5710872 0.5762644 0.009353026
7  0.008761450      8 0.5587415 0.5587415 0.009253271
```

[485] The meaning of CAP115 is: Capabilities analysis (CAP) of 1st step (1) for 5th target group (5).

```
8  0.007964954      9 0.5499801 0.5561529 0.009238212
9  0.007865392     10 0.5420151 0.5525687 0.009217221
10 0.006969335     12 0.5262843 0.5410195 0.009148474
11 0.006571087     13 0.5193150 0.5340502 0.009106155
12 0.005177220     14 0.5127439 0.5073676 0.008938161
13 0.004314350     15 0.5075667 0.4958184 0.008862414
14 0.004181601     20 0.4737157 0.4844683 0.008786123
15 0.004082039     21 0.4695341 0.4776981 0.008739723
16 0.003982477     24 0.4551971 0.4761051 0.008728706
17 0.003285544     25 0.4512147 0.4601752 0.008616440
18 0.003185982     27 0.4446436 0.4504182 0.008545744
19 0.002787734     31 0.4318996 0.4474313 0.008523802
20 0.002721359     33 0.4263242 0.4388690 0.008460106
21 0.002588610     36 0.4181601 0.4358821 0.008437605
22 0.002500000     38 0.4129829 0.4356830 0.008436100
```

```
> summary(dtreeCAP115.pruned)
Call:
rpart(formula = CAP115 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 12178

            CP nsplit rel error    xerror        xstd
1  0.236160892      0 1.0000000 1.0000000 0.010817050
2  0.083632019      1 0.7638391 0.7638391 0.010207272
3  0.031859817      2 0.6802071 0.6802071 0.009871789
4  0.016129032      4 0.6164875 0.6164875 0.009568118
5  0.013142174      6 0.5842294 0.5842294 0.009397132
6  0.012345679      7 0.5710872 0.5762644 0.009353026
7  0.008761450      8 0.5587415 0.5587415 0.009253271
8  0.007964954      9 0.5499801 0.5561529 0.009238212
9  0.007865392     10 0.5420151 0.5525687 0.009217221
10 0.007865392     12 0.5262843 0.5410195 0.009148474
```

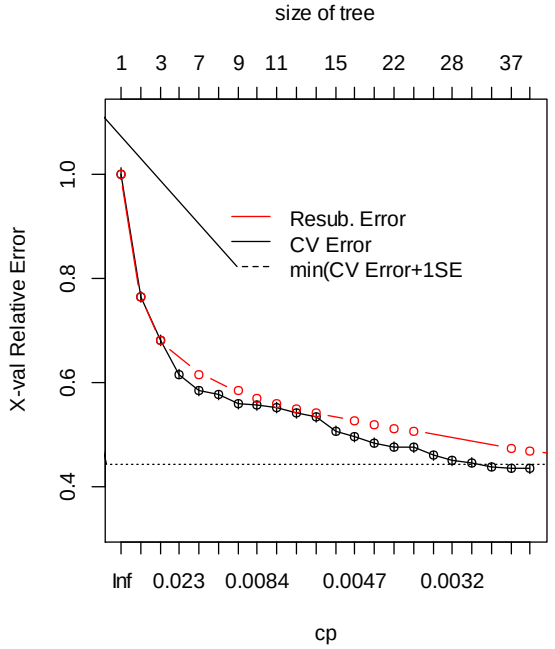


Figure 32. Resubstitution and cross validated error rates by cp and size of tree and 1-SE Rule for CAP115 saturated tree, data: South Africa GHS 2011

As the CAP115 pruned tree shows (see below), there are 13 terminal nodes with 12 splits in the tree. The 13 discovered patterns are: (1) no deprivation regarding the sanitation (toilet facility) type the household uses (SAN) but deprivation regarding nutrition-household runs out of money to buy food (NU1); (2) no deprivation regarding SAN, NU1, condition of wall of the

dwelling (WALC), and air pollution where the dwelling is located/neighborhood, community (AP); (3) no deprivation regarding SAN, NU1, WALC but deprivation regarding AP and water pollution where the dwelling is located/in neighborhood, community (WP); (4) no deprivation regarding SAN, NU1, WALC, WP but deprivation regarding AP; (5) no deprivation regarding SAN, NU1, access to internet (INT), land degradation (LD) but deprivation regarding WALC; (6) no deprivation regarding SAN, NU1, INT but deprivation regarding WALC, and LD; (7) no deprivation regarding SAN, NU1 but deprivation regarding WALC, INT, and the main material used for the roof of the dwelling (ROF); (8) no deprivation regarding SAN, NU1, ROF but deprivation regarding WALC, and INT; (9) deprivation regarding SAN and access to drinking water (WATD); (10) no deprivation regarding WATD but deprivation regarding SAN, and roof condition of the dwelling (ROFC); (11) no deprivation regarding WATD, and ROFC but deprivation regarding SAN, and NU1; (12) no deprivation regarding WATD, ROFC, NU1 but deprivation regarding SAN, and littering situation where the dwelling is located (LIT); (13) no deprivation regarding WATD, ROFC, NU1, LIT but deprivation regarding SAN.

The patterns of GR2 are 2^{nd}, 13^{th}, 5^{th}, 4^{th}, and 8^{th} with 40, 8.8, 7.6, 6.7, 1.7 per centage of coverage respectively (see the table below). For instance, with the 2^{nd} pattern, 40 per cent of people from GR2 are detected. The patterns of GR3 are 9^{th}, 1th, 7^{th}, 10^{th}, 3^{rd}, 6^{th}, 12^{th}, and 11^{th} with 57.5, 9.4, 8.8, 7.6, 5.6, 2.6, 2, 2, and 1.2 per centage of coverage respectively (see the table below). For instance, with the 9^{th} pattern, 57.5 per cent of people from GR3 are captured.

In the training data, by using the CAP115 pruned tree classification patterns, we lose 35.2 per cent of information for GR2 and 3 per cent of information for GR3. The overall accuracy may seem quite good, but it is important to note that the patterns of GR2 are not efficient as those of GR3 and the difference between them is about 32 per centage points.

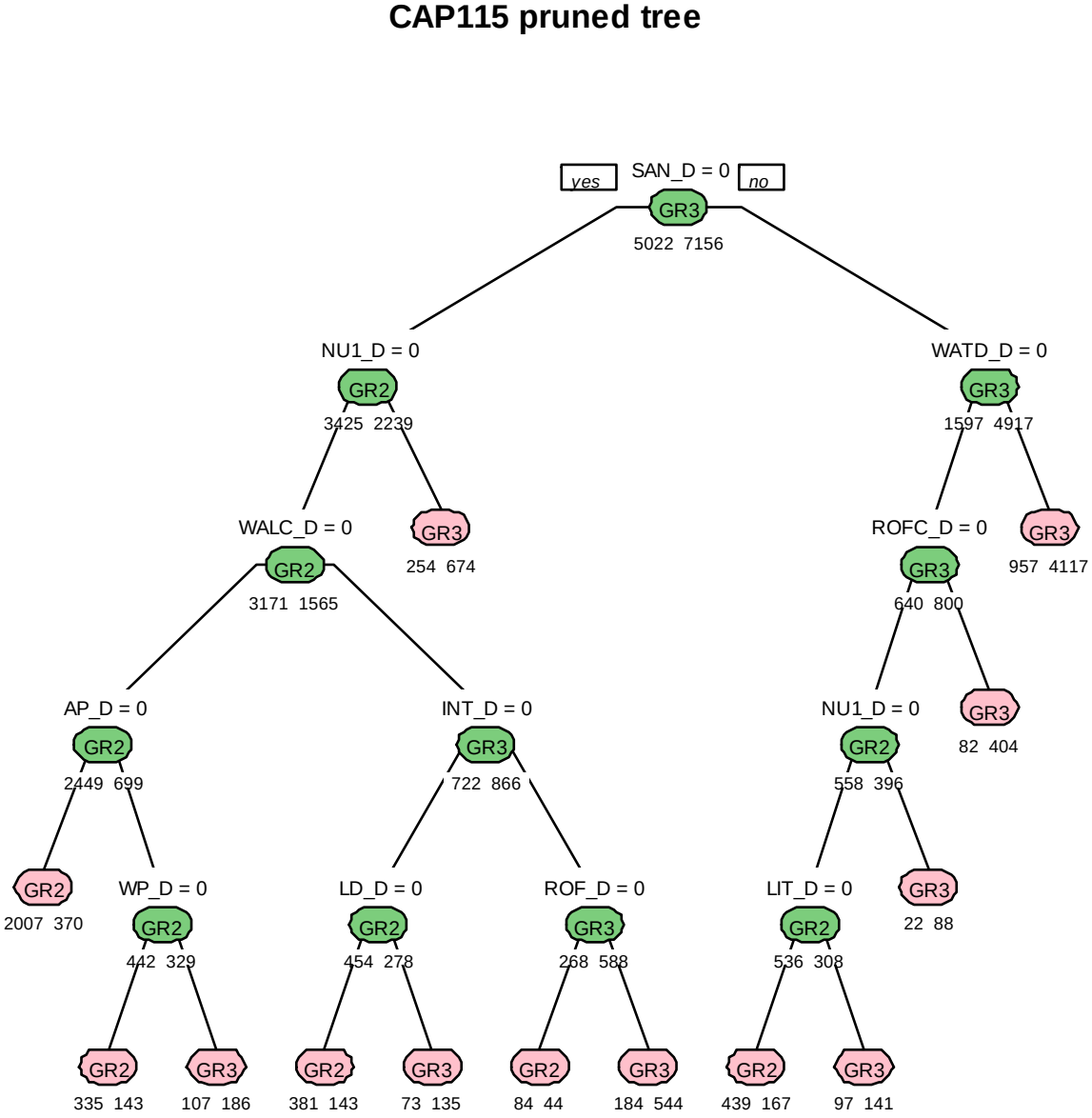


Figure 33. CAP115 pruned tree, data: South Africa GHS 2011

We can also say that through the 5 patterns of GR2, deprivation regarding WALC, SAN, AP, INT concern 9.3, 8.8, 6.7, and 1.7 per cent of the entire population of GR2 (among 5022 persons) respectively. In terms of deficits and disadvantages, we have less information for GR2 compared to GR3. On the other hand, concerning GR3, the 8 patterns show that the variables SAN, WATD, NU1, WALC, INT, ROFC, AP, LD, and LIT target 66.3, 57.5, 10.6, 9.6, 7.6, 5.6, 2.6, 2, and 2 per cent of the entire population of GR3 (among 7156 persons). The 3 deprivation variables, WALC, SAN, and AP, emerging in the patterns of GR2 are the common in terms of deficits for both GR2 and GR3. However, the magnitude of the SAN is very higher for GR3 compared to GR2. But also, the AP appears with a higher influence on GR2 compared to GR3.

Table 50. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP115 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR2	Cumulative coverage (in %)
2	40	40
13	8.8	48.8
5	7.6	56.4
4	6.7	63.1
8	1.7	**64.8**

Table 51. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP115 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR3	Cumulative coverage (in %)
9	57.5	57.5
1	9.4	66.9
13	8.8	75.7
7	7.6	83.3
10	5.6	88.9
3	2.6	91.5
6	2	93.5
12	2	95.5
11	1.2	**96.7~97**

Concerning the confusion matrix statistics, the overall accuracy rate is quite good around 78 per cent but the attention to the sensitivity rate in the sense that the accuracy is lower in prediction of GR2.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR2  GR3
       GR2 1669  459
       GR3  842 3120

               Accuracy : 0.7864
                 95% CI : (0.7759, 0.7966)
    No Information Rate : 0.5877
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5489

 Mcnemar's Test P-Value : < 2.2e-16
```

```
            Sensitivity : 0.6647
            Specificity : 0.8718
         Pos Pred Value : 0.7843
         Neg Pred Value : 0.7875
             Prevalence : 0.4123
         Detection Rate : 0.2741
   Detection Prevalence : 0.3494
      Balanced Accuracy : 0.7682

       'Positive' Class : GR2
```

CAP116[486]: GR2 vs. GR4 analysis

In this analysis, we focus on DEP_GRS2 vs. DEP_GRS4. We use the data containing only people from the poverty group 2 and the poverty group 4. The outcome/target variable is dichotomous, namely CAP116, taking the value 2 for the group 2 and the value 4 for the group 4 (so that people identified in the group 2 take the value 2 and people identified in the group 4 take the value 4 according to CAP116).

Table 52. Training vs. test samples in CAP116 analysis, data: South Africa GHS 2011

	GR2	GR4
Training sample	5022	6146
Test sample	2511	3074

As the table above shows, there are 11168 individuals in the training data and 5585 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 45 per cent for GR2 and 55 per cent for GR4 in both training and test samples. There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree with 25 splits (26 terminal nodes) is associated with 0.0025 complexity parameter. The cross-validation error rates are quite acceptable from the ninth split. For instance, the tree with 11 splits has 18 per cent of cross-validation error rate. We can say that here the optimal level is around 0.0028 cp corresponding to a tree with around 25 splits. However, it is quite big with such cp level. So that, we will prune the tree to around 11 splits as in previous trees.

```
Call:
rpart(formula = CAP116 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 11168

           CP nsplit rel error    xerror        xstd
1  0.496814018      0 1.0000000 1.0000000 0.010468157
2  0.087017125      1 0.5031860 0.5031860 0.008804815
3  0.051042082      2 0.4161689 0.4161689 0.008207366
4  0.033253684      5 0.2630426 0.3243728 0.007427584
5  0.016925528      6 0.2297889 0.2694146 0.006866409
6  0.013440860      7 0.2128634 0.2502987 0.006650617
7  0.009956193      9 0.1859817 0.1943449 0.005942788
8  0.005973716     10 0.1760255 0.1818001 0.005765524
```

[486] The meaning of CAP116 is: Capabilities analysis (CAP) of 1st step (1) for 6th target group (6).

```
9  0.005774592     11 0.1700518 0.1798088 0.005736658
10 0.004978096     12 0.1642772 0.1778176 0.005707583
11 0.004579849     13 0.1592991 0.1712465 0.005610118
12 0.004032258     15 0.1501394 0.1658702 0.005528577
13 0.003185982     19 0.1340104 0.1439665 0.005177963
14 0.003086420     20 0.1308244 0.1403823 0.005117505
15 0.002588610     23 0.1206691 0.1294305 0.004926731
16 0.002500000     25 0.1154918 0.1254480 0.004854952

> printcp(dtreeCAP116.pruned)

Classification tree:
rpart(formula = CAP116 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
 [1] ADN_D   AP_D    ENG_C_D NU1_D   POST_D  ROFC_D  SAN_D   WAL_D   WALC_D
[10] WATD_D

Root node error: 5022/11168 = 0.44968

n= 11168

         CP nsplit rel error  xerror      xstd
1 0.4968140      0   1.00000 1.00000 0.0104682
2 0.0870171      1   0.50319 0.50319 0.0088048
3 0.0510421      2   0.41617 0.41617 0.0082074
4 0.0332537      5   0.26304 0.32437 0.0074276
5 0.0169255      6   0.22979 0.26941 0.0068664
6 0.0134409      7   0.21286 0.25030 0.0066506
7 0.0099562      9   0.18598 0.19434 0.0059428
8 0.0059737     10   0.17603 0.18180 0.0057655
9 0.0057746     11   0.17005 0.17981 0.0057367
```

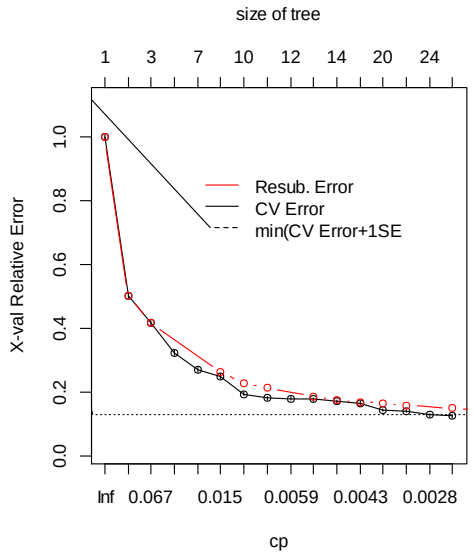


Figure 34. Resubstitution error and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP116 saturated tree, data: South Africa GHS 2011

We prune the tree with a cp equal 0.005774592 giving us a tree with 11 splits. As the CAP116 pruned tree shows (see below), there are 12 terminal nodes in the tree. The 12 discovered patterns are: (1) no deprivation regarding sanitation (toilet facility) type the household uses (SAN), and nutrition-household runs out of money to buy food (NU1); (2) no deprivation regarding SAN but deprivation regarding NU1 and adult nutrition-any adult in the household

go hungry because there was not enough food (ADN); (3) no deprivation regarding SAN, and ADN but deprivation regarding NU1 and wall condition of the dwelling (WALC); (4) no deprivation regarding SAN, ADN, and WALC but deprivation regarding NU1; (5) deprivation regarding SAN and post services (POST); (6) no deprivation regarding POST but deprivation regarding SAN and ROFC; (7) no deprivation regarding POST and ROFC but deprivation regarding SAN, and NU1; (8) no deprivation regarding POST, ROFC , and NU1 but deprivation regarding SAN, and air pollution where the dwelling is located/neighborhood, community(AP), and access to drinking water (WATD); (9) no deprivation regarding POST, ROFC, NU1, and WATD but deprivation regarding SAN, and APD;(10) no deprivation regarding POST, ROFC, NU1, AP but deprivation regarding SAN, and dwelling's main source of energy for cooking (ENG_C); (11) no deprivation regarding POST, ROFC, NU1, AP, ENG_C but deprivation regarding SAN, and the main material used for the wall of the dwelling (WAL); (12) no deprivation regarding POST, ROFC, NU1, AP, ENG_C, WAL but deprivation regarding SAN.

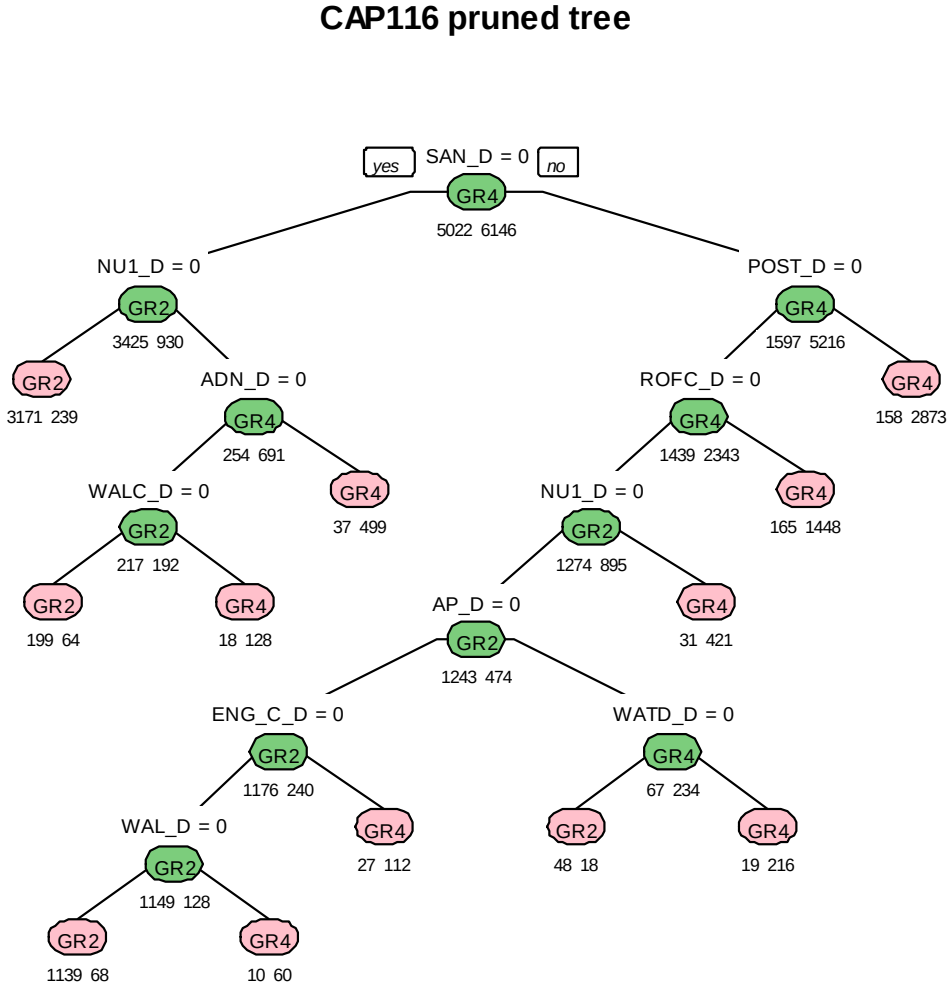


Figure 35. CAP116 pruned tree, data: South Africa GHS 2011

The patterns of GR2 are 1st, 4th,12th, and 9th with 63, 4, 2.8, and 0.95 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 63 per cent of people from GR2 are detected. On the other hand, the patterns of GR4 are 5th, 6th, 2nd, 7th, 8th , 3rd , 10th , and 11th with 47, 23.5, 8, 6.8, 3.5, 2, 1.8, and 1 per centage of coverage respectively (see the table below). For instance, with the 5th pattern, 47 per cent of people from GR4 are captured.

Table 53. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP116 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR2	Cumulative coverage (in %)
1	63	63
4	4	67
11	2.8	69.8
9	0.95	70.75~**71**

Table 54. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP116 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR4	Cumulative coverage (in %)
5	47	47
6	23.5	70.5
2	8	78.5
7	6.8	85.3
8	3.5	88.8
3	2	90.8
10	1.8	92.6
11	1	93.6~**94**

In the training data, by using the CAP116 pruned tree classification patterns, we lose 25 per cent of information for GR2 and 6 per cent of information for GR4.

We can also say that through the 4 patterns of GR2, deprivation regarding NU1, SAN, and AP concern 4, 2.8, and 0.95 per cent of the entire population of GR2 (among 5022 persons) respectively. In terms of deficits and disadvantages, we have less information for GR2 compared to GR4. On the other hand, concerning GR4, the 8 patterns show that the variables SAN, POST, ROFC, NU1, AND, AP, WATD, WALC, ENG_C, WAL affect 83.6, 47, 23.5, 16.8, 8, 3.5, 3.5, 2, 1.8, and 1per cent of the entire population of GR4 (among 6146 persons). The 3 deprivation variables, NU1, SAN, and AP are the common in terms of deficits for both GR2 and GR4; however, their magnitude is higher for GR4 compared to GR2.

Moreover, regarding the confusion matrix statistics, the overall accuracy rate is quite high around 92.7 per cent. The misclassification rate is 8.5 per cent for GR2 and, 6.2 per cent for GR4.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR2  GR4
       GR2 2296  192
       GR4  215 2882

               Accuracy : 0.9271
                 95% CI : (0.92, 0.9338)
    No Information Rate : 0.5504
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.8526

 Mcnemar's Test P-Value : 0.2755

            Sensitivity : 0.9144
            Specificity : 0.9375
         Pos Pred Value : 0.9228
         Neg Pred Value : 0.9306
             Prevalence : 0.4496
         Detection Rate : 0.4111
   Detection Prevalence : 0.4455
      Balanced Accuracy : 0.9260

       'Positive' Class : GR2
```

CAP117[487]: GR2 vs. GR5 analysis

In this analysis, we focus on DEP_GRS2 vs. DEP_GRS5. We use the data containing only people from the poverty group 2 and the poverty group 5. The outcome/target response variable is dichotomous variable, namely CAP117, taking the value 2 for the group 2 and the value 5 for the group 5 (so that people identified in the group 2 take the value 2 and people identified in the group 5 take the value 5 according to CAP117).

As the table below shows, there are 11874 individuals in the training data and 5938 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 42.3 per cent for GR2 and 57.7 per cent for GR5 in both training and test samples.

Table 55. Training vs. test samples in CAP117 analysis, data: South Africa GHS 2011

	GR2	GR5
Training sample	5022	6852
Test sample	2511	3427

There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree with 13 splits (14 terminal nodes) is associated with 0.0025 complexity parameter. The standard error of the smallest cross-validation error rate is quite low. The cross-validation error rate is also low even for smaller trees; for instance, a tree with 11 splits has only 6.6 per cent of cross-validation error rate. Based on the plot of cross-validation error and 1-SE Rule, the optimal pruned tree emerges with around

[487] The meaning of CAP117 is: Capabilities analysis (CAP) of 1st step (1) for 7th target group (7).

0.0053 cp level corresponding to between 9 and 11 splits. The sum of smallest cross-validated error rate and its standard error gives 0.067846694. According to that number, the closest cross-validation error rate is 0.0667 corresponding to the tree with 11 splits. Thus, we prune the tree with a cp equal 0.003086420 keeping other parameters in the model the same. The information in the box below provides the formula used in fitting, the variables appearing in the tree, and the classical complexity parameter table introducing the error rates.

```
Call:
rpart(formula = CAP117 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 11874

            CP nsplit  rel error     xerror        xstd
1  0.614894464      0 1.00000000 1.00000000 0.010719433
2  0.110911987      1 0.38510554 0.38510554 0.008012097
3  0.074273198      2 0.27419355 0.27419355 0.006947433
4  0.033651932      3 0.19992035 0.19992035 0.006036795
5  0.021804062      4 0.16626842 0.16626842 0.005547954
6  0.016726404      6 0.12266029 0.12704102 0.004892614
7  0.013341298      7 0.10593389 0.10911987 0.004552535
8  0.008960573      8 0.09259259 0.09040223 0.004160885
9  0.003086420     11 0.06531262 0.06670649 0.003592785
10 0.002500000     13 0.05913978 0.06431701 0.003529684
```

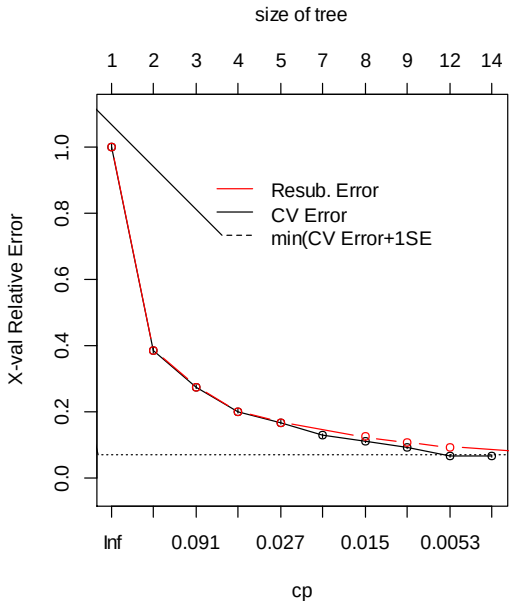


Figure 36. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP117 saturated tree, data: South Africa GHS 2011

```
> dtreeCAP117.pruned=prune(dtreeCAP117, cp=0.003086420)
> summary(dtreeCAP117.pruned)
Call:
rpart(formula = CAP117 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 11874

           CP nsplit  rel error     xerror        xstd
1 0.614894464      0 1.00000000 1.00000000 0.010719433
2 0.110911987      1 0.38510554 0.38510554 0.008012097
3 0.074273198      2 0.27419355 0.27419355 0.006947433
4 0.033651932      3 0.19992035 0.19992035 0.006036795
5 0.021804062      4 0.16626842 0.16626842 0.005547954
6 0.016726404      6 0.12266029 0.12704102 0.004892614
```

```
7 0.013341298      7 0.10593389 0.10911987 0.004552535
8 0.008960573      8 0.09259259 0.09040223 0.004160885
9 0.003086420     11 0.06531262 0.06670649 0.003592785
```

As the CAP117 pruned tree shows (see below), there are 12 terminal nodes with 11 splits in the tree. The 12 discovered patterns are: (1) no deprivation regarding sanitation (toilet facility) type the household uses (SAN), and adult nutrition-any adult in the household go hungry because there was not enough food (ADN); (2) no deprivation regarding SAN, and nutrition-eat a smaller variety of foods because there was not enough food in the house (NU4); (3) no deprivation regarding SAN but deprivation regarding ADN, and NU4; (4) no deprivation regarding dwelling's main energy source for heating (ENG_H) but deprivation regarding SAN and nutrition-household runs out of money to buy food (NU1); (5) no deprivation regarding ENG_H, NU1, and the main material used for the wall of the dwelling (WAL) but deprivation regarding SAN: (6) no deprivation regarding ENG_H, and NU1 but deprivation regarding SAN, and WAL; (7) no deprivation regarding NU1, and access to drinking water (WATD) but deprivation regarding SAN, and ENG_H; (8) no deprivation regarding WATD but deprivation regarding SAN, ENG_H, and NU1; (9) deprivation regarding SAN, ENG_H, WATD, and access to internet (INT); (10) no deprivation regarding INT but deprivation regarding SAN, ENG_H, WATD, and condition of the walls of the dwelling (WALC); (11) no deprivation regarding INT, WALC, and NU1 but deprivation regarding SAN, ENG_H, and WATD; (12) no deprivation regarding INT, WALC but deprivation regarding SAN, ENG_H, WATD, and NU1.

The patterns of GR2 are 1st, 5th, 7th, 11th, and 2nd with 66, 21, 5, 3, 1.7 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 66 per cent of people from GR2 are detected. On the other hand, the patterns of GR5 are 9th, 4th, 3rd, 10th, 8th, 6th, and 12th with 79, 5.85, 3.9, 3.6, 3.4, 1.35, and 0.7 per centage of coverage respectively (see the table below). For instance, with the 9th pattern, 79 per cent of people from GR5 are captured. In the training data, by using the CAP117 pruned tree classification patterns, we lose only 3 per cent of information for GR2, and 2 per cent of information for GR5.

Table 56. Coverage of GR2's patterns among the total population of GR2 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP117 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR2	Cumulative coverage (in %)
1	66	66
5	21	87
7	5	92
11	3	95
2	1.7	96.7~**97**

Table 57. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP117 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR5	Cumulative coverage (in %)
9	79	79
4	5.85	84.85
3	3.9	88.75
10	3.6	92.35
8	3.4	95.75
6	1.35	97.1
12	0.7	97.8~**98**

Moreover, through the 5 patterns of GR2, deprivation regarding SAN, ENG_H, WATD, and ADN concern 29, 8,3, and 1.7 per cent of the entire population of GR2 (among 5022 persons) respectively. On the other hand, concerning GR5, the 7 patterns show that the variables SAN, ENG:H, WALD, INT, NU1, WAT, ADN, NU4, and WALC affect 94, 87, 80, 79, 10, 4.3, 3.9, and 3.6 per cent of the entire population of GR5 (among 6852 persons). The deprivation regarding SAN, ENG_H, WATD, and ADN are the common problems for both GR2 and GR5. However, the deprivation cover of the variables SAN, and ENG_H are much higher in GR5. In terms of confusion matrix statistics, the overall accuracy rate is quite high around 97 per cent. The misclassification rate is around 23.5 per cent for GR2 and around 2 per cent for GR5.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR2  GR5
       GR2 2422   66
       GR5   89 3361

               Accuracy : 0.9739
                 95% CI : (0.9695, 0.9778)
    No Information Rate : 0.5771
    P-Value [Acc > NIR] : < 2e-16

                  Kappa : 0.9465

 Mcnemar's Test P-Value : 0.07721

            Sensitivity : 0.9646
            Specificity : 0.9807
         Pos Pred Value : 0.9735
         Neg Pred Value : 0.9742
             Prevalence : 0.4229
         Detection Rate : 0.4079
   Detection Prevalence : 0.4190
      Balanced Accuracy : 0.9726

       'Positive' Class : GR2
```

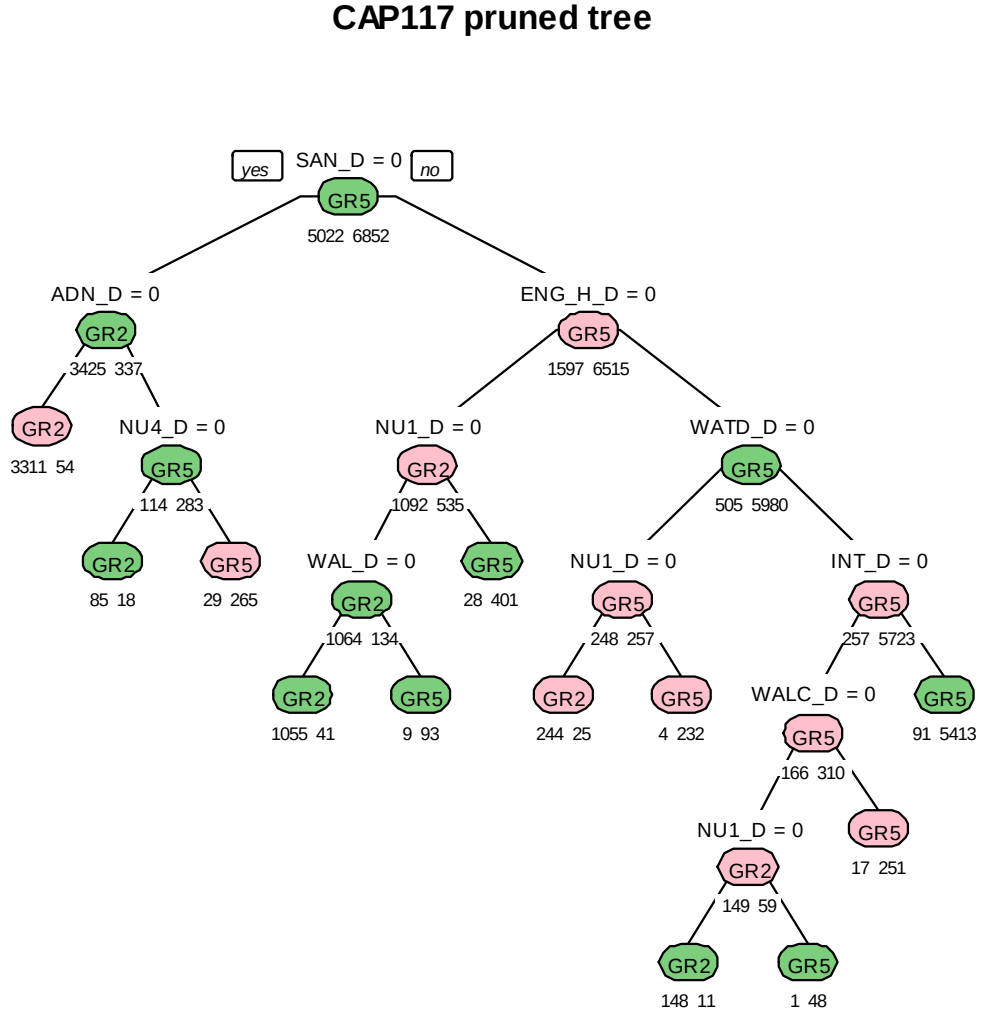


Figure 37. CAP117 pruned tree, data: South Africa GHS 2011

CAP118[488]: GR3 vs. GR4 analysis

In this analysis, we focus on DEP_GRS3 vs. DEP_GRS4. We use the data containing only people from the poverty group 3 and the poverty group 4. The outcome/target variable is dichotomous, namely CAP118, taking the value 3 for the group 3 and the value 4 for the group 4 (so that people identified in the group 3 take the value 3 and people identified in the group 4 take the value 4 according to CAP118).

Table 58. Training vs. test samples in CAP118 analysis, data: South Africa GHS 2011

	GR3	GR4
Training sample	7156	6146
Test sample	3579	3074

As the table above shows, there are 13302 individuals in the training data and 6653 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite

[488] The meaning of CAP118 is: Capabilities analysis (CAP) of 1st step (1) for 8th target group (8).

balanced with 53.8 per cent for GR3 and 46.2 per cent for GR4 in both training and test samples. There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree with 31 splits (32 terminal nodes) is associated with 0.0025 complexity parameter. The cross-validation error rate does not decrease significantly as the tree grows up. With 31 splits, the cross-validation error rate of the tree is equal to 0.4331272. We will proceed with pruning phase, but it is important to note that the cross-validation error rate won't be very small. According to 1-SE Rule, the tree can be pruned to around 25 splits; however, it will be still generating a substantial number of patterns. As in the previous trees, we will set the tree for around 9-11 splits (by keeping other parameters the same in the model), which gives us a cp around 0.009 according to the cp table below.

```
> summary(dtreeCAP118)#print a summary which examines each node in depth.
Call:
rpart(formula = CAP118 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 13302

            CP nsplit rel error    xerror        xstd
1  0.138789457      0 1.0000000 1.0000000 0.009355789
2  0.077774162      2 0.7224211 0.7557761 0.008945931
3  0.014643671      3 0.6446469 0.6836967 0.008723639
4  0.013423365      8 0.5650830 0.6065734 0.008428177
5  0.009030264     10 0.5382363 0.5589001 0.008213058
6  0.008867556     14 0.5021152 0.5525545 0.008182424
7  0.007321835     16 0.4843801 0.5405142 0.008122955
8  0.006671006     18 0.4697364 0.5095997 0.007961966
9  0.004393101     20 0.4563944 0.4866580 0.007834433
10 0.003796507     21 0.4520013 0.4648552 0.007706520
11 0.003579564     24 0.4406118 0.4503742 0.007617784
12 0.003498210     25 0.4370322 0.4469574 0.007596393
13 0.003254149     27 0.4300358 0.4435405 0.007574825
14 0.002684673     29 0.4235275 0.4355678 0.007523807
15 0.002603319     31 0.4181582 0.4331272 0.007507992
```

As the CAP118 pruned tree shows (see below), there are 15 terminal nodes with 14 splits in the tree. The 15 discovered patterns are: (1) no deprivation regarding the main material of walls of the dwelling (WAL), nutrition-eat a smaller variety of foods because there was not enough food in the house (NU4), dwelling's main energy source for cooking (ENG_C), having a refrigerator in the dwelling (REF), and land degradation LD); (2) no deprivation regarding WAL, NU4, ENG_C, REF, and post services (POST) but no deprivation regarding LD; (3) no deprivation regarding WAL, NU4, ENG_C, REF but regarding LD, POST, and the condition of the dwelling's walls (WALC); (4) no deprivation regarding WAL, NU4, ENG_C, REF, WALC, littering (LIT) but deprivation regarding LD, POST; (5) no deprivation regarding WAL, NU4, ENG_C, REF, WALC but deprivation regarding LD, POST, and LIT; (6) no deprivation regarding WAL, NU4, ENG_C, rubbish collection or removement where the dwelling is located/neighborhood, community (RUB) but deprivation regarding REF; (7) no deprivation regarding WAL, NU4, ENG_C but deprivation regarding REF, and RUB; (8) no deprivation regarding WAL, NU4 but deprivation regarding ENG_C, and WAL_C; (9) no deprivation regarding WAL, NU4, WALC but deprivation regarding ENG_C, and REF; (10) no deprivation regarding WAL, NU4, WALC, REF but deprivation regarding ENG_C; (11) no deprivation regarding WAL, NU4, WALC, REF but deprivation regarding ENG_C, and LD; (12)

deprivation regarding WAL; (13) no deprivation regarding WAL but deprivation regarding NU4, and sanitation (toilet facility) type the household uses (SAN); (14) no deprivation regarding WAL, SAN, and the main material used for the roof of the dwelling (ROF) but deprivation regarding NU4; (15) no deprivation regarding WAL, and SAN but deprivation regarding NU4, and ROF. The patterns of GR3 are 1st, 2nd, 10th, 6th, 14th, and 4th with 44, 17.5, 7.6, 3.6, 3.3, and 2.45 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 44 per cent of people from GR3 are detected. On the other hand, the patterns of GR4 are 12th, 13th, 15th, 8th, 7th, 11th, 9th, 3rd, and 5th with 20, 17, 9, 9, 4.8, 4.8, 3.4, 3.3., and 2.6 per centage of coverage respectively (see the table below). For instance, with the 12th pattern, 20 per cent of people from GR4 are captured.

```
> summary(dtreeCAP118.pruned)
Call:
rpart(formula = CAP118 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 13302
          CP nsplit rel error    xerror        xstd
1 0.138789457      0 1.0000000 1.0000000 0.009355789
2 0.077774162      2 0.7224211 0.7557761 0.008945931
3 0.014643671      3 0.6446469 0.6836967 0.008723639
4 0.013423365      8 0.5650830 0.6065734 0.008428177
5 0.009030264     10 0.5382363 0.5589001 0.008213058
6 0.009000000     14 0.5021152 0.5525545 0.008182424
```

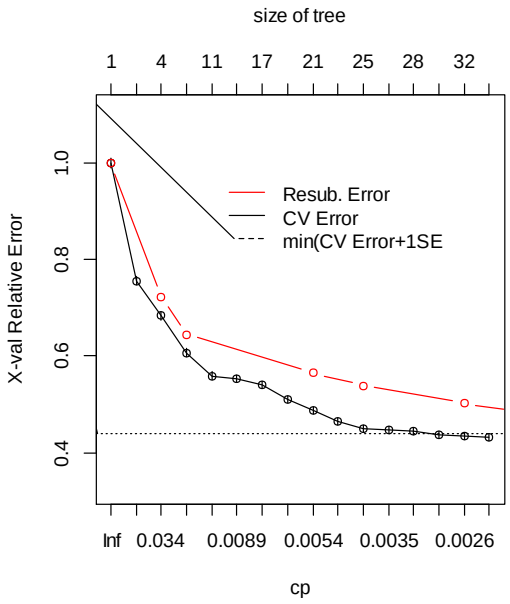


Figure 38. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP118 saturated tree, data: South Africa GHS 2011

Table 59. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP118 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR3	Cumulative coverage (in %)
1	44	44
2	17.5	61.5
10	7.6	69.1
6	3.6	72.7
14	3.3	76
4	2.45	78.45~**78.5**

Table 60. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP118 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR4	Cumulative coverage (in %)
12	20	20
13	17	37
15	9	46
8	9	55
7	4.8	59.8
11	4.8	64.6
9	3.4	68
3	3.3	71.3
5	2.6	73.9~**74**

In the training data, by using the CAP118 pruned tree classification patterns, we lose 21.5 per cent of information for GR3 and 26 per cent of information for GR4.

Furthermore, through the 5 patterns of GR3, deprivation regarding LD, ENG_C, REF, NU4, and POST concern 20, 7.6, 3.6, 3.3, and 2.45 per cent of the entire population of GR3 (among 7156 persons) respectively. On the other hand, concerning GR4, the 9 patterns show that the variables NU4, WAL, SAN, ENG_C, WALC, LD, ROF, POST, RUB, and LIT affect 26, 20, 17, 16.7, 12.3, 10.7, 9, 8.2, 6, 4.8, and 2.6 per cent of the entire population of GR4 (among 6146 persons). The deprivation regarding LD, POST, REF, ENG_C, and NU4 are the common for both GR3 and GR4; however, concerning LD its magnitude is two times higher but other variables cover more people in GR4 such as the influence of NU4 is 8 times higher in GR4.

Concerning confusion matrix statistics, the overall accuracy rate is good around 76 per cent but it is important to note that the misclassification rate is 21.5 per cent for GR3 and 26 per cent for GR4.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR3  GR4
       GR3 2809  803
       GR4  770 2271

               Accuracy : 0.7636
                 95% CI : (0.7532, 0.7737)
    No Information Rate : 0.538
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.524

 Mcnemar's Test P-Value : 0.4198

            Sensitivity : 0.7849
            Specificity : 0.7388
         Pos Pred Value : 0.7777
         Neg Pred Value : 0.7468
             Prevalence : 0.5380
         Detection Rate : 0.4222
   Detection Prevalence : 0.5429
      Balanced Accuracy : 0.7618

       'Positive' Class : GR3
```

CAP118 pruned tree

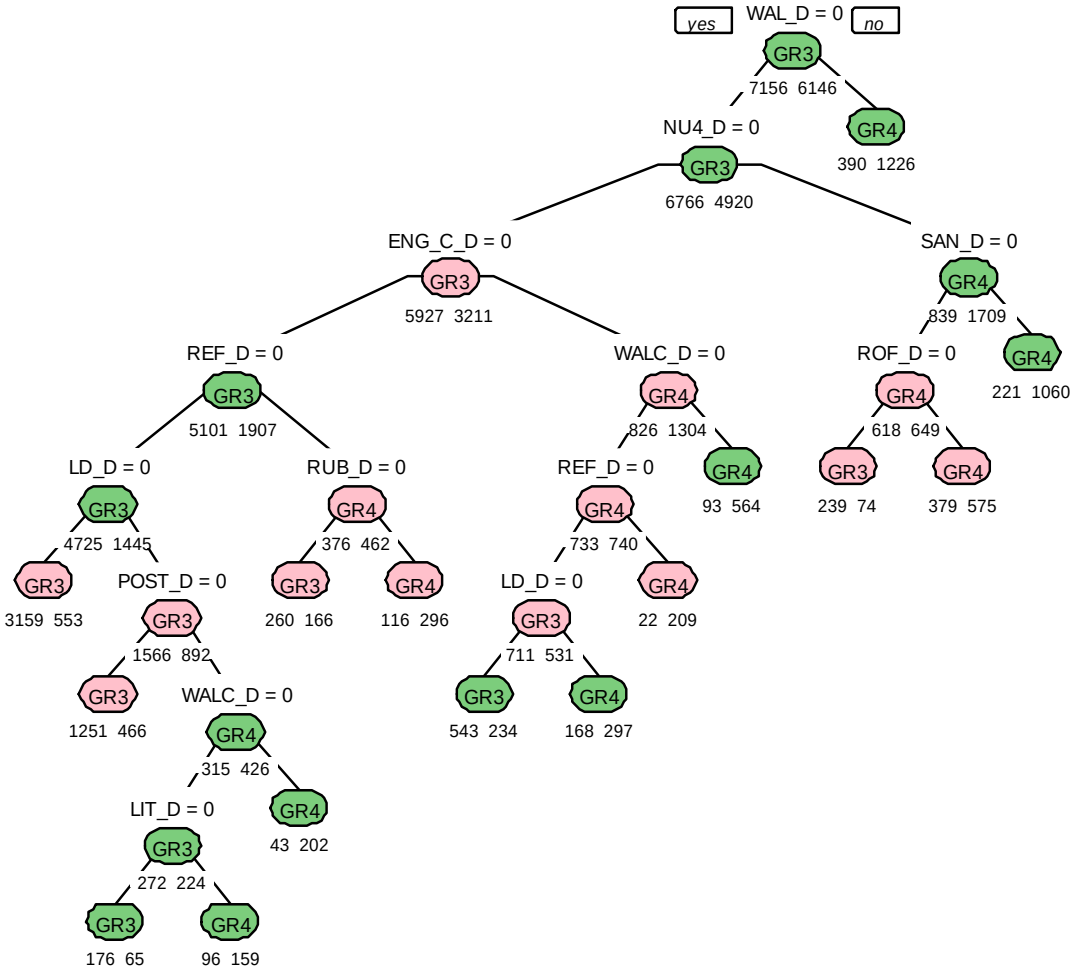


Figure 39. CAP118 pruned tree, data: South Africa GHS 2011

CAP119[489]: GR3 vs. GR5 analysis

In this analysis, we focus on DEP_GRS3 vs. DEP_GRS5. We use the data containing only people from the poverty group 3 and the poverty group 5. The outcome/target response variable is dichotomous, namely CAP119, taking the value 3 for the group 3 and the value 5 for the group 5 (so that people identified in the group 3 take the value 3 and people identified in the group 5 take the value 5 according to CAP119).

Table 61. Training vs. test samples in CAP119 analysis, data: South Africa GHS 2011

	GR3	GR5
Training sample	7156	6852
Test sample	3579	3427

[489] The meaning of CAP119 is: Capabilities analysis (CAP) of 1st step (1) for 9th target group (9).

As the table above shows, there are 14008 individuals in the training data and 6090 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 51 per cent for GR3 and 49 per cent for GR5 in both training and test samples. There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree is with 21 splits (22 terminal nodes). The smallest cross-validation error rate is equal to 0.1123759. According to the plot of the cross-validation error rate by cp and size of tree, the optimal pruned tree can be around with 15 splits but as in previous cases we will accept a little more misclassification rate and set the pruned tree with a cp equal to 0.005983654 (by keeping other parameters the same) to obtain a tree with around 11 splits.

```
> summary(dtreeCAP119)#print a summary which examines each node in depth.
Call:
rpart(formula = CAP119 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 14008
            CP nsplit rel error    xerror        xstd
1  0.551225919      0 1.0000000 1.0000000 0.008634524
2  0.150467017      1 0.4487741 0.4487741 0.007149683
3  0.066112084      2 0.2983071 0.2994746 0.006107679
4  0.017513135      3 0.2321950 0.2348219 0.005507645
5  0.011529480      5 0.1971687 0.1971687 0.005099031
6  0.007953882      7 0.1741097 0.1742557 0.004823243
7  0.007661996      9 0.1582020 0.1590776 0.004627064
8  0.005983654     11 0.1428780 0.1494454 0.004496233
9  0.005837712     12 0.1368943 0.1468184 0.004459628
10 0.003867484     13 0.1310566 0.1364565 0.004311099
11 0.003575598     15 0.1233217 0.1297431 0.004211103
12 0.003162094     17 0.1161705 0.1218622 0.004089594
13 0.002500000     21 0.1030356 0.1123759 0.003936871
```

As the CAP119 pruned tree shows (see below), there are 13 terminal nodes with 12 splits in the tree. The 13 discovered patterns are: (1) no deprivation regarding having a refrigerator in the dwelling (REF), nutrition-household runs out of money to buy food (NU1), and the main material used for the walls of the dwelling (WAL); (2) no deprivation regarding REF, and NU1 but deprivation regarding WAL and littering situation where the dwelling is located (LIT); (3) no deprivation regarding REF, NU1, LIT, and household's main energy source for cooking (ENG_C) but deprivation regarding WAL; (4) no deprivation regarding REF, NU1, and LIT but deprivation regarding WAL, and ENG_C; (5) no deprivation regarding REF, LIT, and sanitation (toilet facility) type the household uses (SAN) but deprivation regarding NU1; (6) no deprivation regarding REF, and SAN but deprivation regarding NU1, LIT, and adult nutrition-any adult in the household go hungry because there was not enough food (ADN); (7) no deprivation regarding REF, SAN, and ADN but deprivation regarding NU1, and LIT; (8) no deprivation regarding REF but deprivation regarding NU1, SAN, and nutrition-eat a smaller variety of foods because there was not enough food in the house (NU4); (9) no deprivation regarding REF, NU4, and the condition of roofs of the main dwelling (ROFC) but deprivation regarding NU1, and SAN; (10) no deprivation regarding REF, and NU4 but deprivation regarding NU1, SAN, and ROFC; (11) deprivation regarding REF, and household's main source of energy for heating (ENG_H); (12) no deprivation regarding ENG_H, and nutrition-

cut the size of meals because there was not enough food in the house (NU2) but deprivation regarding REF; (13) no deprivation regarding ENG_H but deprivation regarding REF, and NU2.

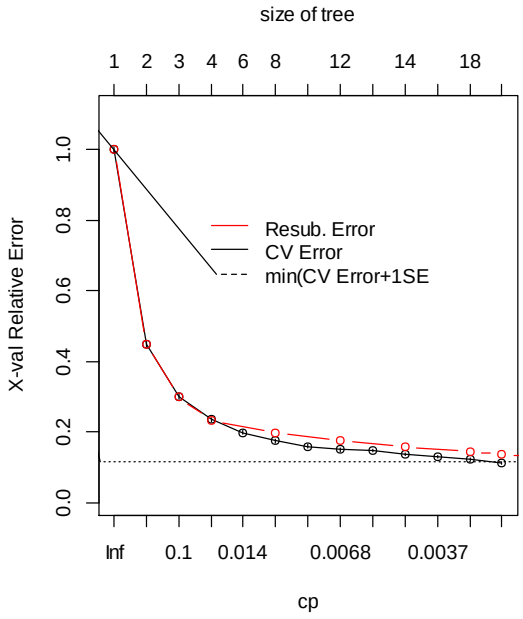


Figure 40. Resubstitution and cross-validated error rates by cp and size of tree and 1-SE Rule for CAP119 saturated tree, data: South Africa GHS 2011

```
Summary(dtreeCAP119.pruned)
Call:
rpart(formula = CAP119 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 14008

           CP nsplit rel error    xerror        xstd
1 0.551225919      0 1.0000000 1.0000000 0.008634524
2 0.150467017      1 0.4487741 0.4487741 0.007149683
3 0.066112084      2 0.2983071 0.2994746 0.006107679
4 0.017513135      3 0.2321950 0.2348219 0.005507645
5 0.011529480      5 0.1971687 0.1971687 0.005099031
6 0.007953882      7 0.1741097 0.1742557 0.004823243
7 0.007661996      9 0.1582020 0.1590776 0.004627064
8 0.005983654     11 0.1428780 0.1494454 0.004496233
9 0.005983654     12 0.1368943 0.1468184 0.004459628
```

The patterns of GR3 are 1st, 5th, 3rd, 12th, 9th, and 7th with 75.8, 7, 4, 3.2, 2, and 1.4 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 75.8 per cent of people from GR3 are detected. On the other hand, the patterns of GR5 are 11th, 8th, 2nd, 10th, 6th, 13th, and 4th with 58.5, 23.75, 4, 2, 2, 1.86, and 0.9 per centage of coverage respectively (see the table below). For instance, with the 11th pattern, 58.5 per cent of people from GR5 are captured. In the training data, by using the CAP119 pruned tree classification patterns, we lose only 7 per cent of information for both GR3 and GR5.

Table 62. Coverage of GR3's patterns among the total population of GR3 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP119 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR3	Cumulative coverage (in %)
1	75.8	75.8
5	7	82.8
3	4	86.8
12	3.2	90
9	2	92
7	1.4	93.4~**93**

Table 63. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP119 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR5	Cumulative coverage (in %)
11	58.5	58.5
8	23.75	82.25
2	4	86.25
10	2	88.25
6	2	90.25
13	1.86	92.11
4	0.9	93.01~**93**

Moreover, through the 6 patterns of GR3, deprivation regarding NU1, WAL, REF, SAN, and LIT concern 10.4, 4, 3.2, 2, and 1.4 per cent of the entire population of GR3 (among 7156 persons) respectively. On the other hand, concerning GR5, the 7 patterns show that the variables REF, ENG_H, NU1, SAN, NU4, LIT, WAL, ROFC, AND, NU2, and ENG_C affect 60.36, 58.5, 27.75, 25.75, 23.75, 6, 4.9, 2, 1.86, and 0.9 per cent of the entire population of GR5 (among 6852 persons). The deprivation in REF, NU1, SAN, LIT, and WAL are the common for both GR3 and GR5; however, their magnitude is higher for GR5 compared to GR3 notably for the variables REF, NU1, and SAN.

Concerning the confusion matrix statistics, the overall accuracy rate is quite high around 93 per cent. The misclassification rate is around 6.3 per cent for GR3, and 6.7 per cent for GR5.

```
   Confusion Matrix and Statistics

          Reference
Prediction  GR3  GR5
       GR3 3353  232
       GR5  226 3195

               Accuracy : 0.9346
                 95% CI : (0.9286, 0.9403)
    No Information Rate : 0.5108
    P-Value [Acc > NIR] : <2e-16

                  Kappa : 0.8692

 Mcnemar's Test P-Value : 0.8153

            Sensitivity : 0.9369
            Specificity : 0.9323
         Pos Pred Value : 0.9353
         Neg Pred Value : 0.9339
             Prevalence : 0.5108
         Detection Rate : 0.4786
   Detection Prevalence : 0.5117
      Balanced Accuracy : 0.9346

       'Positive' Class : GR3
```

CAP1110[490]: GR4 vs. GR5 analysis

In this analysis, we focus on DEP_GRS4 vs. DEP_GRS5. We use the data containing only people from the poverty group 4 and the poverty group 5. The outcome/target variable is dichotomous, namely CAP1110, taking the value 4 for the group 4 and the value 5 for the group 5 (so that people identified in the group 4 take the value 4 and people identified in the group 5 take the value 5 according to CAP1110).

Table 64. Training vs. test samples in CAP1110 analysis, data: South Africa GHS 2011

	GR4	GR5
Training sample	6146	6852
Test sample	3074	3427

[490] The meaning of CAP1110 is: Capabilities analysis (CAP) of 1st step (1) for 10th target group (10).

CAP119 pruned tree

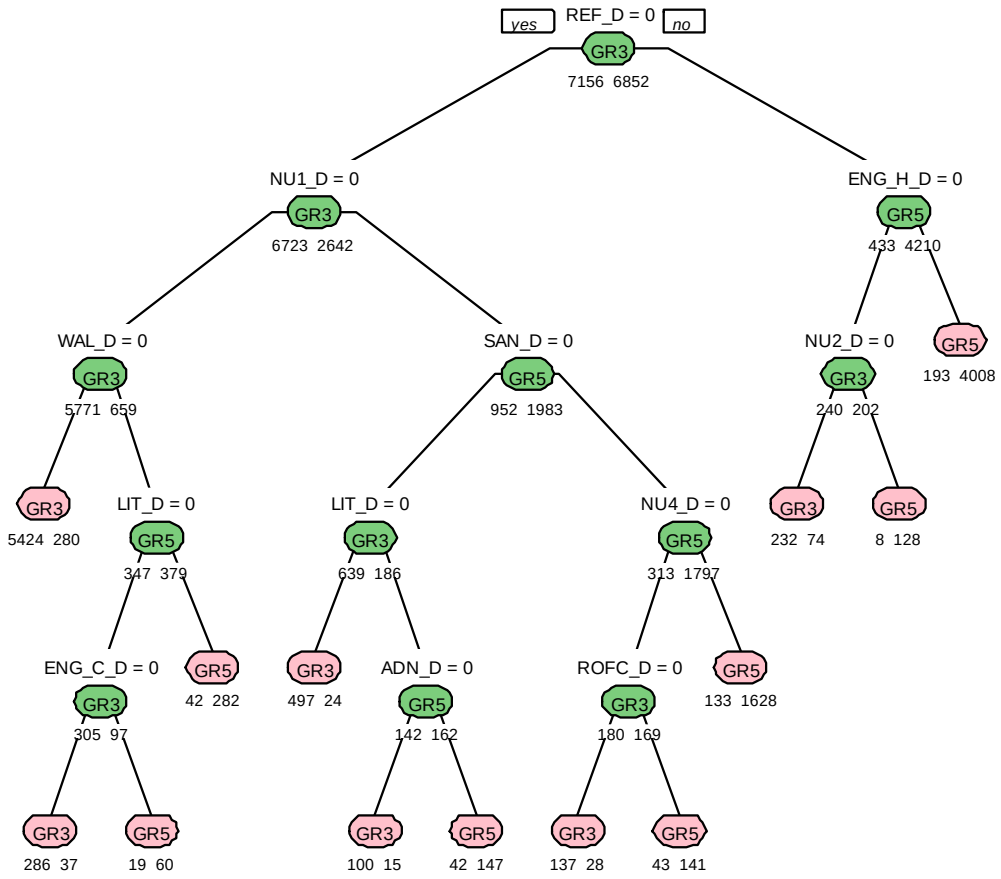


Figure 41. CAP119 pruned tree, data: South Africa GHS 2011

As the table above shows, there are 12998 individuals in the training data and 6501 individuals in the test sample. Training and test data are obtained by randomly partitioning the initial data/total sample into two subsets by 70 per cent vs. 30 per cent respectively. The data is quite balanced with 47.3 per cent for GR4 and 52.7 per cent for GR5 in both training and test samples. There are 33 regressors in the model. Concerning parameters, the gini index for the split, 0.025 as complexity parameter are used, and the analysis is realized on the training data to obtain the classification rules. The information in the box below contains the resubstitution, cross-validation error rates by complexity parameter. The biggest tree contains 34 splits (35 terminal nodes). The plot of cross-validation error by cp and size of tree and of 1-SE Rule shows that the optimal pruned tree can be around of cp=0.0028; however, with such level the tree will be still quite complicated to interpret even if the accuracy will be better. As in previous trees, we will apply the pruning phase to obtain a tree with around 10-12 splits, which means that the cp equal 0.006345591 will be used in the pruning model (by keeping other parameters the same).

```
> summary(dtreeCAP1110)#print a summary which examines each node in depth.
Call:
rpart(formula = CAP1110 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 12998

           CP nsplit rel error    xerror        xstd
1  0.363163033      0 1.0000000 1.0000000 0.009261346
2  0.072079401      1 0.6368370 0.6368370 0.008509776
3  0.025707777      3 0.4926782 0.4983729 0.007872753
4  0.014969086      5 0.4412626 0.4469574 0.007573238
5  0.007809958      6 0.4262935 0.4259681 0.007439653
6  0.007728604      8 0.4106736 0.4176700 0.007384908
7  0.006671006     10 0.3952164 0.3999349 0.007264077
8  0.006345591     11 0.3885454 0.3911487 0.007202220
9  0.005911704     15 0.3612105 0.3776440 0.007104465
10 0.004881224     18 0.3434754 0.3499837 0.006893570
11 0.004718516     20 0.3337130 0.3462415 0.006863881
12 0.003579564     21 0.3289945 0.3390823 0.006806283
13 0.003254149     22 0.3254149 0.3384315 0.006800994
14 0.003091442     24 0.3189066 0.3359909 0.006781080
15 0.002928734     26 0.3127237 0.3330622 0.006757017
16 0.002766027     27 0.3097950 0.3263911 0.006701515
17 0.002603319     28 0.3070290 0.3232997 0.006675465
18 0.002521966     30 0.3018223 0.3192320 0.006640863
19 0.002500000     34 0.2917345 0.3156525 0.006610105
```

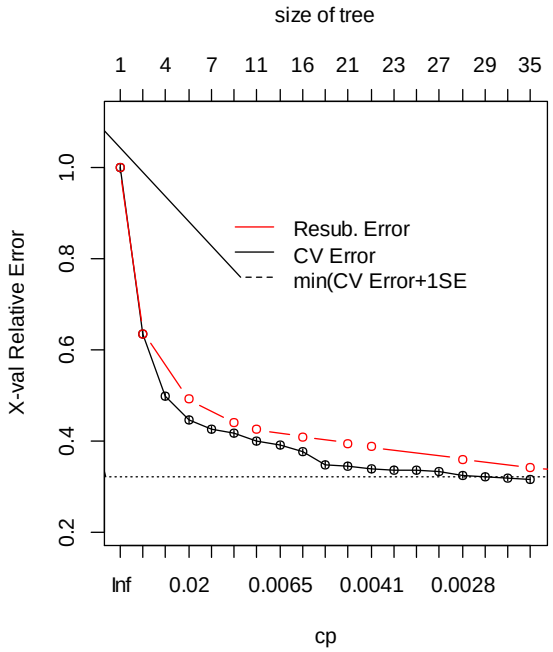


Figure 42. Resubstitution and cross-validated error rate by cp and size of tree and 1-SE Rule for CAP1110 saturated tree, data: South Africa GHS 2011

```
> summary(dtreeCAP1110.pruned)
Call:
rpart(formula = CAP1110 ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)
  n= 12998

           CP nsplit rel error    xerror        xstd
1 0.363163033      0 1.0000000 1.0000000 0.009261346
2 0.072079401      1 0.6368370 0.6368370 0.008509776
3 0.025707777      3 0.4926782 0.4983729 0.007872753
4 0.014969086      5 0.4412626 0.4469574 0.007573238
5 0.007809958      6 0.4262935 0.4259681 0.007439653
6 0.007728604      8 0.4106736 0.4176700 0.007384908
7 0.006671006     10 0.3952164 0.3999349 0.007264077
8 0.006345591     11 0.3885454 0.3911487 0.007202220
```

As the CAP1110 pruned tree shows (see below), there are 12 terminal nodes with 11 splits in the tree. The 12 discovered patterns are: (1) no deprivation regarding having a refrigerator in the dwelling (REF), and nutrition-cut the size of meals because there was not enough food in the house (NU2); (2) no deprivation regarding REF, rubbish collection where the dwelling is (RUB), and the main material used for the walls of the dwelling (WAL) but no deprivation regarding NU2; (3) no deprivation regarding REF, and RUB but deprivation regarding NU2, and WAL; (4) no deprivation regarding REF but deprivation regarding NU2, RUB, the condition of the walls of the main dwelling (WALC); (5) no deprivation regarding REF, WALC but deprivation regarding NU2, RUB, and land degradation (LD); (6) no deprivation regarding REF, WALC, LD, dwelling's main source of energy for cooking(ENG_C) but deprivation regarding NU2, and RUB; (7) no deprivation regarding REF, WALC, and LD but deprivation regarding NU2, RUB, ENG_C; (8) no deprivation regarding REF, and ENG_C; (9) no deprivation regarding ENG_C but deprivation regarding REF, and NU2; (10) no deprivation regarding ENG_C, NU2, and WAL but deprivation regarding REF; (11) no deprivation regarding ENG_C, and NU2 but deprivation regarding REF, WAL, and littering situation where the dwelling is located/neighborhood, community (LIT); (12) no deprivation regarding ENG_C, NU2, and LIT but deprivation regarding REF, and WAL.

The patterns of GR4 are 1st, 2nd, 10th, 6th, and 12th with 54.3, 16, 8, 3.2, and 3 per centage of coverage respectively (see the table below). For instance, with the 1st pattern, 54.3 per cent of people from GR4 are detected. On the other hand, the patterns of GR5 are 8th, 4th, 9th, 5th, 3rd, 11th, and 7th with 44.86, 14, 10.4, 3.85, 2.75, 1.94, and 1.8 per centage of coverage respectively (see the table below). For instance, with the 8th pattern, 44.86 per cent of people from GR5 are captured. In the training data, by using the CAP1110 pruned tree classification patterns, we lose 15.52 per cent of information for GR4 and 20 per cent of information for GR5.

In addition, through the 5 patterns of GR4, deprivation regarding NU2, REF, and WAL concern 19.2, 11, and 3 per cent of the entire population of GR4 (among 6146 persons) respectively. On the other hand, concerning GR5, the 7 patterns show that the variables REF, ENG_C, NU2, RUB, WAL, LD, and LIT affect 57.2, 46.66, 42.8, 19.65, 4.69, 3.85, and 1. 94 per cent of the entire population of GR5 (among 6852 persons). The deprivation regarding NU2, REF, and WAL are the common for both GR4 and GR5; however, the coverage of REF and NU2 are much higher regarding GR5 compared to GR4.

Table 65. Coverage of GR4's patterns among the total population of GR4 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP1110 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR4	Cumulative coverage (in %)
1	54.3	54.3
2	16	70.3
10	8	78.3
6	3.2	81.5
12	3	**84.5**

Table 66. Coverage of GR5's patterns among the total population of GR5 (in per centage) and their cumulative frequencies in per centage from the most to the least frequent patterns of CAP1110 pruned tree, data: South Africa GHS 2011

Pattern no	Coverage of pattern (in %) among the total population of GR5	Cumulative coverage (in %)
8	44.86	44.86
4	14	58.86
9	10.4	69.26
5	3.85	73.11
3	2.75	75.86
11	1.94	77.8
7	1.8	79.6~**80**

CAP1110 pruned tree

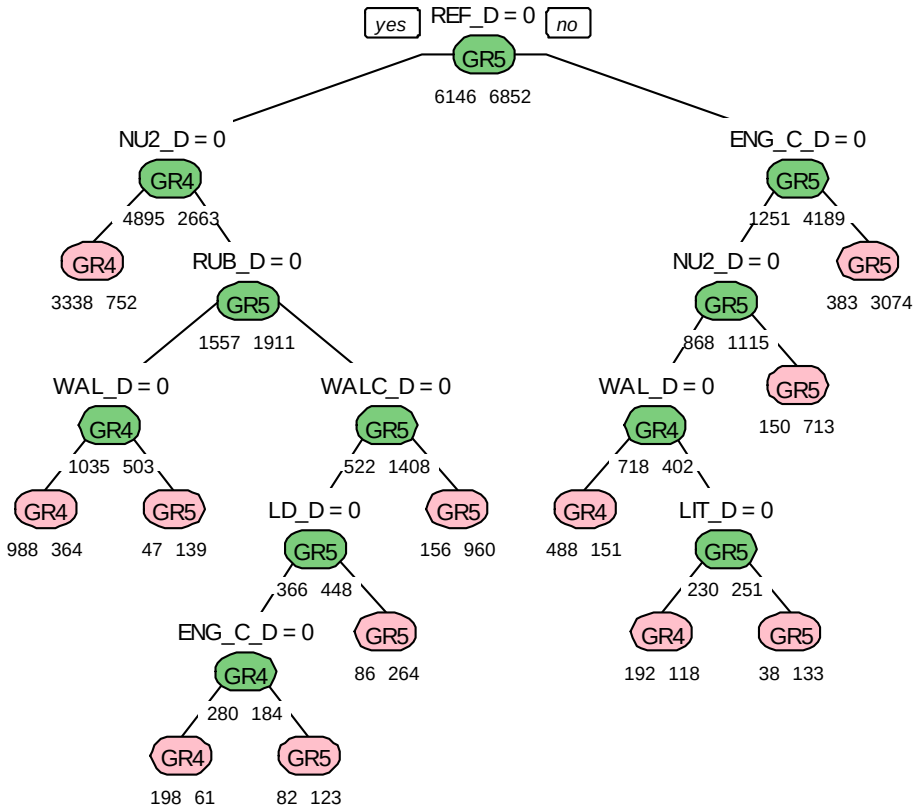


Figure 43. CAP1110 pruned tree, data: South Africa GHS 2011

Concerning the accuracy of the prediction via the classification rules obtained in fitting the mode on the training data, the confusion matrix statistics results mentioned in the box below yield a comprehensive picture on the results of testing. The overall accuracy rate is quite good around 81 per cent. The misclassification rate is 16 per cent for GR4, and around 22 per cent for GR5.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR4  GR5
       GR4 2571  748
       GR5  503 2679

               Accuracy : 0.8076
                 95% CI : (0.7978, 0.8171)
    No Information Rate : 0.5271
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.6156

 Mcnemar's Test P-Value : 5.252e-12

            Sensitivity : 0.8364
            Specificity : 0.7817
         Pos Pred Value : 0.7746
         Neg Pred Value : 0.8419
             Prevalence : 0.4729
```

```
          Detection Rate : 0.3955
    Detection Prevalence : 0.5105
       Balanced Accuracy : 0.8091

        'Positive' Class : GR4
```

We used a binary target response variable in our models. However, there is also possibility to find out the patterns of a categorical variable through the regressors. When the response variable has a greater number of categories, predictive accuracy of classification becomes lower. For testing it, we used the variable DEP_GRS (1=GR1, 2=GR2, 3=GR3, 4=GR4, 5=GR5) and keep the same regressors and the parameters in fitting the model. When the saturated tree is pruned at cp=0.006 level, 12 patterns in total are emerged. We plot the error rates, cp of this model as well as confusion matrix statistics below. We can see that cross-validation error rate is moderately high and overall accuracy is moderately not enough for having a good prediction. These results, therefore, show that running classification for 10 different models give us a more comprehensive picture on the distinction between groups but especially a relatively better prediction result except the models CAP111, CAP115, CAP118, and CAP1110.

```
Classification tree:
rpart(formula = DEP_GRS ~ ., data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
 [1] AP_D   MC_D   NU1_D  NU2_D  POST_D REF_D  ROF_D  ROFC_D SAN_D  WALC_D

Root node error: 23567/30723 = 0.76708

n= 30723
          CP nsplit rel error  xerror      xstd
1  0.1948063      0   1.00000 1.00000 0.0031438
2  0.0901260      1   0.80519 0.80519 0.0036143
3  0.0648364      2   0.71507 0.71507 0.0037012
4  0.0243419      3   0.65023 0.65023 0.0037187
5  0.0172699      6   0.57721 0.57250 0.0036911
6  0.0140451      7   0.55994 0.55357 0.0036763
7  0.0138329      8   0.54589 0.54687 0.0036702
8  0.0089108      9   0.53206 0.53282 0.0036563
9  0.0078924     10   0.52315 0.52480 0.0036475
10 0.0061102     11   0.51525 0.51691 0.0036382
```

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR2  GR3  GR4  GR5
       GR1 2202  647  198   16    1
       GR2  393  784  291   34    1
       GR3  178 1032 2534 1123  123
       GR4    0   14  291  898  393
       GR5    1   34  265 1003 2909

Overall Statistics

               Accuracy : 0.607
                 95% CI : (0.5993, 0.6148)
    No Information Rate : 0.2329
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5028

 Mcnemar's Test P-Value : < 2.2e-16
```

```
Statistics by Class:

                     Class: GR1 Class: GR2 Class: GR3 Class: GR4 Class: GR5
Sensitivity              0.7938    0.31223     0.7080    0.29213     0.8488
Specificity              0.9315    0.94406     0.7916    0.94321     0.8909
Pos Pred Value           0.7187    0.52162     0.5078    0.56266     0.6906
Neg Pred Value           0.9535    0.87541     0.8993    0.84196     0.9536
Prevalence               0.1805    0.16342     0.2329    0.20007     0.2230
Detection Rate           0.1433    0.05103     0.1649    0.05844     0.1893
Detection Prevalence     0.1994    0.09782     0.3248    0.10387     0.2741
Balanced Accuracy        0.8627    0.62815     0.7498    0.61767     0.8699
```

Comparative summary

After presented in detail the results of classification trees through 5 different deprivation groups, we will provide in this sub-section a summary of the discovered patterns regarding content and predictive accuracy.

CAP111

The combination of deprivations regarding both medical coverage and adequate dwelling roof is found as the most important pattern for detecting GR2. In other words, 61 per cent of training sample of GR2 is found deprived regarding medical coverage and adequate dwelling roof. Or we can also say that people lacking medical coverage and living in a dwelling with a bad/inadequate quality of dwelling roof are at most risk of being in GR2. Subsequently, the combination of lack of an adequate dwelling roof and of an adequate access to drinking water, with non-deprivation regarding medical coverage emerges as the second most important pattern for capturing GR2. So that, people lacking an adequate dwelling roof and an adequate access to drinking water but having at the same time a medical coverage are at risk of being in GR2. As GR2 suffers more deprivation than GR1, it is important to consider that experiencing one of these two patterns increases the probability of living in more severe deprivation. Moreover, the pattern of living with a lack of an adequate access to sanitation (toilet) facilities and of medical coverage but living at the same time with a good/adequate quality of dwelling roof corresponds to be at risk of being in GR2. On the other hand, people *not* suffering a bad/inadequate dwelling roof, a non-access/bad quality access to sanitation (toilet) facilities, a bad/not adequate condition of dwelling roof, land degradation, and food shortage/cut of the size of the meals in the house are most likely to be part of GR1. In addition, people *not* lacking a medical coverage, an (adequate) access to drinking water, and a regular collection of rubbish (i.e., littering) in the streets where they live (neighborhood, community), but living at the same time in a dwelling with a bad quality/inadequate roof are likely to be part of GR1. Moreover, people experiencing land degradation, but not having any deprivation regarding roof conditions of dwelling, medical coverage, and sanitation (toilet) facilities are likely to be in GR1. Based on these patterns, the likelihood of being part of more deprived (i.e., being in GR2 compared to GR1) becomes higher especially for those living with deprivation regarding medical coverage, access to drinking water, and access to sanitation (toilet) facilities. Or, we can predict that people at risk of living with deprivation in terms of these three health and living standard related items are at most risk of living in more severe poverty (i.e., being in GR2). The roof material of dwelling emerges as a common deprivation issue for both GR1 and GR2, given that the magnitude of such deprivation is 4.5 times higher for GR2 than GR1 based on the 12 patterns extracted on the training data. Therefore, improvements vis à vis dwellings' constructional/architectural properties, procuration of essential sanitation facilities, health services, lack of food, and

municipality/governmental services for maintaining precautions and measures against land degradation should be top priorities for reducing the likelihood of living poverty as GR2 compared to GR1.

CAP112

People living with an adequate access to sanitation (toilet) facilities are at most risk of being in GR3. Subsequently, people lacking a good condition of dwelling roof, an adequate access to sanitation (toilet) facilities, and a medical coverage are at risk of being in GR3. Moreover, people *not* deprived regarding a bad/not adequate condition of dwelling roof, an adequate access to sanitation (toilet) facilities but suffering at the same time a food shortage-household runs out of money to buy food are at risk of being in GR3. So that, experiencing one of these three patterns increases the probability of living in more severe deprivation, notably the first pattern covers 68.3 per cent of the training sample of GR3. On the other hand, concerning GR1, people *not* suffering a bad/inadequate dwelling roof condition, non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food), an air pollution in their neighborhood/community, living in an environment (neighborhood/community) where there is littering, and no refrigerator in the dwelling are most likely to be in GR1. People *not* suffering a bad/inadequate dwelling roof condition, a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food), an air pollution in their neighborhood/community but living at the same time in an environment (neighborhood/community) where littering appears are likely to be in GR1. Moreover, people *not* experiencing a bad/inadequate dwelling roof condition, a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food) but witnessing at the same time an air pollution in their neighborhood/community and not access to medical coverage are likely to be in GR1. And, people *not* experiencing a bad/inadequate dwelling roof condition, a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food), a non-medical coverage but living at the same time with an air pollution in their neighborhood/community are likely to be in GR1. Based on these patterns, the likelihood of being part of more deprived GR3 (compared to GR1) increases especially for those living with deprivation regarding access to sanitation (toilet) facilities, roof condition, medical coverage, nutrition, education, and littering.

CAP113

People lacking an adequate access to sanitation (toilet) facilities, a medical coverage, and an access to drinking water at the same time are at most risk of being in GR4. Subsequently, people lacking a medical coverage and a shortage of food (household runs out of money), but not lacking an adequate access to sanitation (toilet) facilities at the same time are at risk of being in GR4. Moreover, people lacking an adequate access to sanitation (toilet) facilities, a medical coverage, and a good level of educational attainment, but not lacking an access to drinking water at the same are at risk of being in GR4. So that, experiencing one of these three patterns increases the probability of living in more severe deprivation, notably the first pattern covers 76 per cent of the training sample of GR4. On the other hand, concerning GR1, people *not* suffering a bad/inadequate dwelling roof condition, a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food), and a no energy source for cooking in the dwelling are most likely to be in GR1. People *not* suffering a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food), an access to internet but living at the same time with bad/not

adequate dwelling roof condition are likely to be in GR1. Moreover, people *not* experiencing a lack of medical coverage and a bad/not adequate littering situation where the dwelling is located (neighborhood, community), but suffering at the same time a non-access/bad quality access to sanitation (toilet) facilities are likely to be in GR1. Based on these patterns, the likelihood of being part of more deprived GR3 (compared to GR1) increases especially for those living with deprivation regarding access to sanitation (toilet) facilities, medical coverage, drinking water, and food. The three variables SAN, MC, and NU1 appear in both patterns of GR1 and GR4 however the magnitude of SAN deprivation is 18 times, of MC deprivation is 63 times, and of NU1 deprivation is 11 times higher for GR2 than GR1 (in terms of per centage point) based on the 12 patterns extracted on the training data.

CAP114

People suffering a lack of an adequate access to sanitation (toilet) facilities, of medical coverage, and of adequate roof dwelling are at most risk of be in GR5. Subsequently, people living in not adequate dwelling roof and in a household where any adult went hungry because there was not enough food but not experiencing a lack of an adequate access to sanitation (toilet) facilities are at risk of be in GR5. Notably the former pattern covers 93 per cent of the training sample of G5. On the other hand, having adequate access to sanitation (toilet) facilities, lack of nutrition for the adults in the household, and having a refrigerator in the dwelling at the same time emerges as the most frequent pattern of GR1. According to that, 94 per cent of people of GR1 on the training data are most likely to be in such situation. Moreover, people living in a household witnessing lack of nutrition among the adults of the household and lack of adequate roof dwelling but at the same time not experiencing a no access to not adequate sanitation (toilet) services are likely to be part of GR1. Efforts for diminishing deprivation regarding dwellings' constructional/architectural properties, essential sanitation facilities, health services, food appear likely to be the top priorities in order to reduce the likelihood of living severe poverty as GR5 compared to GR1.

CAP115

People experiencing deprivation regarding an adequate access to sanitation (toilet) facilities, and drinking water are at most risk of being in GR3. Subsequently, people having an adequate access to sanitation (toilet) facilities but suffering food shortage because household run out of money to buy food are at risk of being in GR3. Moreover, (i) living with deprivation regarding dwelling wall condition, access to internet, and a good quality/adequate dwelling roof but also at the same time not experiencing any deprivation regarding access to adequate sanitation (toilet) facilities and nutrition (household run out of money to buy food); and (ii) living with no deprivation regarding access to drinking water but suffering access to adequate sanitation (toilet) facilities and good dwelling roof condition, make likely to be in GR3. So that, experiencing one of these four patterns increases the probability of living in more severe deprivation compared to GR2 on the basis of the training data. On the other hand, concerning GR2, people *not* suffering a bad/inadequate dwelling wall condition, a non-access/bad quality access to sanitation (toilet) facilities, a food shortage (household runs out of money to buy food, and any adult in the household went hungry because there was not enough food) are most likely to be in GR1. People suffering a non-access/bad quality access to sanitation (toilet) facilities, and living with deprivation regarding access to drinking water, food (household run out of money to buy food), with bad littering situation where the dwelling is (neighborhood, community), and with not adequate dwelling roof condition are likely to be in GR2. Moreover, (i) people *not* experiencing a lack of nutrition (NU1), access to adequate sanitation (toilet)

services, access to internet, land degradation but suffering at the same time a not adequate dwelling wall condition; and (ii) people not experiencing no/not adequate access to drinking water but at the same time living with not adequate sanitation (toilet) facilities and non-adequate dwelling roof condition are likely to be in GR2. Based on these patterns, the likelihood of being part of more deprived GR3 (compared to GR2) increases especially for those living with deprivation regarding access to sanitation (toilet) facilities, drinking water, food, wall and roof conditions, and access to internet.

CAP116

People living with non-adequate access to sanitation (toilet) services and post services are at most risk of being in GR4. Subsequently, people having not access to adequate sanitation (toilet) facilities, living with not adequate/bad quality dwelling roof condition, and not deprived regarding post services are likely to be part of GR4. People having access to adequate sanitation (toilet) services but living with food shortage (NU1 and ADN) are likely to be in GR4. Moreover, people experiencing not adequate sanitation (toilet) facilities, food shortage (NU1) but at the same time not suffering regarding post services and good quality dwelling roof condition are at risk of be part of GR4. On the other hand, the pattern of not suffering lack of access to adequate sanitation (toilet) facilities and food shortage (NU1) is the most important for GR1 as 63 per cent of the GR1 on the training data are captured by this pattern. In addition, people suffering food shortage (NU1) and no deprivation at the same time regarding access to adequate sanitation (toilet) facilities, adult nutrition (AND), and dwelling wall condition are likely to be part of GR1. According to the patterns of the training data, the variables SAN and NU1 play important role in identification of GR4 compared to GR2.

CAP117

The pattern of living with deprivation in access to adequate sanitation (toilet) facilities, having energy source for heating in the dwelling, access to drinking water, and access to internet reaches to capture 79 per cent of the most deprived GR5 on the training data. This result means that deprivation in terms of these four items simultaneously indicates the high likelihood of being part of the most deprived group GR5 compared to GR2. The other pattern including deprivation regarding access to adequate sanitation (toilet) facilities, run out of money in the household to buy food but at the same time no deprivation regarding energy source for heating in the dwelling indicates the risk of being in GR5. In addition the following three patterns mean also the risk of be in GR5 but with on average sixteen times less magnitude in terms of per centage point compared to the most frequent pattern): no deprivation regarding access to adequate sanitation (toilet) facilities but suffering lack of food in terms both adult nutrition (ADN) and eating a smaller variety of foods than the person would have liked to because there was not enough food in the house); no deprivation regarding access to internet but suffering lack of access to adequate sanitation (toilet) facilities, energy source for heating in the dwelling, access to drinking water, and dwelling's wall condition; no deprivation in terms of access to drinking water but experiencing lack of access to adequate sanitation (toilet) facilities, energy source for heating in the dwelling, and nutrition in terms of household run out of money to buy food. On the other hand, people suffering lack of access to adequate sanitation (toilet) facilities and lack of food for adults (any adult in the household went hungry because there was not enough food) are the most likely to be in the most deprived GR5. Subsequently, people suffering lack of access to adequate sanitation (toilet) facilities but at the same time not experiencing deprivation regarding the main energy source for heating in the dwelling, nutrition in terms of the household where he or she lives has run out of money to buy food, and the main

material used for the wall of the dwelling are likely to be in GR5. Moreover, people living with no access to adequate sanitation (toilet) facilities and no access to energy source for heating in the dwelling but at the same time not being deprived regarding access to drinking water and nutrition (household runs out of money to buy food) are likely also to be part of GR5 given that the probability of the former pattern is higher than this one. We can say that the inequality between GR2 and GR5 is being shaped notably by the lack of access to sanitation (toilet) facilities, energy source for heating, dwelling's structural situation in terms of quality of wall and lack of food to the detriment of GR5.

CAP118

People living in a dwelling with a bad quality/inadequate wall are at most likely group to be part of GR4. Bu on the other side, people who are living in a dwelling with appropriate wall quality but suffering access to adequate sanitation (toilet) facilities and food (eat a smaller variety of foods than he or she would have liked to, because there was not enough food in the house) are also at most likely to be part of GR4. These two patterns cover 20, and 17 per cent of the population of GR4 on the training data, respectively. In addition, people suffering eating a smaller variety of foods due to food lack in the house and living in a dwelling with bad quality/inadequate roof but not having deprivation regarding the access to adequate sanitation (toilet) facilities and bad quality/inadequate dwelling wall are at risk of being part of GR4. People who are living with bad wall conditions and no proper energy source for cooking in their dwelling but at the same time not suffering regarding nutrition in terms of eating smaller variety than he or she would have liked, are at risk of being part of GR4. At lower proportion, some people of GR4 living with no appropriate main source of energy for cooking, and experienced with land degradation, but living in a dwelling with adequate wall condition, having a refrigerator, and not obligated to reduce her/his eating to a smaller variety are classified as GR4. With same proportion as the former pattern, around 5 per cent of GR4 are captured as living with no deprivation regarding bad quality/inadequate dwelling wall, reduction of variety of foods, main energy source for cooking but not having a refrigerator in their dwelling which is located in an environment where rubbish are not regularly or appropriately collected. On the other hand, the middle deprived GR3 is mostly characterized by no deprivation regarding dwelling wall material, reduction of food variety, main source of energy for cooking, having a refrigerator in the dwelling, and land degradation. After that, people experiencing land degradation but living with no deprivation regarding food variety, main energy source for cooking, refrigerator, and proper access to post/mail services are also at risk of being part of GR3. Moreover, people suffering a no access to adequate main source of energy for cooking but living in a dwelling with adequate/good quality wall condition, having a refrigerator in the dwelling, with no land degradation impact, and no reduction of meals are likely to being part of GR3 compared to GR4. Based on the patterns on the training data, we can say that deprivation regarding dwelling wall in terms of conditions and main material, adequate sanitation (toilet) facilities, and dwelling roof in terms of material are the most important elements making more likely to be part of GR4. Deprivation regarding land degradation, mail services, refrigerator in the dwelling, main energy source for cooking in the dwelling, and nutrition (NU4) flourish as the common attributes with different magnitudes. Except land degradation, the frequency of other 4 items are higher in the patterns of GR4.

CAP119

Classification on the training data shows that 58.5 per centage of GR5 are captured by the pattern of deprivation in both having refrigerator in the dwelling and main energy source for

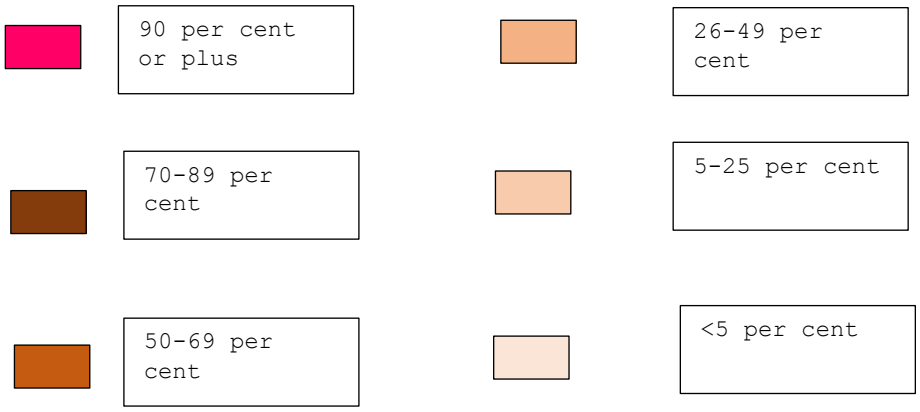
heating in the dwelling. With this pattern, it is most likely to be in GR5 compared to GR3. Subsequently, with around 24 per cent, the pattern of deprivation regarding nutrition (NU1 and NU4: household runs out of money to buy food, and household members consumed a smaller variety of foods than they would have liked to, because there was not enough food in the house), and access to adequate sanitation facilities, and of no deprivation regarding having refrigerator in the dwelling is classified by GR5. In addition, 4 per cent of people of GR5 are classified by the pattern of living in a dwelling with bad quality/not adequate wall in an environment where littering appears, and of living with a refrigerator in the dwelling and with no nutrition (NU1) problem. On the other hand, people living with no deprivation regarding nutrition (NU1), dwelling wall quality, and refrigerator in the dwelling are at most likely to be part of GR3. Subsequently, people not suffering deprivation regarding refrigerator in the dwelling, access to adequate sanitation (toilet) facilities, littering situation where he or she lives (neighborhood, community) but at the same time suffering deprivation in terms of nutrition (NU1) are also likely to be part of GR3. In addition, with 4 per cent of coverage, the pattern of not lacking refrigerator in the dwelling, food shortage (NU1) but lacking at the same time good dwelling wall quality and living in an environment with littering problem is classified with GR3. Notably, GR5 distinguishes from GR4 in terms of deprivation regarding nutrition (NU4), and main energy source for heating. Other items, REF, NU1, SAN, LIT, and WAL exist in both GR3 and GR5 pattern however their frequency is higher among people of GR5.

CAP1110

Concerning the most deprived two groups GR4 and GR5, people living with deprivation in terms of having refrigerator in the dwelling and main energy source for cooking are at most risk of being in GR5 (which concern 47 per cent of population of GR5-almost one in two persons on the training sample). Subsequently, people suffering food shortage (cut the size of meals because there was not enough food in the house), bad/not adequate dwelling wall condition, rubbish in the environment where the person lives and no suffering regarding having a refrigerator in the dwelling are likely to be in GR5. In addition, people having adequate energy source for cooking but not having a refrigerator in their dwelling and living also with food shortage (cut the size of meal) are likely to be in GR5. On the other hand, people having a refrigerator in their dwelling and not cutting the size of meals because there was not enough food in the house are at most likely to be in GR4 (which concern 54 per cent of population of GR4 on the training sample). People having a refrigerator in their dwelling, living with good/adequate dwelling wall, and not in an environment suffering no collection or irregular or seldom collection of rubbishes but suffering in terms of nutrition (cutting the size of meals) are likely to be in GR4. In addition, people living in a dwelling with good/ adequate wall, a proper energy source for cooking, and with no cutting the size of meals but not having a refrigerator in their dwelling are likely to be in GR4. Notably, based on the patterns, we can say that GR5 differentiates from GR4 in terms of energy source for cooking, and living in an environment with no proper collection of rubbishes. The other 3 items, REF, NU2, and WAL exist in patterns of both GR4 and GR5; however, the frequency of REF and NU2 is much higher for GR5.

After giving the main results per mode from CAP111 to CAP1110, now we will illustrate these configurations in Figure 44, and Figure 45. Figure 44 below focuses on how much of the population (in per centage) the deprivation regarding the variable concerns in the training data based on the group patterns in which the variable appears. There are 36 columns labelled as model, group, and variables-from DWE to MC, respectively. The rows in the figure include the name of models (e.g., CAP111) and the groups existing in the model/target groups (e.g., GR1

vs. GR2). For instance, in pruned CAP111 tree, the deprivation in terms of medical coverage (MC) appears in 4 patterns of GR2. When we count the frequency of individuals classified by these 4 patterns as GR2 in the training data, we reach a total of 70 per cent. So that, the deprivation regarding MC concerns 70 per cent of the population of GR2 in the training data. These numbers are already illustrated in a table format in the presentation of results per model. The aim of creating this new figure below is to provide a summary look at the numerous results flourished in the analysis, which may be used in social reports for informing public and non-experts. We defined arbitrarily 6 criteria in terms of per centage for reflecting the deprivation magnitude experienced by each group regarding each single important variable appeared in the extracted patterns: 90 per cent +, 70-89 per cent, 50-69 per cent, 26-49 per cent, 5-25 per cent, and < 5 per cent. Each criterion is presented in a different color as follows:



We can see that in CAP111 model, the deprivation in medical coverage (MC) for GR2 is colored in dark brown according to the labels mentioned above. The deprivation regarding the roof dwelling (ROF) for GR2 is colored in middle ton brown, which concerns 67 per cent of the training population of GR2, and so on.

From the information below, we can see that the variables DWE, RO, KIT, ENG_L, NOI, TEL, TV, RAD, CHN, and NU3 are not used in any pruned tree to distinguish the poverty groups. It is important to mention that the variables appearing in the pruned trees are the most influential ones in terms of information gain to distinguish the two groups of the target response variable. However, it does not mean that the groups in the model are not deprived regarding other variables which do not appear in the tree.

The two advantages come to surface when we use classification tree method in poverty analysis, which are: (i) as the life is composed of many different items reflecting different living conditions, such plurality brings intrinsically a large variety of indicators to be employed in such examination. Due to a high number of items, a first question might come into mind of a policy maker might be ‘Where do I have to begin? Which indicator should I take in consideration to start off the poverty reduction project?’. In a dashboard case, a variable having the most deprivation frequency in the data may be selected through simple tabulation process. However, when there are many variables and several groups, tabulation method will be a very long process making difficult interpretation but especially getting summary information. In addition, such method is not robust for prediction. Classification tree method enables policy

maker to be aware of which variables and their interactions bring the most information for predicting a poverty group; (ii) according to our theoretical framework, we explained that within a society there are different ways that poverty manifest itself which leads to the fact that a specific poverty line is not sufficient to define who is poor. So that, we determined 5 poverty groups as a function of a simple overall deprivation index. Of course, the situation of people in poverty group number 5 is clearly harder than the first group in terms of number of deprivations experienced at the moment of the survey. By examining how different groups differ gives also a picture of inequality in terms of whether a deprivation regarding a variable or experiencing a combination of deprivations or both deprivation and non-deprivation regarding at least two variables (i.e., interaction among the variables) is more relevant for a particular group than other or others. If a government wants to reduce inequality between groups in a society, which is very important for safety and health of the people in a general sense for avoiding polarization, stigmatization, marginalization, or discrimination, or love lessness, or hate between social classes, it is crucial to start off by examining how different lives these groups experience in same society. On the other side, discovering the fact that all or many groups witness the same deprivation patterns or combination of deprivation and non-deprivation patterns is also informative to understand the magnitude of such patterns affecting the whole or the majority of the society. In addition, a policy maker can also discover the fact that a particular group experiences more severely some types of deprivations or living in some particular patterns compared to other groups of society in examination. Policy makers, in poverty analysis, need some effective and summary information, given that such social phenomenon is substantially complex and cannot be seen easily in all angles, not only in terms of per centage or numbers but also in terms of shape of data or more specifically in terms of shape of both deprivation and non-deprivation. In this regard, we argue that classification tree method deals with complexity and yield the shapes of poverty.

Table 68 below shows the main confusion matrix statistics clarifying the predictive accuracy of our models. In the table, columns are the models and rows correspond to accuracy (acc), sensitivity (sens), specificity (spec), positive predictive value (pos pred value), negative predictive value (neg pred value), balanced accuracy (balanced acc) computed on the test data, and misclassification rate (misc rate) computed in the training data, respectively. From the results, we can see that the models containing the groups which are close in terms of total number of deprivations have less accuracy. This includes CAP111 (GR1 vs. GR2), CAP115 (GR2 vs. GR3), CAP118 (GR3 vs. GR4), and CAP1110 (GR4 vs. GR5). Notably, the models CAP113 (GR1 vs. GR4), CAP114 (GR1 vs. GR5), CAP117 (GR2 vs. GR5), and CAP119 (GR3 vs. GR5) have the best predictive accuracy. The two other models CAP116 (GR2 vs. GR4), and CAP112 (GR1 vs. GR3) can be considered having a fair predictive accuracy.

Based on the results displayed in Table 68 and the models with good/fair predictive accuracy, we can as a summary say that the variable SAN, having access to adequate sanitation (toilet) facilities, emerges as the most substantial deprivation item for GR3, GR4, and GR5 (in other words, for the most deprived groups). These 3 groups distinguish from the two more prosperous groups mainly by deprivation regarding access to sanitation (toilet) services and to drinking water (SAN, and WATD). GR2, GR3, GR4, and GR5 are notably emerging deprived regarding medical coverage, and dwelling structural/architectural roof and wall material and conditions compared to GR1. GR4, and GR5 appear having less opportunities in terms of having a refrigerator in the dwelling and access to adequate post services compared to GR1, GR2, and GR3. And GR5 differs from the rest regarding having access to adequate internet services and

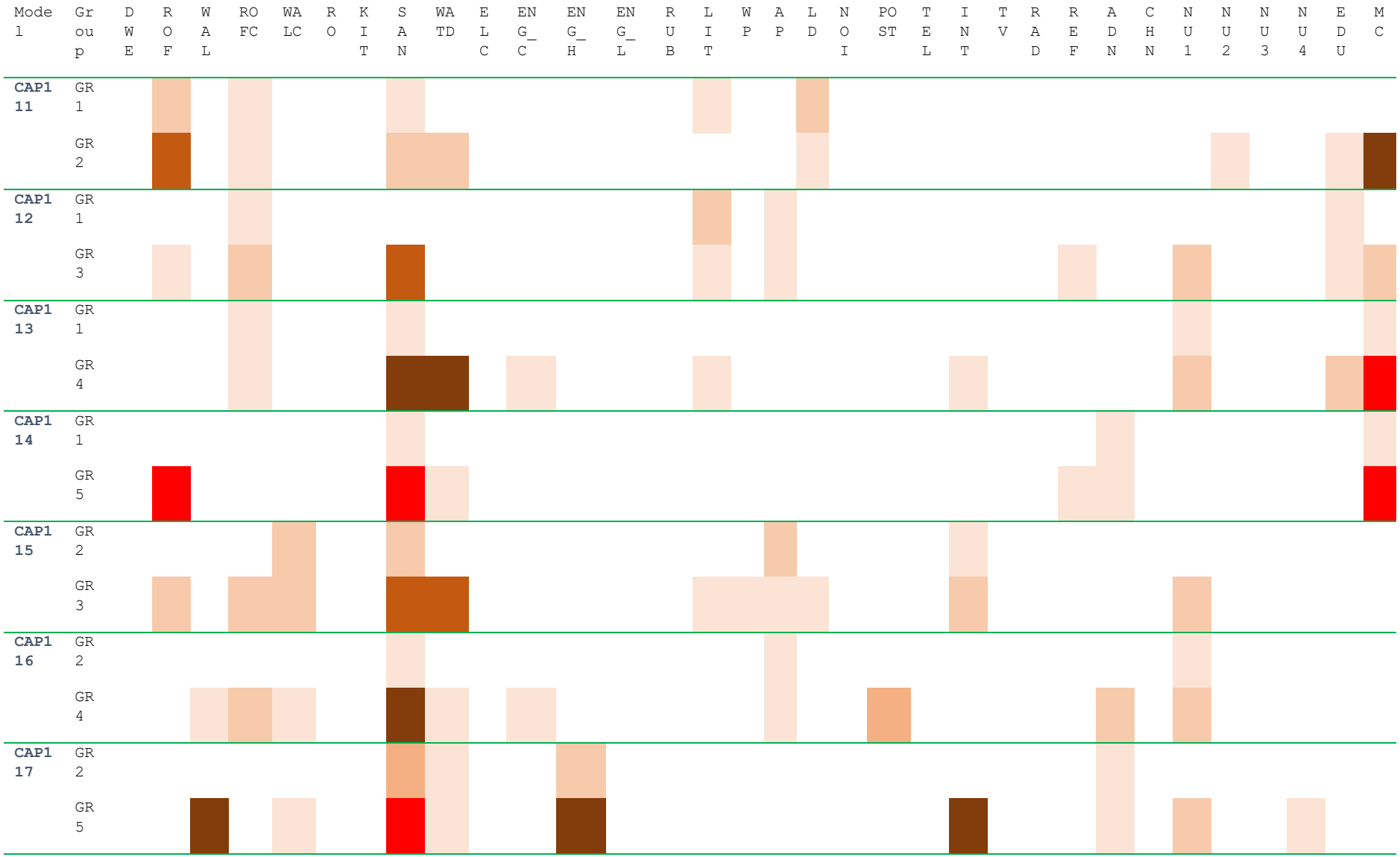
main energy source for heating in the dwelling. It is important to note that based on the pattern's structures, groups cannot be clearly distinguished in terms of nutrition, and education. We will check in more detail the dimensional profile of the group patterns later to understand whether a low educational attainment is a global problem or not even affecting the most prosperous group number 1 among the 5 groups. Concerning nutrition, there are 6 different variables at household level employed in the analysis according to survey questionnaire: run out of money for buying food (NU1), cut the size of meals due to lack of food in the house (NU2), skip any meals due to not enough food in the house (NU3), eat a smaller variety of foods due to not enough food in the house (NU4), whether any adult in the household went hungry because of not enough food in the house (ADN), and whether a child in the household went hungry because of not enough food in the house (CHN). In fact, a negative answer to at least one of NU1, NU2, NU3, and NU4 indicates a food shortage for the household, and AND and CHN are more specific variables composing together a view on food situation of the household. CHN does not appear in any pruned tree; however, ADN could be used instead of NU1, NU2, NU3, and NU4 or NU1, NU2, NU3, and NU4 could be composed together as a one categorical variable instead of ADN. The patterns we obtain include in a heterogenic way the nutrition variables which makes difficult to see a real influence of food shortage among the groups. These 6 variables are all of them integrated as the regressors in the model for a comprehensive view on the reasons behind the nutrition situation. Nevertheless, there is a room for discussing whether these 6 variables should be employed simultaneously or not in fitting the classification model. Similarly, the patterns have a heterogenic view on the dwelling's architectural/structural features across the 4 variables, namely ROF (main material used for the roof of the dwelling), WAL (main material used for the walls of the dwelling), ROFC (roof condition of the dwelling), and WALC (wall condition of the dwelling), among the groups. ROFC and WALC could be rather used without ROF and WAL in the model, or inverse, or ROF and WAL could be combined or ROFC and WALC could be combined to create a composed dwelling structure variable. Thus, there is a room for discussion on the selection of these variables in fitting the model.

In Figure 45 below, the patterns of groups are shown per model. The green color means no deprivation and the red color means deprivation. Rows are the models, and columns include the title of model (Mod), groups (Gr), pattern (Pat), frequency of pattern (Fre) or in other words the coverage of pattern among the training population in per centage, and variables from ROF to MC, respectively (including only variables emerging in the patterns). For instance, the first pattern (P1) of GR1 in CAP111 concerns 51.4 per cent of GR1 people in the training data. This pattern includes no deprivation regarding ROF, ROFC, SAN, LD, and NU2 at the same time. We already gave the full patterns information with their frequency. In addition to that, the goal of this figure is to facilitate the understanding of patterns and their comparison between among models and groups. In CAP112, deprivation regarding sanitation facilities is alone the biggest problem for GR3. For whom are not deprived in access to adequate sanitation facilities, the combination of bad/not adequate dwelling roof conditions and not being covered for medical services emerges as the poorest situation. Then, for who is not deprived in sanitation facilities and dwelling roof conditions, food shortage appears as the most substantial poverty element in the life of GR3. In addition to that, living with air pollution and without a medical coverage is another issue for those of GR3 not suffering in terms of sanitation facilities and dwelling roof conditions. However, for those are not living in bad roof conditions and air pollution risks experience of living with deprivation regarding the material used for dwelling

and littering in their community/neighborhood. As we move to more deprived group GR4 and its comparison with GR1, drinking water adds to sanitation and medical coverage issues. So that, in CAP113, the most alarming situation is for GR4 as being deprived simultaneously regarding sanitation facilities, drinking water, and medical coverage. After that, deprivation in terms of nutrition and medical coverage comes to surface. In CAP114, GR5 members are mostly living with deprivation regarding sanitation, dwelling roof material, and medical coverage at the same time. We still see the most frequent pattern of GR4 in CAP114 for GR5 as well. And deprivation regarding nutrition among adults in the house combined with bad/not adequate dwelling roof is another type of poverty the GR5, the most deprived group, experiences. On the other hand, GR1 as the most prosperous one flourishes as not at risk of experiencing deprivation simultaneously in terms of dwelling roof, sanitation, nutrition, littering, and air pollution in CAP112. When the distance between GR1 and other group increases (i.e., moving from CAP112 to CAP114), GR1 flourishes as not at risk of being deprived simultaneously in terms of dwelling roof, sanitation, nutrition, and main source of energy for cooking in the dwelling in CAP113, and as not at risk of being deprived simultaneously in terms of sanitation, nutrition, and having a refrigerator in the dwelling. From that we can also say that sanitation and nutrition are the main difference drivers between GR1 and 3 other groups. However, energy source for cooking should be considered for GR4 and GR5 and the lack of a refrigerator in the dwelling for GR5 as the two are other inequality sources between them and GR1. Furthermore, GR1 also experiences deprivation regarding littering, air pollution, education, dwelling roof, sanitation, medical coverage. These results lead to the fact that sanitation facilities, dwelling roof conditions, medical coverage, and educational attainment should be considered for the entire country. For the poorest (GR4 and GR5), issues of drinking water, food (nutrition), energy source for cooking, and refrigerator should be the top priorities in national poverty reduction agenda. In terms of environment, littering and air pollution should be also discussed for the health of entire population. When we move to GR2 and its comparison with GR4, and GR5 in CAP116, and CAP117 respectively, we can see that GR4 is mainly captured of being at risk of being deprived regarding sanitation, and adequate post/mail services, and that GR5 is mostly at risk of being deprived regarding sanitation, drinking water, main energy source for heating in the dwelling, and an access to internet. The results of these two models bring additional information showing the importance of heating possibility in dwellings, and internet access for especially the most suffering groups in the country. Therefore, the Figure 44 below sheds light on the importance of variables used in the analysis as regressors for distinguishing groups. However, it is crucial to complement this information by the information on patterns shown in the figures of pruned trees (see chapter 5) and Figure 45 below. This second output shows the most relevant capabilities/combinations of various functionings, in other words the most relevant living conditions emerging in the society among different poverty groups. The first output is necessary as an introductory step but not enough to get the detailed picture on capabilities, which means that the second output is in fact vital in examining different configurations of poverty according to our framework.

Table 67. Main confusion matrix statistics per model, data: South Africa GHS 2011

	CAP111 (GR1 vs. GR2)	CAP112 (GR1 vs. GR3)	CAP113 (GR1 vs. GR4)	CAP114 (GR1 vs. GR5)	CAP115 (GR2 vs. GR3)	CAP116 (GR2 vs. GR4)	CAP117 (GR2 vs. GR5)	CAP118 (GR3 vs. GR4)	CAP119 (GR3 vs. GR5)	CAP1110 (GR4 vs. GR5)
Acc	0.788	0.929	0.978	0.990	0.786	0.927	0.973	0.763	0.934	0.807
Sens	0.749	0.895	0.978	0.989	0.664	0.914	0.964	0.784	0.936	0.836
Spec	0.831	0.956	0.977	0.990	0.871	0.937	0.980	0.738	0.932	0.781
Pos pred value	0.830	0.941	0.975	0.988	0.784	0.922	0.973	0.777	0.935	0.774
Neg pred value	0.750	0.921	0.980	0.991	0.787	0.930	0.974	0.746	0.933	0.841
Balanced acc	0.790	0.925	0.978	0.990	0.768	0.926	0.972	0.761	0.934	0.809
Non-Misc rate (test data)	74.6 83.55	90 95.9	98 98	98.65 99	64.8 97	71 94	97 98	78.5 74	93 93	84.5 80


Model
Group
DWE
ROF
WAL
ROFC
WALC
RO
KIT
SAN
WATD
ELC
ENG_C
ENG_H
ENG_L
RUB
LIT
WP
AP
LD
NOI
POST
TEL
INT
TV
RAD
REF
ADN
CHN
NU1
NU2
NU3
NU4
EDU
MC
CAP111
GR1
GR2
CAP112
GR1
GR3
CAP113
GR1
GR4
CAP114
GR1
GR5
CAP115
GR2
GR3
CAP116
GR2
GR4
CAP117
GR2
GR5

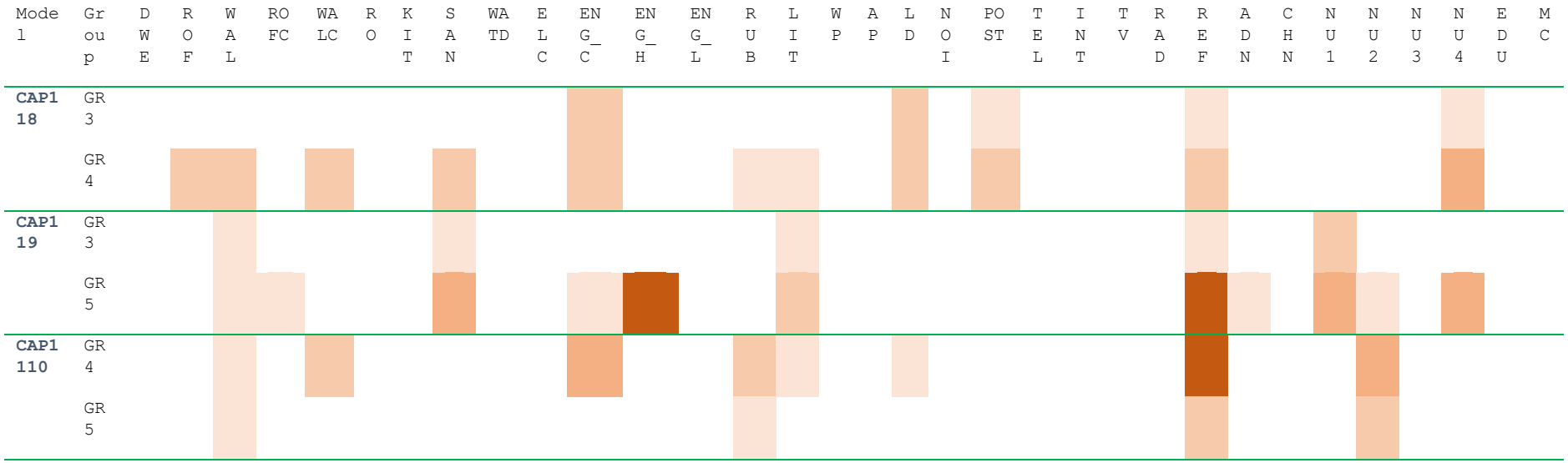
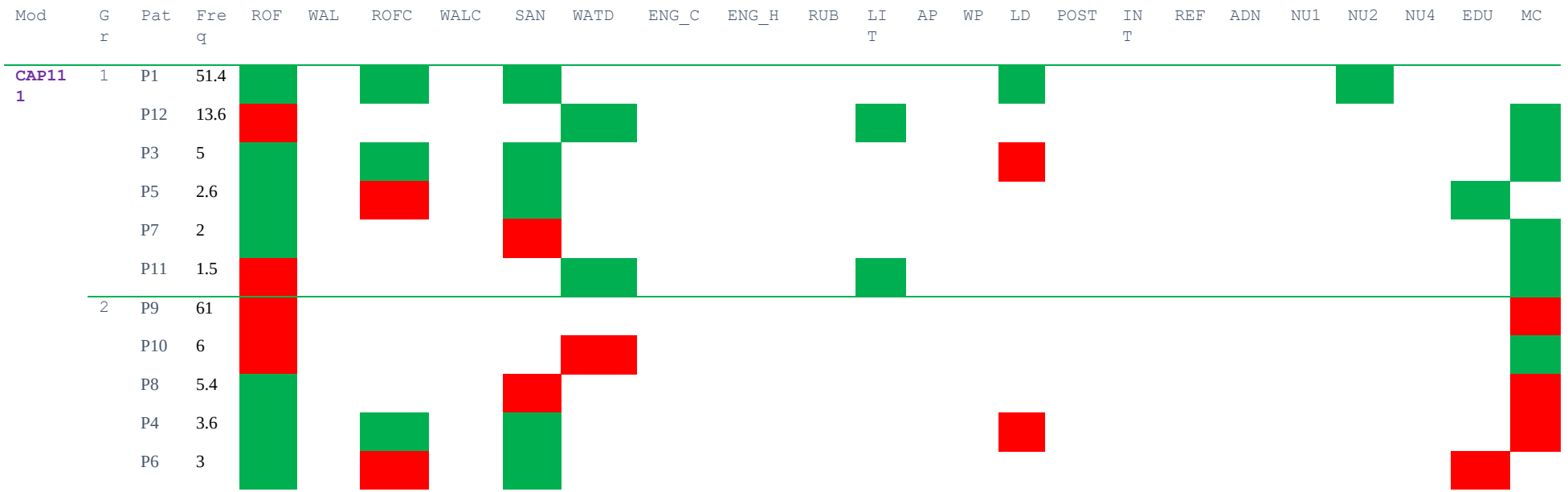
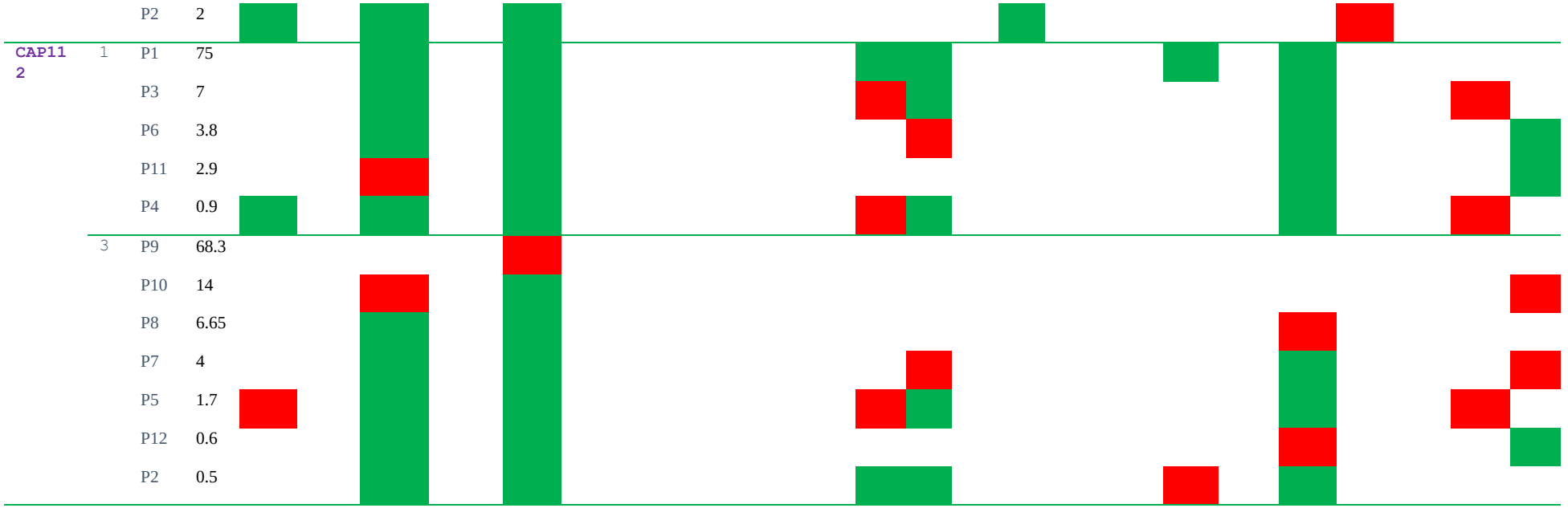


Figure 44. Deprivation per centages represented by different colors in terms of variables emerging in the patterns by group within each model, data: South Africa GHS 2011

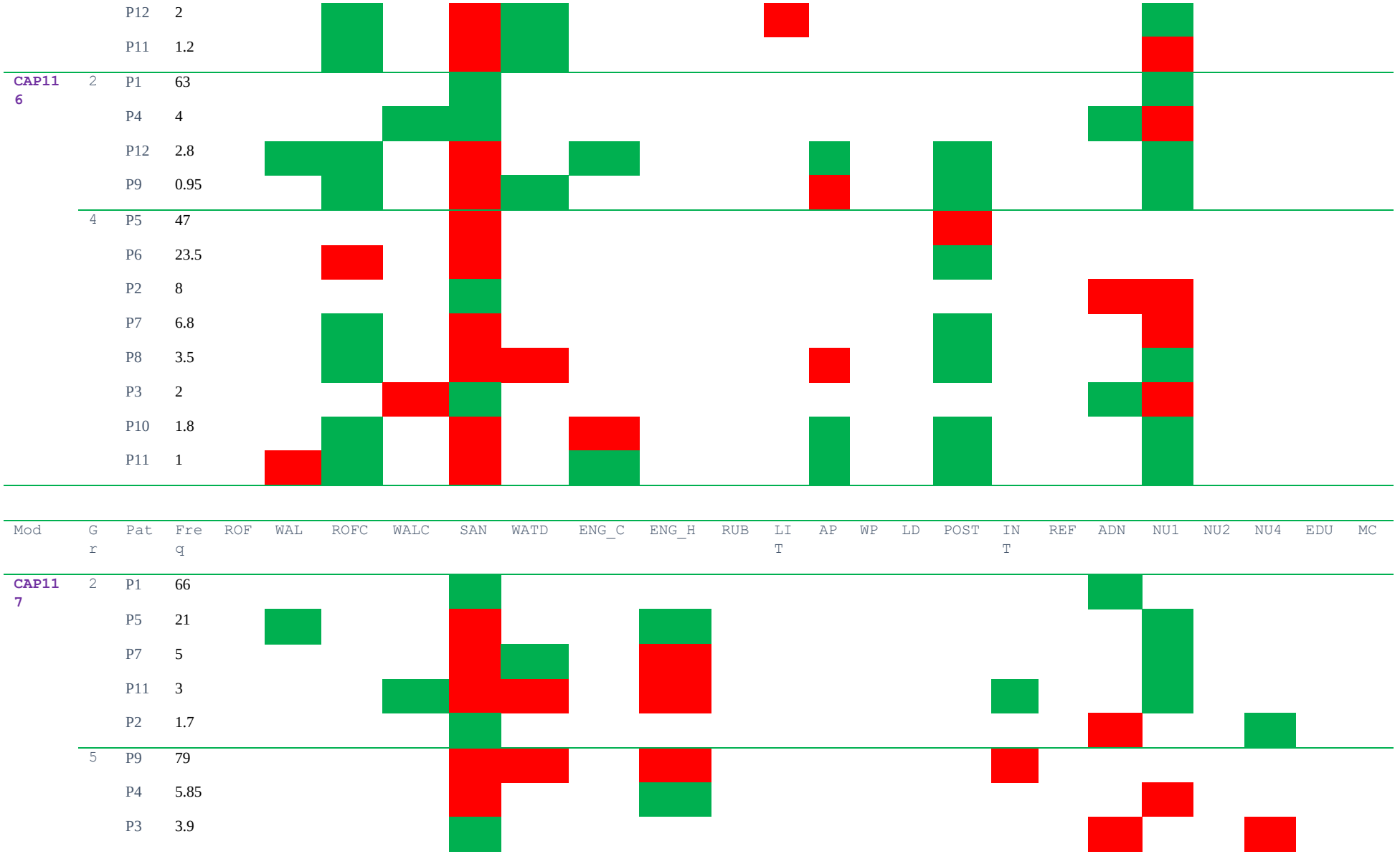




				ROF	WAL	ROFC	WALC	SAN	WATD	ENG_C	ENG_H	RUB	LIT	AP	WP	LD	POST	INT	REF	ADN	NU1	NU2	NU4	EDU	MC
		P2	2																						
CAP112	1	P1	75																						
		P3	7																						
		P6	3.8																						
		P11	2.9																						
		P4	0.9																						
	3	P9	68.3																						
		P10	14																						
		P8	6.65																						
		P7	4																						
		P5	1.7																						
		P12	0.6																						
		P2	0.5																						

Mod	Gr	Pat	Freq	ROF	WAL	ROFC	WALC	SAN	WATD	ENG_C	ENG_H	RUB	LIT	AP	WP	LD	POST	INT	REF	ADN	NU1	NU2	NU4	EDU	MC
CAP113	1	P1	88.5																						
		P3	4																						
		P7	3.2																						
		P11	1.5																						
		P5	1																						
	4	P9	76																						
		P6	11.2																						
		P10	6																						
		P4	2.2																						

P8 1.13
P12 0.7
P2 0.5
CAP114
1
P1 94
P5 3.2
P9 1
P3 0.45
5
P7 93
P4 4
P8 1.2
P6 0.5
P2 0.36
Mod
Gr
Pat
Freq
ROF
WAL
ROFC
WALC
SAN
WATD
ENG_C
ENG_H
RUB
LIT
AP
WP
LD
POST
INT
REF
ADN
NU1
NU2
NU4
EDU
MC
CAP115
2
P2 40
P13 8.8
P5 7.6
P4 6.7
P8 1.7
3
P9 57.5
P1 9.4
P7 7.6
P10 5.6
P3 2.6
P6 2



Mod	Gr	Pat	Freq	ROF	WAL	ROFC	WALC	SAN	WATD	ENG_C	ENG_H	RUB	LIT	AP	WP	LD	POST	INT	REF	ADN	NU1	NU2	NU4	EDU	MC
CAP116	2	P12	2																						
		P11	1.2																						
	2	P1	63																						
		P4	4																						
		P12	2.8																						
		P9	0.95																						
	4	P5	47																						
		P6	23.5																						
		P2	8																						
		P7	6.8																						
		P8	3.5																						
		P3	2																						
		P10	1.8																						
		P11	1																						
CAP117	2	P1	66																						
		P5	21																						
		P7	5																						
		P11	3																						
		P2	1.7																						
	5	P9	79																						
		P4	5.85																						
		P3	3.9																						

P10 3.6
P8 3.4
P6 1.35
P12 0.7
CAP118 3 P1 44
P2 17.5
P10 7.6
P6 3.6
P14 3.3
P4 2.45
4 P12 20
P13 17
P15 9
P8 9
P7 5
P11 5
P9 3.4
P3 3.3
P5 2.6
Mod Gr Pat Freq ROF WAL ROFC WALC SAN WATD ENG_C ENG_H RUB LIT AP WP LD POST INT REF ADN NU1 NU2 NU4 EDU MC
CAP119 3 P1 75.8
P5 7
P3 4
P12 3.2

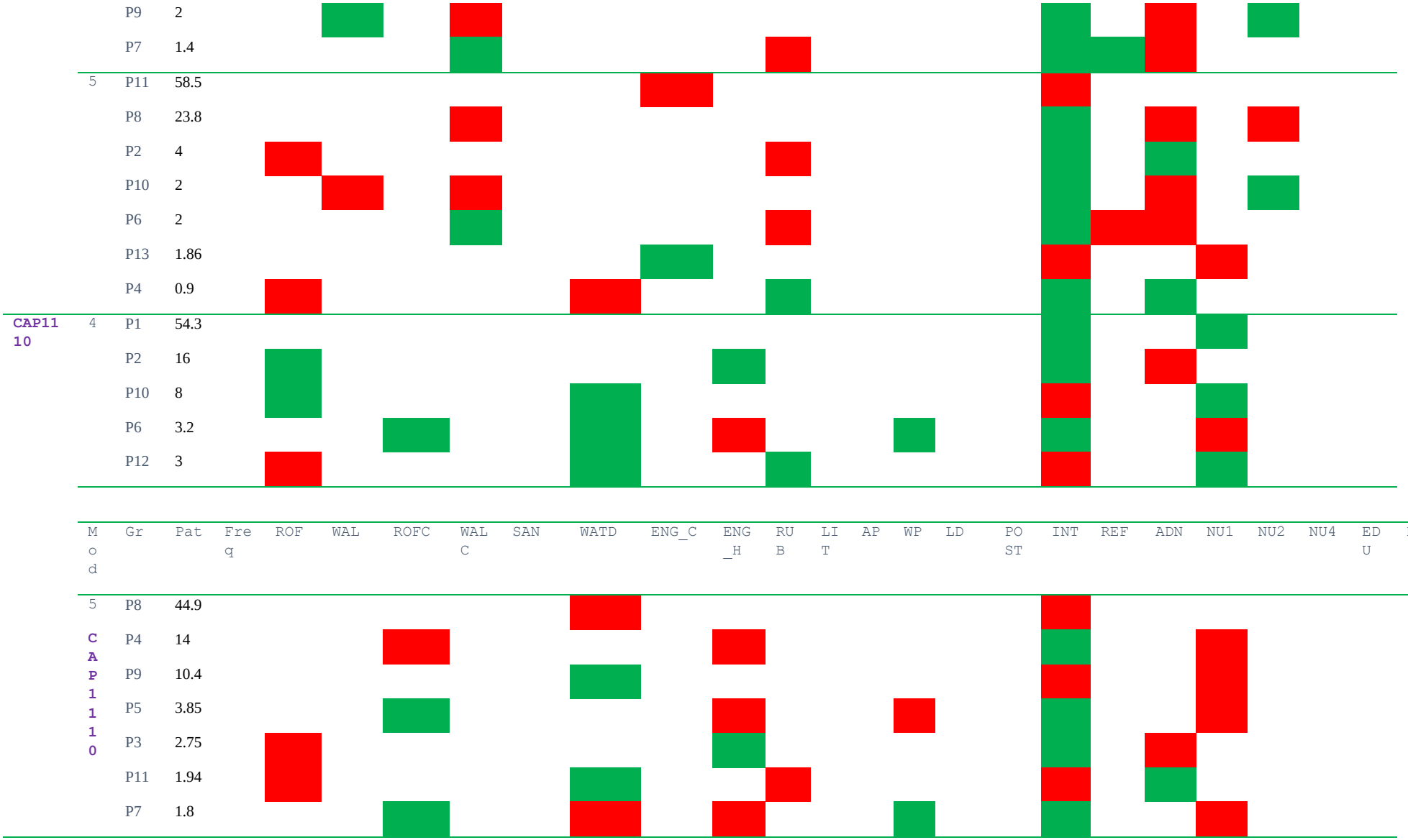


Figure 45. Frequency and composition of classification patterns per model, data: South Africa GHS 2011

5.3 Second step (step2 analysis): detecting dimensional poverty profiles of poverty groups

After having examined in detail the patterns discovered through the classification tree analysis which gave us already the most influential functionings, and their combinations, in this section, we will in particular focus on GR3, GR4, and GR5 in terms of their dimensional profiles. By taking in consideration the performance of models (i.e., predictive accuracy), we can see that for GR3 the model CAP112, for GR4 the model CAP113, and for GR5 the model CAP114 bring the most efficient information concerning their most relevant capabilities patterns. Within each model, by selecting the patterns capturing at least 93 per cent of group population (GR1 is not concerned), we will attempt to find out the variety of profiles of the three groups regarding dwelling, environment, material, nutrition, education, health aspects respectively. The aim of doing this is to detect more information on living conditions of these three most deprived groups.

The dimension of *dwelling conditions* is separated into two sub-sets: (i) DWE1: ROF, WAL, ROFC, WALC, RO; (ii) DWE2: KIT, SAN, WATD, ELC, ENG_C, ENG_L, and ENG_H. The first covers deprivation regarding roof, wall situation of the dwelling as well as the number of rooms available per person. The second includes deprivation regarding availability of kitchen in the dwelling, adequate sanitation (toilet) facilities, access to drinking water, access to electricity, main source of energy available for cooking, lighting, and heating in the dwelling. The main reason behind this separation is profiles become easily complex to interpret when many variables are included, as we have seen in the beginning of the analysis. The dimension of *environment* (ENV) includes deprivation regarding rubbish removal, littering situation in the streets, water pollution, air pollution, land degradation, noise pollution, and available post services of the community, neighborhood where the dwelling is located (where the household, therefore the person lives). The dimension of *material* (MAT) comprises deprivation regarding available telephone, internet services, and radio, tv, and refrigerator in the dwelling. The dimension of *nutrition* (NUT) is composed of 6 variables which are ADN (adult nutrition), CHN (child nutrition), NU1 (run out of money to buy food), NU2 (cut the size of meals), NU3 (skip any meals), NU4 (eat a smaller variety of foods) focusing on household's food situation. The dimension of *education* (EDUC) includes only deprivation regarding educational attainment. And the dimension of *health* (HEALTH) includes only deprivation regarding medical coverage.

We start off by detecting the population of GR3, GR4, and GR5 conforming to the most informative patterns (see below) in the full data (training plus test data), and then we find out their dimensional profiles according to the variables indicated above. This exercise is performed in statistical software STATA, for each group separately.

GR3 dimensional profiles based on CAP112 model

With this model, we saw that the following 4 patterns capture 93 per cent of population of GR3 in the training data: (i) deprivation regarding SAN; (ii) no deprivation regarding SAN but deprivation regarding ROFC, and MC; (iii) no deprivation regarding SAN, ROFC but deprivation regarding NU1; and (iv) no deprivation regarding SAN, ROFC, NU1 but deprivation regarding AP, and MC. When we applied these 4 classification rules in the full data, we lose only 6.8 per cent of information reaching to 93.2 per cent (as in the case of training data), which shows that the model predicts very well GR3.

The table below illustrates the dimensional profiles of GR3 based on the extracted 4 patterns. We note that only profiles concerning at least 5 per cent of the pattern population are shown. The table has 8 columns representing pattern number, and the most substantial profiles of dwelling conditions (DWE1, and DWE2), environment (ENV), material (MAT), nutrition (NUT), education (EDUC), and health (HEALTH) per each pattern. In first column, next to the pattern number, the frequency of the pattern is mentioned in parenthesis in terms of per centage. For instance, 1st pattern concern 68.5 per cent of population of GR3 in the entire data. Profiles are composed of 0 and 1 where 0 means no deprivation and 1 means deprivation regarding the variable. The profiles are composed in the following way:

DWE1= ROF_D + WAL_D + ROFC_D + WALC_D + RO_D where the 5 variables are string character according to STATA language which allows the combination of values and not necessary summing the values to obtain a numeric total value. For instance, the profile 10001 means having deprivation regarding ROF, and RO but no deprivation regarding WAL, ROFC, and WALC. Please see the table of deprivation variables presented in the beginning of the case study. We also note that ROFC_D is removed from the DWE1 equation in case of 2nd, 3rd, and 4th patterns as we have already information on the situation of that variable in the pattern, so there is no necessary to repeat it again in the construction of profiles.

DWE2= KIT_D+ SAN_D + WATD_D + ELC_D + ENG_C_D + ENG_L_D + ENG_H_D. The variable SAN_D is being removed from that equation as the 4 patterns include already information on that.

ENV=RUB_D+LIT_D+WP_D+AP_D+LD_D+NOI_D+POST_D. The variable AP_D is being removed from the equation of ENV in the 4th pattern as it includes already information on that.

MAT=TEL_D+INT_D+TV_D+RAD_D+REF_D.

NUT=ADN_D+CHN_D+NU1_D+NU2_D+NU3_D+NU4_D. The variable NU1_D is being removed from the equation of NUT in the 3rd and 4th patterns as these 2 patterns include already information on that.

EDU=EDU_D

HEALTH=MC_D

In the table, next to the profiles, their frequency (i.e., Freq) within the pattern population is mentioned in parenthesis in terms of per centage. For instance, the profile 10001 in DWE1 concerns 32.5 per cent of population of pattern 1. The sum of frequencies of dimensional profiles in each column for each pattern equals maximum 100 per cent. In case of DWE1 profile for the 1st pattern, there are 4 profiles are selected as the most substantial ones with at least 5 per cent of coverage, and the sum of these frequencies reach 78.6 per cent. Thus, the rest 21.4 per cent is not included. We kept only profiles having at least 100 observations, this is why some loss of information happens in some cases. On the other hand, concerning EDUC and HEALTH profiles, for the 1st pattern, there is no loss of information (the total of frequency of selected patterns reach 100 per cent for both EDUC and HEALTH). The last row in the table indicates the total information captured in each dimension for the ensemble of GR3. For instance, for the DWE1 profile, we do computation in following way:

(68.5*0.786) + (13.8*0.84) + (6.6*0.826) + (4*0.715) = **74**

where 68.5, 13.8, 6.6, and 4 are the frequencies of each single 4 patterns of the GR3 respectively. And 0.786 means 78.6 per cent corresponding to the sum of the frequencies (in per centage) of the DWE1 profiles for the 1st pattern (32.5 + 21.2 + 15.2 + 9.7 = 78.6). Or, 0.84 means 84 per cent corresponding to the sum of frequencies (in per centage) of the DWE1 profiles for the 2nd pattern (54.9+18.6+7.9=84), and so on. At the end, through the selected DWE1 profiles based on the 4 patterns of GR3 in CAP112, we reach to gather information for 74 per cent of the population of GR3 in the full data. We don't necessarily obtain 100 per cent of information. As we know due to variety of patterns, it is quite complex to capture full information. In this regard, the selected DWE1, DWE2, MAT, NUT profiles bring a quite good amount of information, above 70 per cent. The EDUC, and HEALTH profiles reach 93 per cent, i.e., which covers all the 4 patterns as they concern all together 93 per cent of population of GR3 in the full data. However, there are so many profiles found for the ENV dimension which did not allow to obtain a figure above 70 per cent, so that we can only reach 42 per cent of the population of GR3 in the full data based on the dimensional profiles having at least 100 observations in the data.

The pattern 1 (P1) concerns 68.5 per cent of population of GR3 who suffers deprivation regarding access to adequate sanitation (toilet) facilities. 78.6 per cent of them (i.e., 78.6 per cent of GR3 detected by P1) suffer also deprivation regarding roof quality/material of their dwelling but not regarding wall quality/material, and 48 per cent of them live in a dwelling with less than one room per person. In terms of DWE2 aspect, 80.7 per cent of them don't have deprivation regarding kitchen, electricity, and energy source for lighting in their dwelling; but on the other hand, 51 per cent of them lack access to adequate energy source for heating, and 73 per cent of them lack adequate access to drinking water. Nearly half of them don't have deprivation regarding littering, water pollution, air pollution, land degradation, and noise pollution where they do live (around their community, neighborhood). 85 per cent of them don't have deprivation regarding having a television, a telephone, and a refrigerator in their dwelling, but 65 per cent of them experience deprivation regarding access to internet. Most importantly, 88.6 per cent of them don't suffer food shortage. However, a substantial part of that population doesn't have a medical coverage, 70 per cent of them don't have enough or sufficiently enough educational attainment.

The pattern 2 (P2) concerns 13.8 per cent of population of GR3 who suffer deprivation regarding both roof conditions of their dwelling and medical coverage. 81 per cent of them (i.e., 81 of GR3 detected by P2) suffer in terms of wall conditions but not in terms of wall quality/material (this is also the case for 17 per cent of GR3 in P1), and 63 per cent of them live in a dwelling with less than 1 room per person. Regarding DWE2 dimension, 87 per cent of them don't lack kitchen, electricity, energy source for cooking and lighting in their dwelling. However, 46 per cent of them are deprived in terms of energy source for heating. In terms of the dimension of material, 83 per cent of them don't lack a telephone, a tv, and a refrigerator in their dwelling, but on the other hand 55.2 per cent of them suffer deprivation regarding access to internet. In terms of nutrition, 73 per cent of them does not face food shortage. In addition, 100 per cent of that group don't suffer regarding any aspect of environment but in opposite lack a medical coverage. Moreover, 74 per cent of them witness deprivation regarding educational attainment.

The pattern 3 (P3) concerns 6.6 per cent of population of GR3 who suffer living in a household did run out of money to buy food but are not deprived regarding sanitation and roof conditions

of their dwelling. 60 per cent of that group i.e., 60 per cent of GR3 detected by P3) encounter severe nutrition problems in terms of cutting the size of meals, skipping any meals, and eating a smaller variety of foods because there was not enough food in the house. In terms of material, 82 per cent of that group don't have deprivation regarding telephone, television, radio, and refrigerator in the household except 45 per cent face deprivation regarding access to internet. When it comes to dwelling conditions, 72 per cent of them don't have deprivation regarding kitchen, sanitation facilities, drinking water, electricity, energy source for cooking and lighting except 36 per cent witness only deprivation in terms of energy source for heating. Moreover, 73.5 per cent of them live in a dwelling composed of a not adequate/bad roof material but at the same time they do not suffer in terms of wall material and conditions, and 57 per cent of them live in a dwelling with less than 1 room per person.

Throughout the collected main profiles displayed in the table below per dimension and per pattern of GR3, we achieve to capture 93 per cent of population of GR3 for the dimensions of education and health which show that 65 per cent of GR3 suffer deprivation regarding educational attainment and 88 of GR3 lack a medical coverage. In addition to that two aspects, clearly the lack of access to adequate sanitation (toilet) facilities, to proper way of getting drinking water, and the lack of living in a good roof and with an adequate number of room per person (34 per cent of GR3) dwelling are the other functionings reducing substantially the capabilities of the person in her/his life. And, there are 5 dimensional profiles experienced most among the GR3 which are: (i) living deprived regarding dwelling roof material, roof conditions, wall conditions, room per person and not deprived regarding dwelling wall material (48 per cent of GR3); (ii) living deprived regarding dwelling roof material, roof conditions and wall conditions and not deprived regarding dwelling wall material and number of room per person in the dwelling (4 per cent of GR3); (iii) no deprivation regarding all items of the dimension of material (21.5 per cent of GR3); (iv) no deprivation regarding all items of the dimension of material except access to internet (51 per cent of GR3); (v) no deprivation regarding all items of the dimension of nutrition (74 per cent of GR3).

When we adjust the frequency of these deprivations at GR3 population level, we find the following per centage of deprivations:

Regarding the dimension DWE1, 68.5 for SAN_D, 59 for ROF_D, 45 for RO_D, 12 for ROFC_D, and 11 for WALC_D. Regarding the dimension DWE2, 50 for WATD_D, and 43 for ENG_H_D. Regarding the dimension ENV, 19.7 for RUB, and 12 for POST. Regarding the dimension MAT, 55 for INT_D. Regarding the dimension NUT, 6.6 for NU1_D, and 4 for NU2_D, NU3_D, and NU4_D. Moreover, 64 for the dimension of EDUC concerning EDUC, and 69 per cent for the dimension of HEALTH concerning MC. These figures give the following rank from the greatest deprivation to the least among these 12 items: MC, SAN, EDU, ROF, INT, WATD, RO, ENG_H, RUB, ROFC, POST, WALC, NU1, and NU2, NU3, NU4 (with NU2=NU3=NU4 in terms of frequency). For reducing poverty within GR3 this rank can be considered. Previously, through the results of our classification model CAP112, we have noticed that SAN is the most important element, then it is followed by MC, ROFC, and NU1; however, it is important note that this slight difference is due to the fact that CAP112 gives us the most important patterns (combinations) distinguishing GR1 and GR3. So that, for reducing the inequality between GR1 and GR3, it is essential to keep the 4 patterns of GR3 from CAP112 as the basis. Moreover, the fact that for instance the variable MC did not show up at the top of the (pruned) classification tree of CAP112 is due to the fact that the variable SAN distinguishes

at the best level GR1 and GR3 (or in other words the highest difference/inequality between GR1 and GR3 is based on that variable. The difference in terms of not having access to adequate sanitation facilities is the most dramatic factor separating the capabilities of GR1 and GR3).

Table 68. Dimensional profiles of GR3 based on the CAP112 model, data: South Africa GHS 2011

Pattern number (Freq in % in full data)	DWE1 profile (Freq in % in full data)	DWE2 profile (Freq in % in full data)	ENV profile (Freq in % in full data)	MAT profile (Freq in % in full data)	NUT profile (Freq in % in full data)	EDUC profile (Freq in % in full data)	HEALTH profile (Freq in % in full data)
1 (68.5)	10001 (32.5), 10000 (21.2), 10111 (15.2), 10110 (9.7)	010001 (30.6), 010000 (29.8), 010101 (12.8), 000001 (7.5)	0000000 (19.7), 1000001 (18), 1000000 (10.8)	01000 (58), 00000 (20.4), 01010 (6.8)	000000 (88.6)	0 (30.4), 1 (69.6)	0 (6.3), 1 (93.7)
2 (13.8)	1011 (54.9), 1010 (18.6), 0011 (7.9)	000001 (38), 000000 (28.5), 010000 (12.3), 010001 (8)	0000000 (37.8)	01000 (45.2), 00000 (27.9), 01010 (10)	000000 (73)	0 (26.2), 1 (73.75)	0 (0), 1 (100)
3 (6.6)	1001 (33.5), 1000 (25.4), 0001 (23.7)	000000 (36), 000001 (36)	0000000 (45.9)	01000 (45), 00000 (37)	00111 (28.8), 11111 (16), 10111 (15)	0 (36.9), 1 (63.11)	0 (7), 1 (93)
4 (4)	1001 (47.5), 1000 (24)	000000 (25.7), 010000 (22.5)	-	01000 (56), 00000 (29)	00000 (92.4)	0 (28), 1 (72)	0 (0), 1 (100)
Total info/GR3 population coverage in %	(68.5*0.786) + (13.8*0.814) + (6.6*0.826) + (4*0.715) = **73**	(68.5*0.807) + (13.8*0.868) + (6.6*0.720) + (4*0.482) = **74**	(68.5*0.485) + (13.8*0.378) + (6.6*0.459) = **42**	(68.5*0.852) + (13.8*0.831) + (6.6*0.820) + (4*0.850) = **79**	(68.5*0.886) + (13.8*0.730) + (6.6*0.598) + (4*0.924) = **78**	**93**	**93**

GR4 dimensional profiles based on CAP113 model

This model reaches to capture 94 per cent of GR4 population in the training data via the 3 following patterns: (i) deprivation regarding SAN, MC, and WATD; (ii) no deprivation regarding SAN but deprivation regarding NU1, and MC; and (iii) deprivation regarding SAN, MC, EDU but no deprivation regarding WATD. When we applied these 3 classification rules in the full data, we lose only 7.3 per cent of information reaching to 92.7~93 per cent (very close to the score obtained in the training data), which shows that the model predicts very well

GR4. We provide below the same table as in GR3 dimensional analysis, so that we don't comment again the structure of the table.

We note that the variable SAN_D is removed from the equation of DWE2, the variable WATD_D is removed from the equation of DWE2 concerning 1st and 2nd patterns, the variable NU1_D is removed from the equation of NUT concerning 2nd pattern as we have already information on the situation of these variables in the related patterns, so there is no necessary to repeat it again in the construction of profiles.

The pattern 1 (P1) covers almost 76 per cent of population of GR4 who are deprived in terms of sanitation facilities, medical coverage, and drinking water. Compared to GR3, we can see here that the situation aggravates. Concerning the dimension DWE1, all persons in that group (i.e., 76 per cent of GR4 detected by P1) suffer deprivation regarding the roof material/quality in their dwelling; 59 per cent of them live in a dwelling with less than 1 room per person; and 40 per cent of them suffer regarding both roof and wall conditions in their dwelling; and on the other hand the half of them don't suffer regarding the wall quality/material of their dwelling. In terms of the dimension DWE2, 66.5 per cent of them have a deprivation regarding source of energy for heating in their dwelling; 12.6 per cent of them have a deprivation regarding kitchen in their dwelling; on the other hand, half of that group don't experience deprivation regarding kitchen, electricity, energy source for cooking, and energy source for lighting in their dwelling; all persons in that group don't suffer regarding electricity and source of energy for lighting in their dwelling; 63 of them have no deprivation regarding energy source for cooking in their dwelling. In terms of the dimension ENV, all population of that group don't have deprivation regarding water pollution, air pollution, and noise pollution in their neighborhood, community; on the other hand, 42 of them suffer regarding no removal or not sufficient removal of rubbish in their neighborhood, community; 35.5 of them suffer deprivation regarding no removal or not sufficient removal of rubbish, and no sufficient or adequate post services in their neighborhood, community. When it comes to materials, everyone in that group has a television, and a telephone in their dwelling; on the other hand, 63 per cent of them has a deprivation regarding internet. Concerning nutrition, 72.5 per cent of them don't experience any food shortage except 5 per cent of them have run out of money to buy food, cut the size of meals, skip any meals, and eaten a smaller variety of foods because there was not enough food in the house. Most importantly, 82 per cent of them don't have enough or sufficiently enough educational attainment.

The pattern 2 (P2) includes 11 per cent of population of GR4 who are deprived in terms of medical coverage, and nutrition (run out of money to buy food in the household) but not deprived regarding sanitation facilities. In terms of the dimension DWE1, all of them have deprivation in dwelling roof material/quality but not for wall material/quality; half of them have deprivation regarding roof material and roof and wall conditions of their dwelling; 62 per cent of the have deprivation in terms of roof per person in their dwelling; and none of them have deprivation regarding their dwelling's wall quality/material. Regarding the dimension DWE2, all of that population don't have deprivation in terms of kitchen, electricity, energy source for cooking and lighting in their dwelling; but 60 per cent live with a deprivation in terms of energy source for heating in their dwelling. When it comes to environment, 32 per cent of them don't have any deprivation; however, as in the case of GR3 not very much information can be collected in this dimension which is mainly due to its complexity and high variety of patterns of environmental conditions. Among materials, the most important deprivation is in terms of access to internet affecting 55 per cent of that population; on the other hand, all people in that

group don't experience deprivation regarding television, tv, and refrigerator in their dwelling. Food shortage appears as an important issue here meaning that all of them have cut the size of meals, skip any meals, and eat a smaller variety of foods because there was not enough food in the house. And almost 80 per cent of them don't have enough or sufficiently enough educational attainment.

The pattern 3 (P3) concerns 6 per cent of GR4 who are deprived in terms of sanitation facilities, medical coverage, educational attainment but not deprived in terms of drinking water. All population of that group don't suffer regarding having a kitchen, electricity, energy source for cooking and lighting in their dwelling but only 43 per cent of them have a deprivation in terms of energy source for heating in their dwelling. Regarding the dimension DWE1, 59.5 per cent of them live in dwelling with not adequate/bad quality/material roof but also with adequate/good quality/material wall, and with less than 1 room per person; and 43 per cent of them live in a dwelling with deprivation regarding roof and wall conditions, and number of room per person. In terms of material, 36 per cent of them have a deprivation in term of access to internet. In addition, 24 of them suffer food shortage.

When we look at the profiles across all patterns (i.e., concerning 93 per cent of GR4), we can see the following most frequent patterns: (i) for 20 per cent of them, deprivation regarding both roof quality and number of room per person in the dwelling and no deprivation in other 3 items of the dimension DWE1; (ii) for 22.4 per cent of them, no deprivation regarding dwelling wall quality/material but on the other hand deprivation regarding all other 4 items of the dimension DWE1; (iii) for 10.5 per cent of them, no deprivation in all variables of the dimension of ENV, (iv) for 100 per cent of them, no deprivation regarding both air and water pollution, (v) for 100 per cent of them, no deprivation regarding having a television, and a telephone in their dwelling, (vi) 12.5 per cent of them have run out of money to buy food, cut the size of meals, skip any meals, and eat a smaller variety of foods because there was not enough food in the house, (vii) for 100 per cent of them, deprivation regarding medical coverage, (viii) and 77 per cent of them don't have enough or sufficiently enough educational attainment.

When we adjust the frequency of these deprivations at GR4 population level, we find the following per centage of deprivations:

In DWE1 dimension, 70.6 for ROF_D, 54.7 for RO_D, and 35.5 for ROFC_D and WALC_D. In DWE2 dimension, 82 for SAN_D, 76 for WATD_D, 56.8 for ENG_H_D. In ENV dimension, 37 for RUB_D, 27 for POST_D, and 9 for LD_D. In MAT dimension, 56 for INT_D. In HEATLH dimension, 94 for MC_D. In EDUC dimension, 77 for EDU. In NUT dimension, 13.4 for NU1, 7.4 for AP, NU2, NU3, and NU4. These figures give the following rank from the greatest deprivation to the least among these 18 items: MC, SAN, EDU, WATD, ROF, ENG_H, INT, RO, RUB, ROFC, WALC, POST, NU1, LD, AP, NU2, NU3, and NU4. Compared to the dimensional profiles of GR3, we can see that for GR4 as well the three first items, namely medical coverage (MC), sanitation (toilet) facilities (SAN), and educational attainment (EDU) emerge as the most important deprivations. In addition, deprivation regarding drinking water becomes more substantial among GR4 compared to GR3.

Table 69. Dimensional profiles of GR4 based on the CAP113 model, data: South Africa GHS 2011

Pattern number of GR4 (Freq in % in full data)	DWE1 profile (Freq in % in full data)	DWE2 profile (Freq in % in full data)	ENV profile (Freq in % in full data)	MAT profile (Freq in % in full data)	NUT profile (Freq in % in full data)	EDUC profile (Freq in % in full data)	HEALTH profile (Freq in % in full data)
1 (75.7)	10001 (21.7), 10111 (20.7), 10000 (10), 11111 (9.8), 10110 (9.6), 11001 (6.4)	00001 (33.7), 00101 (25.8), 00000 (16.7), 10001 (7), 10000 (5.6)	1000001 (23.5), 0000000 (7.2), 1000101 (7), 1000000 (6.4), 1100101 (5)	01000 (46.4), 01010 (10.7), 00000 (9), 01001 (6.2)	000000 (72.5), 001111 (5)	1 (82.2), 0 (17.8)	1 (100), 0 (0)
2 (11)	10111 (38), 10001 (24), 10110 (12)	000001 (41.8), 000000 (18.4), 010001 (17.6)	0000000 (32)	01000 (44.8), 00000 (16.7), 01010 (10)	11111 (38), 00111 (15), 10111 (14)	1 (79.5), 0 (20.5)	1 (100),0 (0)
3 (6)	10111 (40.3), 10001 (19.2)	000001 (43.4), 000000 (25.4)	0000000 (26)	01000 (36.3), 00000 (19.8)	000000 (32.7), 111111 (23.8)	1 (100), 0 (0)	1 (100), 0 (0)
Total info/GR4 population coverage in %	(75.7*0.217) + (75.7*0.207) + (75.7*0.1) + (75.7*0.098) + (75.7*0.096) + (75.7*0.064) + (11*0.41) + (11*0.28) + (11*0.12) + (6*0.403) + (6*0.227) = **72**	(75.7*0.337) + (75.7*0.258) + (75.7*0.167) + (75.7*0.07) + (75.7*0.056) + (11*0.418) + (11*0.184) + (11*0.176) + (6*0.434) + (6*0.254) = **79**	(75.7*0.235) + (75.7*0.072) + (75.7*0.07) + (75.7*0.064) + (75.7*0.05) + (11*0.32) + (6*0.26) = **42**	(75.7*0.464) + (75.7*0.107) + (75.7*0.09) + (75.7*0.062) + (11*0.448) + (11*0.167) + (11*0.1) + (6*0.363) + (6*0.198) = **64.7**	(75.7*0.725) + (75.7*0.05) + (11*0.38) + (11*0.15) + (11*0.14) + (6*0.327) + (6*0.238) = **69**	**93**	**93**

GR5 dimensional profiles based on CAP114 model

The CAP114 model achieves to capture 97 per cent of GR5 population in the training data with the 2 following patterns: (i) deprivation regarding SAN, MC, and ROF; (ii) no deprivation regarding SAN but deprivation regarding ADN, and ROF. When we applied these 2 classification rules in the full data, we lose only 2.8 per cent of information reaching to 97.2 per cent (the same score obtained in the training data), which shows that the model predicts very

well GR5. We provide below the same table as in GR3 and GR4 dimensional analysis, so that we don't comment again the structure of the table and meanings of figures inside.

We note that the variable SAN_D is removed from the equation of DWE2, the variable ROF_D is removed from the equation of DWE1 concerning 1st and 2nd patterns, the variable ADN_D is removed from the equation of NUT concerning 2nd pattern as we have already information on the situation of these variables in the related patterns, so there is no necessary to repeat it again in the construction of profiles.

The pattern 1 (P1) covers almost 93 per cent of population of GR5 who are deprived in terms of sanitation facilities, medical coverage, and roof material of their dwelling. In terms of the dimension DWE1, almost 32 per cent of that population suffer deprivation regarding roof and wall material of the dwelling, roof and wall conditions of the dwelling, and number of rooms per person in the dwelling. But also, 71.6 per cent of them live in a dwelling with a less than 1 room per person, 68 per cent of them live in a dwelling with not adequate roof and wall conditions, 54 per cent of them live in a dwelling with bad wall conditions and a less than 1 room per person, and on the other hand 38 per cent of them are not deprived regarding the wall quality/material of their dwelling. Regarding the dimension DWE2, 35.6 per cent of them are deprived regarding adequate access to drinking water, having electricity, energy source for cooking, lighting, and heating in the dwelling; 84 per cent of them lack energy source for heating in their dwelling; 23.8 per cent of them lack a kitchen in their dwelling; 77 per cent of them don't have a proper access to drinking water; and almost half of them have electricity and source of energy for lighting in their dwelling. Regarding the dimension ENV, 38 per cent of them live in an environment with a no regular or not frequent collection of rubbish and with a not proper access to post services; 38 per cent of them live in an environment with no air pollution; 19.65 per cent of them live in an environment with no air and noise pollution but with land degradation and lack of adequate post services; 27 per cent of them live with no deprivation regarding littering, water and air pollution; and 18.4 per cent of them suffer deprivation regarding rubbish collection and post services but at the same time they live in an environment with no problem regarding littering, water, air, noise pollution, land degradation. Regarding the dimension MAT, 82 per cent of them lack proper access to internet; 76 per cent of them have a telephone in their dwelling; 38 per cent of them lack proper access to internet and a radio in their dwelling; and 56 per cent of them don't have proper access to internet and don't have a refrigerator in their dwelling. Regarding the dimension NUT, 35.4 per cent of them don't suffer any food shortage but 44 per cent of them suffer in terms of run out of money to buy food, cut the size of meals, skip any melas, and eat a smaller variety of foods because there was not enough food in the house. In addition, 90 per cent of them lack enough or sufficiently enough educational attainment.

When we adjust the frequency of these deprivations at GR5 population level, we find the following per centage of deprivations: in DWE1 dimension, 97 for ROF_D, 69.4 for RO_D, 66 for ROFC_D and WALC_D, 45 for WAL_D; in DWE2 dimension, 93 for SAN_D, 78.5 for ENG_H_D, 72 for WATD_D, 55 for ENG_C_D, and 22 for KIT_D; in ENV dimension, 35.5 for RUB_D, POST_D, and 18 for LD_D; in MAT dimension, 78 for INT_D, 53 for REF_D, 39.4 for TV_D; in NUT deprivation, 48.4 for NU1_D, NU2_D, NU3_D, NU4_D, and 37 for ADN_D; in EDUC dimension, 87.4 for EDUC_D; and in HEALTH dimension, 97 for MC_D. These figures give the following rank from the greatest deprivation to the least among these 23 items: MC, ROF (MC=ROF), SAN, EDU, ENG_H, INT, WATD, RO, ROFC, WALC

(ROFC=WALC), ENG_C, REF, WAL, NU1=NU2=NU3=NU4, TV, AND, RUB=POST, KIT, and LD. As in the case of GR3, and GR4, here also GR5 suffer mostly in terms of medical coverage (MC), access to adequate sanitation (toilet) facilities, and education. Moreover, it is important to note that the roof material/quality of dwelling emerges as the most important deprivation item with MC.

Table 70. Dimensional profiles of GR5 based on the CAP114 model, data: South Africa GHS 2011

Pattern number of GR5 (Freq in % in full data)	DWE1 profile (Freq in % in full data)	DWE2 profile (Freq in % in full data)	ENV profile (Freq in % in full data)	MAT profile (Freq in % in full data)	NUT profile (Freq in % in full data)	EDUC profile (Freq in % in full data)	HEALTH profile (Freq in % in full data)
1 (93.4)	1111 (31.8), 0111 (22), 0001 (9.4), 1001 (8.4), 1110 (8), 0110 (6.2)	011111 (24.6), 010001 (18), 010101 (17.8), 111111 (11), 100001 (7.4), 110101 (5.4)	1000001 (18.4), 1000101 (8.6), 1100101 (5.75), 1110101 (5.3)	01000 (19.4), 01111 (18.7), 01101 (17.2), 01001 (8), 01010 (7), 11111 (6.3), 01011 (5.6)	000000 (35.4), 111111 (26), 001111 (9.3), 101111 (9)	1 (90), 0 (10)	1 (100), 0 (0)
2 (4)	0111 (60.8)	-	-	01000 (28.9)	11111 (60.8)	85.6 (1), 0 (14.4)	1 (99), 0 (1)
Total info/GR5 population coverage in %	(93.4*0.318) + (93.4*0.22) + (93.4*0.094) (93.4*0.084) + (93.4*0.08) + (93.4*0.062) + (4*0.608) = **82**	(93.4*0.246) + (93.4*0.197) + (93.4*0.182) + (93.4*0.11) + (93.4*0.083) + (93.4*0.055) = **81.5**	(93.4*0.184) + (93.4*0.086) + (93.4*0.0575) + (93.4*0.053) = **35.5**	(93.4*0.194) + (93.4*0.187) + (93.4*0.172) + (93.4*0.08) + (93.4*0.07) + (93.4*0.063) + (93.4*0.056) + (4*0.289) = **78**	(93.4*0.365) + (93.4*0.2625) + (93.4*0.184) + (4*0.608) =**78**	**98**	**98**

5.4 Third step (step3 analysis): detecting various profiles of poverty groups

The goal of the decomposition analysis presented here is to find out various profiles of GR3, GR4, and GR5 based on their most efficient patterns in the models CAP112, CAP113, and CAP114 respectively. At this point, sub-group profiles will be searched in terms of grant, employment, individual characteristics, geographical situation, and some household features.

More specifically, the *grant* (GRANT) profile includes the variables, DISG (disability grant), OLDG (old age grant), WARG (war veterans grant), FOSG (foster child grant), DEPG (care dependency grant), AIDG (grant-in-aid), and SRG (social relief or distress grant); the *employment* (EMP) profile contains the variables ES1 (employment status), and SEC (sector of work); the *individual characteristics* (INDC) profile covers the variables GEN (gender), POP (population group/race), and AGE (age group); the *geographical situation* (GS) profile focuses on the variables PRO (province), and SET (urban vs. rural/setting); and the *household features* (HF) profile is composed of the variables DWELS (RDP or subsidized dwelling), HS (household size), hh_child (children in the household), GH (gender of head of household), AH (Age of head of household), POPH (population group of head of household), and INDG (indigent household) respectively. The description of these variables is given in section 4.2.

GR3 sub-group profiles based on CAP112 model[491]

As we can see from the table below, we obtain information for around 90-92 per cent of GR3 (in full data including both training and test sets) for GRANT, and EMP profiles. In other words, with the GRANT profiles presented in the table, we achieve to capture 92 per cent of the population of GR3, so that information for only 8 per cent of the population of GR3 is missing. Concerning INDC and GS profiles, the per centage can be considered quite good, above 70 per cent; however, for the HF, the performance was not good in the sense that the most frequent profiles (with at least 50 observations in the data) are not enough to reach at least 70 per cent of GR3 due to high variety of combinations of 7 variables included in the HF.

In the first pattern of GR3 which concerns 68.5 per cent of GR3 who are deprived in access to adequate sanitation (toilet) facilities, there are 3 GRANT profiles emerging respectively: (i) no receiving any social grant (82.3 per cent); (ii) receiving only old age grant (13.8 per cent); and (iii) receiving only disability grant (3.3 per cent). In other patterns of GR3, we see also these 3 patterns with different per centages. In total, across 4 patterns, we find that 77 per cent of GR3 do not receive any social grant, 12 per cent of GR3 receive only old age grant, and 3 per cent of GR3 receive only disability grant.

Concerning the EMP profile, across the 4 patterns of GR3, we can say that almost 30 per cent of GR3 population did not seek a job and he or she would not have liked to work, and even if a suitable job had been offered or circumstances had allowed he or she would not have been able to start work or a business in the last calendar week of the interview. In addition to that, 26.5 per cent is employed (working for a wage, a commission or a salary or engaged in business activities or temporarily absent but will return to work or business or work without remuneration) and is registered to perform activity in a formal sector; 11.2 per cent did not search but would have liked to work during the last calendar week of the interview; 12.6 per cent sough a job during the last 4 weeks of the interview and available or ready to start to work; 7.5 per cent is employed but works/register to perform activity in an informal sector; and 5 per cent did not seek a job and would not have liked to work but if a suitable job had been offered or circumstances had allowed, he or she would have been able to start work or a business in the last calendar week of the interview.

Concerning the INDC profile, across the 4 patterns of GR3, the profiles 112, 212, 211, 111, 113, 213, 114, and 214 are more encountered compared to the profiles 115, 116, 215, and 216 appearing only in 1st pattern and to the profiles 221, 124, 122, 121, 222, and 123 appearing only

[491] Only profiles with at least 50 observations in the data are revealed in the table.

in 2nd pattern. Each of the profiles 112, 212, 211, 111 cover between 9 and 10 per cent of GR3, which are: female & African/black & 25-34 yrs.; male & African/Black & 25-34 yrs.; male & African/black & 18-24 yrs.; female & African/black & 18-24 yrs. respectively. In addition, each of the profiles 113, 213, 114, and 116 include between 7.7 and 5 per cent of GR3, which are: female & African/black & 35-44 yrs.; male & African/black & 35-44 yrs.; female & African/black & 45-54 yrs.; and female & African/black & 65+ respectively. With these 8 profiles, we reach to capture 61.5 per cent of GR3.

The GS is composed of various sub-profiles with relatively small per centages compared to those of GRANT and EMP. Within the 1st pattern, the patterns 94, 41, 84, 71, 54, and 64 has the adjusted per centages of 13, 7, 7, 4.8, 5, and 4.65 respectively. This leads to a coverage of 41.45 per centage among GR3. Due to complex structure of the data, we cannot obtain efficient information through the profiles of 2nd and 3rd pattern (their percentages are very small with the highest is around 4 per cent). And, when we take in consideration the most informative profiles of 1st pattern (profiles with around 4.5 and 13 per centages), we cannot even reach the half of the GR3. However, the most frequented 6 profiles of 1st pattern are: living in Limpopo in rural formal area; living in Free State in urban formal area; living in Mpumalanga in rural formal area; living in Gauteng in urban formal area; living in KwaZulu-Natal in rural formal area; and living in North West in rural formal area. As the GS profile is composed of two items, namely province and setting, we can also proceed with a simple cross-tabulation which is shown in the table below[492]. We can see that the table is quite long (distribution of 9 provinces and 5 settings across the population of GR3), and it is not very practical to get summary information at first glance. If we start by the rows (province), we can say that: (i) in the province 1 (Western Cape), 90 per cent of population lives in urban formal areas; (ii) in the province 2 (Eastern Cape), 61.6 per cent of population lives in urban formal areas and 31 per cent lives in tribal areas; (iii) in the province 3 (Northern Cape), 85 per cent of population lives in urban formal areas; (iv) in the province 4 (Free State), 85 per cent of population lives in urban formal areas; (v) in the province 5 (KwaZulu-Natal), 43 per cent of population lives in tribal areas, 39.5 per cent lives in urban formal areas, and 15.6 per cent lives in urban informal areas; (vi) in the province 6 (North West), 51 percent of population lives in tribal areas, and 41 per cent lives in rural formal areas; (vii) in the province 7 (Gauteng), 80 per cent of population lives in urban formal areas; (viii) in the province 8 (Mpumalanga), 59 per cent of population lives in tribal areas and 36 per cent lives in urban formal areas; (xi) in the province 9 (Limpopo), 90 per cent of population lives in tribal areas, and 9 per cent lives in urban formal areas. Across the country, the majority of GR3 lives in the provinces of Gauteng and Limpopo, and less live in the provinces of Eastern, and Northern Cape. Regarding type of setting, 56.5 per cent of GR3 lives in urban formal areas and 34 per cent of them lives in tribal areas. In terms of population of GR3 living in urban

[492] We produce simple cross tabulation based on the entire population of GR3 in the full data. To find out the profiles of GRANT, EMP, INDC we employed the data of the most frequented 4 patterns capturing 93 per cent of population of GR3 as a result of the CAP112 classification tree. However, for tabulation process, it is not important to base on the 4 patterns because even if we use 93 per cent or 100 per cent of GR3 population we will in any case have a long table composed of 9 provinces and 5 settings according to the questionnaire. In addition, in case of two variables, there is not very much difference between profile extraction and tabulation because the number of variables is few. Notably, in case of more than 2 variables, the tabulation becomes problematic this is why we focus on profiles to see the combination of variables. Thus, we can say that for the EMP, and GS profiles, both of these 2 methods can be applied; however, for GRANT, INDC, and HHC that are composed of at least 3 variables, it is more efficient to proceed with profiles extraction.

formal areas, Gauteng (where the cities of Johannesburg and Pretoria are located) and Western Cape (where the city of Cape Town is located) have the highest scores; on the other hand, clearly Limpopo (north east of the country sharing in international borders with Botswana, Zimbabwe, and Mozambique, and includes Kruger National Park) is the province where the most of population of GR3 lives in tribal areas, and in Western Cape, there is no population living in tribal areas. In terms of urban informal areas, the province of Gauteng and KwaZulu-Natal (where the city of Durban is located) have the highest scores. And, regarding rural formal areas, the province of Western Cape has the highest population of GR3 but also the province of Northern Cape comes at second place with a per centage of 17.5.

If we look at pattern levels, the majority of GR3 captured by 1st pattern lives in the tribal areas of the province of Limpopo, followed by urban formal areas in Free State, tribal areas in Mpumalanga, urban formal areas in Gauteng, tribal areas in KwaZulu-Natal, and tribal areas in North West. These 6 locations form 62 per cent of that group. With the following other 9 locations, 93 per cent of that group is reached (i.e., corresponding to 63 per cent of the entire population of GR3), respectively: Western Cape urban formal, Eastern Cape urban formal, KwaZulu-Natal urban formal, Eastern Cape tribal areas, Northern Cape urban formal, Mpumalanga urban formal, North West urban formal, Gauteng urban informal, and KwaZulu-Natal urban informal. As we mentioned in previous paragraphs, we also get the same results with profiles extraction process (see the table below) which is an easier way of obtaining a first comprehensive glance compared to cross tables. Nevertheless, when the data is complex and generates a large variety of profiles, both the two methods lead to take in consideration several profiles to obtain a large coverage of the population.

To reduce poverty among GR3, these 15 locations can be put on priority in terms of the functionings appearing in the 1st pattern and its dwelling, material, health, education, nutrition, environment, and economic activity profiles.

Table 71. Sub-group profiles of GR3 based on the CAP112 model, data: South Africa GHS 2011

Pattern number (Freq in % in full data)	GRANT [493]profile (Freq in % in full data)	EMP[494] profile (Freq in % in full data)	INDC profile (Freq in % in full data)	GS profile (Freq in % in full data)
1 (68.5)	33333333 (82.3), 21222222 (13.8), 12222222 (3.3)	60 (29.8), 11 (27.8), 40 (12.8), 20 (12.3), 12 (9.5), 50 (6), 70 (1.46)	112 (12.8), 212 (11.4), 211 (11.4), 111 (11), 113 (9), 213 (7.4), 114 (7.4), 116 (6.9), 115 (5.9), 214 (4.8), 216 (3.3), 215 (3.2)	94 (17.8), 41 (10.2), 84 (10), 71 (7), 54 (7.4), 64 (6.8), 11 (4.3), 21 (4.2), 51 (4), 24 (3.9), 31 (3.8), 81 (3), 61 (3), 72 (2.4),

[493] 25 individuals were dropped from the data due to missing values regarding all variables of *GRANT* dimension.

[494] 376 individuals were dropped from the data due to missing values regarding all variables of *EMP* dimension.

				52 (2)
2 (13.8)	33333333 (83.2), 21222222 (11), 12222222 (5)	11 (32), 60 (25.6), 20 (16.4), 40 (11), 12 (7.4), 50 (5)	211 (7), 111 (6.8), 212 (6.8), 113 (6.5), 112 (6.4), 114 (4.9), 124 (4.6), 221 (5), 122 (4.4), 213 (4.2), 121 (3.8), 222 (3.8), 123 (3.6), 214 (3.4)	11 (29.3), 71 (12), 31 (9.8), 41 (8.7), 21 (8.5), 51 (8), 15 (5), 81 (4.8), 61 (4)
3 (6.6)	33333333 (86.3), 21222222 (9)	11 (28.3), 60 (23.7), 20 (19.8), 40 (14.8)	111 (9.7), 212 (9.4), 112 (9), 211 (7)	11 (28.9), 71 (22), 61 (11.5), 81 (8), 41 (7.4)
4 (4)	33333333 (83.5), 21222222 (12)	11 (29.3), 60 (29.3), 20 (15)	-	71 (25.7), 81 (17.4)
Total info/GR3 population coverage in %	(68.5*0.994) + (13.8*0.992) + (6.6*0.953) + (4*0.995) = **92**	(68.5*0.9966) + (13.8*0.974) + (6.6*0.866) + (4*0.736) = **90.3**	(68.5*0.945) + (13.8 * 0.712) + (6.6*0.351) = **76.9**	(68.5*0.898) + (13.8*0.901) + (6.6*0.778) + (4*0.431) = **81**

As we have seen due to complexity of the data, we are not able to obtain a large coverage of GR3's geographical situation (GS), individual characteristics (INDC), and household characteristics (HHC) profiles. To deal with this, we will effectuate classification trees. We won't be able to capture all combinations of variables but at least we will attempt to find out the most relevant variables and their combinations for a more coverage among GR3. The classification tree method does not produce good prediction when the target response variable has more than 2 categories, and when few regressors are included in fitting the model. In this line, we will use the same target response variable of the model CAP112 (GR1 vs. GR3) and all variables of GS, INDC, and HHC will be included as regressors. The parameters are the same as in the model CAP112. It is important to note that 925 observations were deleted from the data, corresponding to 4.85 per cent of the data[495]. In other words, 326 observations were deleted from GR1 and 599 observations were deleted from GR3. The ratio of sample of GR1 to GR3 remains still the same after deleting some missing values. The new sample comprises

[495] The number of missing values in parenthesis for variables is as follows: DISG (23), OLDG (29), WARG (22), FOSG (22), DEPG (22), AIDG (22), SRG (25), ES1 (478), SEC (66), DWELS (45), INDG (316).

of 7995 observations from GR1 and 10136 observations from GR3 which gives a data of 18131 observations in total (GR1 is 44 per cent and GR3 is 56 per cent of that data). We obtain the following tree named CAP112DEC pruned tree:

```
Classification tree:
rpart(formula = CAP112 ~ GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] AH       DWELS    hh_child HS       INDG     POP      POPH     PRO
[9] SET

Root node error: 5330/12087 = 0.44097

n= 12087

         CP nsplit rel error  xerror      xstd
1 0.3712946      0   1.00000 1.00000 0.0102413
2 0.0453408      1   0.62871 0.62871 0.0092333
3 0.0131332      4   0.49268 0.49512 0.0085212
4 0.0088180      5   0.47955 0.48799 0.0084767
5 0.0065666      6   0.47073 0.48687 0.0084696
6 0.0054409      7   0.46417 0.48330 0.0084469
7 0.0033771     10   0.44784 0.45722 0.0082757
8 0.0028143     11   0.44447 0.45178 0.0082387
9 0.0025000     15   0.43246 0.44934 0.0082219
```

CAP112DEC pruned tree

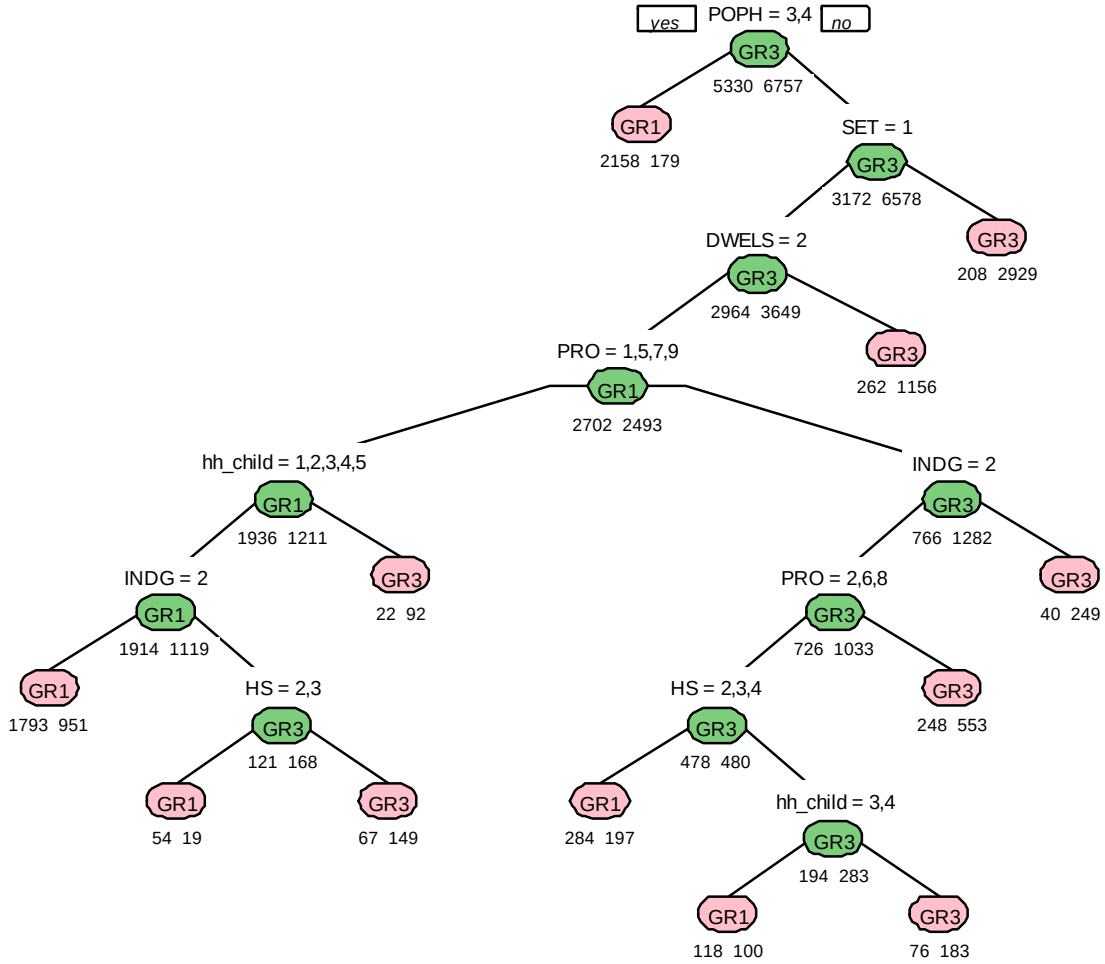


Figure 46. CAP112DEC pruned tree, data: South Africa GHS 2011

From the CAP112DEC pruned tree above, we can see that the following 4 patterns capture almost 70 per cent of the population of GR3: (i) head of household is of origin African/Black/Colored, and household lives in urban informal, or tribal, or rural formal areas (43 per cent of GR3); (ii) head of household is of origin African/Black/Colored, household lives in urban formal areas, and household receives RDP or lives in a subsidized dwelling (17 per cent of GR3); (iii) head of household is of origin African/Black/Colored, household lives in urban formal areas of Eastern or Northern Cape, the household is not indigent and does not live in a subsidized dwelling or does not receive RDP (8 per cent of GR3); and (iv) head of household is of origin African/Black/Colored, household lives in urban formal areas of Eastern Cape, Northern Cape, Free State, North West, or Mpumalanga, household does not receive RDP or does not live in a subsidized dwelling, and household is registered with the local municipality as an indigent household (3.6 per cent of GR3). On the other hand, with the following 2 pattern, 74 per cent of GR1 is detected: (i) head of household is of origin White, or Indian, or Asian (40.5 per cent of GR1); and (ii) head of household is of origin African/Black/Colored,

household loves om urban formal areas of Western Cape, KwaZulu-Natal, Gauteng, or Limpopo, household is not indigent and does not receive RDP or does not live in a subsidized dwelling, and maximum child in the family is 4 (34 per cent of GR1). It is interesting that any variable of individual characteristics (INDC) appears in the tree, which means that geographical situation (GS) and some variables of household characteristics (HHC) are more influential in distinguishing GR1 and GR3. Based on the results, we can remark that GR1 which is prosperous than GR3 is likely composed of non-indigent households headed by White/Indian/Asian people living in urban formal settles with no receiving RDP or any subsidies for dwelling. Moreover, the provinces of Western Cape, KwaZulu-Natal, and Gauteng, containing some of the richest cities of the country, Cape Town, Durban (where mostly India/Asian people live), Johannesburg respectively emerge in the second most important pattern of GR1. To reduce poverty in GR3 but also reduce inequality between GR1 and GR3, we already saw in previous classification trees examination the most important functionings and their combinations that policy makers or government or local authorities should consider in their poverty reduction or quality of life improvement agenda. In addition to these findings, in order to reach out these people of GR3 suffering poverty configurations, the results of CAP112DEC pruned tree can be employed notably in terms of race, geographical context and subsidies. Given that some essential elements of life such as adequate dwelling living standards including drinking water, sanitation (toilet) facilities, roof and wall conditions/materials are emerged as the influential variables in identifying and distinguishing GR3 from GR1, the works on poverty reduction for GR3 can be start off for instance in African/Black/Colored headed households living in urban informal, or tribal, or rural formal settlings. For improving medical coverage, and educational attainment, available number of health facilities, and educational institutions in these settlings scan be also investigated. Regarding the tabulation results of geographical situation presented previously, urban formal settlement appears frequently also among GR3; however, to reduce inequality between GR1 and GR3 and taking in consideration some other variables, priority can be given to non-urban formal settlings. Furthermore, the efforts of reducing poverty and inequality between GR1 and GR3 may contribute in middle or long term reducing racial discrimination or racial inequality in large sense in the country.

In terms of predictive accuracy, the confusion matrix statistics show below that the model can be considered quite good with an overall accuracy of 79.5 per cent.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR3
       GR1 2171  745
       GR3  494 2634

               Accuracy : 0.795
                 95% CI : (0.7846, 0.8051)
    No Information Rate : 0.5591
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5883

 Mcnemar's Test P-Value : 1.226e-12

            Sensitivity : 0.8146
            Specificity : 0.7795
         Pos Pred Value : 0.7445
         Neg Pred Value : 0.8421
             Prevalence : 0.4409
```

```
          Detection Rate : 0.3592
    Detection Prevalence : 0.4825
       Balanced Accuracy : 0.7971

        'Positive' Class : GR1
```

In another model called CAP112DEC2. pruned, we added the variables of GRANT, and EMP as the regressors. We obtain the following results:

```
Classification tree:
rpart(formula = CAP112 ~ DISG + OLDG + WARG + FOSG + DEPG + AIDG +
    SRG + ES1 + SEC + GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
 [1] DWELS    ES1      hh_child HS       INDG     OLDG     POPH     PRO
 [9] SEC      SET

Root node error: 5330/12087 = 0.44097

n= 12087

         CP nsplit rel error  xerror      xstd
1 0.3712946      0   1.00000 1.00000 0.0102413
2 0.0453408      1   0.62871 0.62871 0.0092333
3 0.0115385      4   0.49268 0.51163 0.0086217
4 0.0090056      6   0.46961 0.47674 0.0084048
5 0.0062852      7   0.46060 0.46698 0.0083410
6 0.0045028      9   0.44803 0.45985 0.0082934
7 0.0033771     10   0.44353 0.45403 0.0082540
8 0.0033771     11   0.44015 0.45272 0.0082451
```

CAP112DEC2 pruned tree

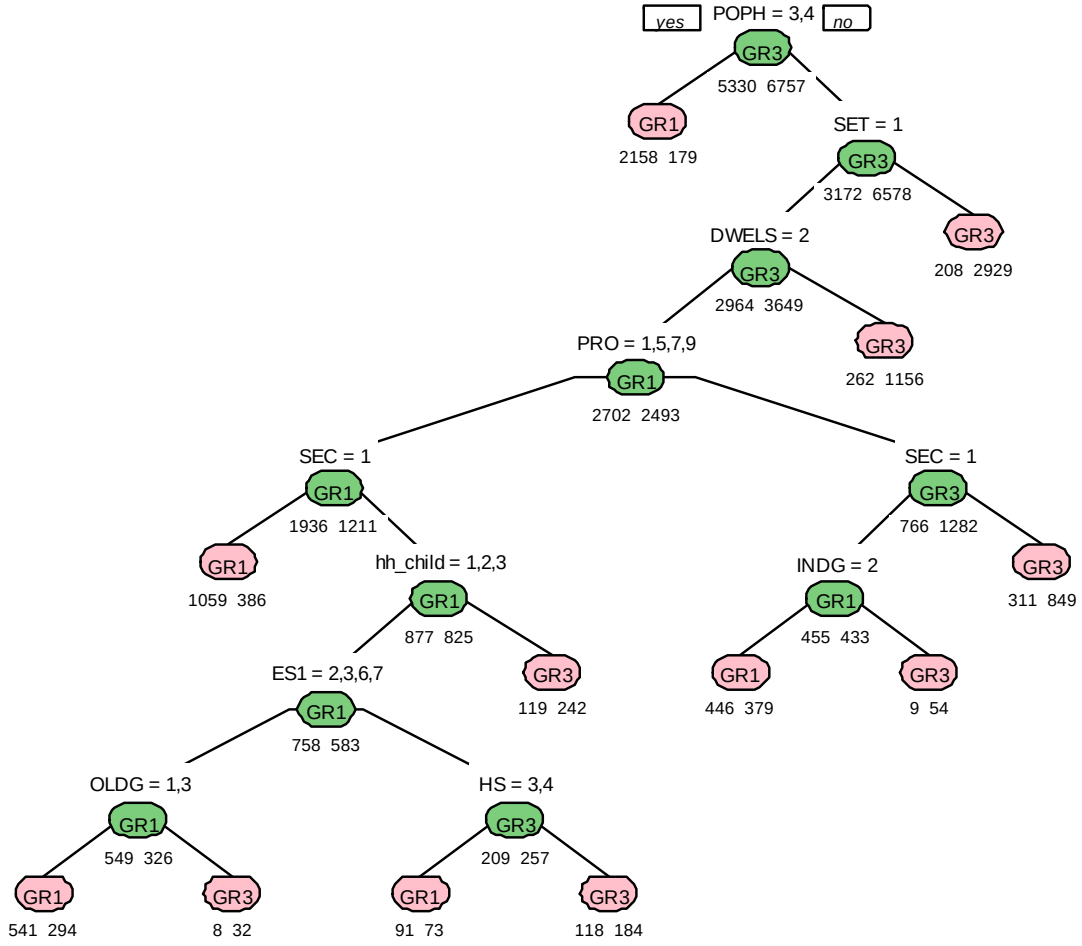


Figure 47. CAP112DEC2 pruned tree, data: South Africa GHS 2011

From the CAP112DEC2 pruned tree we can see that by employing the following 4 patterns, we detect 79 per cent of GR1: (i) households with a head from the origin of White/Indian/Asian (40.5 per cent of GR1); (ii) African/Black/Colored headed households living in a non-subsidized dwelling or not receiving RDP in urban formal areas of Western Cape, KwaZulu-Natal, Gauteng, Limpopo, and working in formal sector (19.9 per cent of GR1); (iii) White/Indian/Asian headed households living in a non-subsidized dwelling or not receiving RDP, in a household of maximum 2 children in urban formal areas of Western Cape, KwaZulu-Natal, Gauteng, Limpopo, working in informal sector or unemployed or not economically active-ES1=2,3,6,or 7 (10 per cent of GR1); and (iv) African/Black/Colored headed households living in non-subsidized or no receiving RDP, not indigent, living in urban formal areas of Eastern Cape, Northern Cape, Free State, Gauteng, Mpumalanga, and working in formal sector (8 per cent of GR1). On the other hand, 73 per cent of GR3 can be identified via the following 3 patterns: (i) African/Black/Colored headed households living in urban informal settlings (43 per cent of GR3); (ii) African/Black/Colored headed households living in a subsidized dwelling or receiving RDP and living in urban formal settlings (17 per cent of GR3); and (iii)

African/Black/Colored headed households living in no-subsidized or no receiving RDP dwellings living in urban formal settings in Eastern Cape, Northern Cape, Free State, North West, Mpumalanga, and working in informal sector (12.5 per cent of GR3). In these results, we can see some impact of employment factors, in particular formal vs. informal sector in the patterns of GR1. Moreover, GR1 is more likely comprising White/Indian/Asian people compared to GR3 which is more likely forming by African/Black/Colored people. In other words, African/Black/Colored people are more likely to be part of more deprived group compared to White/Indian/Asian people appearing in more prosperous situation. GR1 people are more likely to be living in no-subsidized dwelling or no receiving RDP and in urban formal areas compared to GR3. In terms of predictive accuracy, the results are quite close to what we obtained in CAP112DEC pruned tree (see below for confusion matrix statistics).

```
  Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR3
       GR1 2163  701
       GR3  502 2678

               Accuracy : 0.801
                 95% CI : (0.7907, 0.811)
    No Information Rate : 0.5591
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.5994

 Mcnemar's Test P-Value : 1.139e-08

            Sensitivity : 0.8116
            Specificity : 0.7925
         Pos Pred Value : 0.7552
         Neg Pred Value : 0.8421
             Prevalence : 0.4409
         Detection Rate : 0.3579
   Detection Prevalence : 0.4739
      Balanced Accuracy : 0.8021

       'Positive' Class : GR1
```

GR4 sub-group profiles based on CAP113 model

The GRANT, EMP, and GS profiles of GR4 based on the 3 most frequented patterns extracted in CAP113 classification tree analysis are presented in the table below. We can see that through the selected profiles (profiles with at least 50 observations in the data), we arrive to capture 92 per cent of GR4 regarding GRANT, and 93 per cent of GR4 regarding EMP. However, as in the case of GR3 previously, GS profiles are more complex, and we can only achieve to obtain information concerning around 80 per cent of GR4. For that reason, we also created simple cross-table of the two variables, Province and Setting, of the GS for GR4.

Regarding *GRANT*, no receiving any social grant is the most frequent profile concerning 79.4 per cent of the 1st pattern of GR4, or in other words 60 per cent of the entire population of GR4. Receiving only old age grant (OLDG) constitutes around 12 per cent of GR4 and receiving only disability grant includes about 3 per cent of GR4. Thus, with these 3 most frequented profiles from the 1st pattern of GR4, we reach to capture 75 per cent of GR4. The rest, 17 per cent of GR4, are detected through the profiles of 2nd and 3rd patterns of GR4. 82.5 per cent of GR4 population from the 2nd pattern does not receive any social grants; on the other hand, 12 per

cent of them receive old age grant. These two profiles are also found through the 3rd pattern of GR4. These results are quite close to what we obtained for GR3. So that, in total, the most frequented profile is no receiving any social grant concerning 75 per cent of the entire population of GR4, then it is followed by the profile of receiving only old age grant which covers 14 per cent of GR4; and lastly the profile of receiving only disability grant involving 3 per cent of GR4 population in the country. We can say that both GR3 and GR4 have very similar per centages and profiles of GRANT. The most influential item is old age grant, then the second is disability grant, given that the most of the two populations do not receive any social grant.

In terms of *EMP*, people who did not seek a job and would not have liked to work and even if a suitable job had been offered or circumstances had allowed, he or she would not have been able to start work or a business in the last calendar week of the interview emerges as the major group concerning 31.4 per cent of the population of GR4. At second place, active employment status (working for a wage, a commission or a salary; or engaged in business activities; or temporarily absent but will return to work or business; or work without remuneration) with a performance in a formal sector comes covering 18.3 per cent of the population of GR4. Another profile, have sought a job during last 4 weeks and available or ready to start to work, comes at third place concerning 13.3 per cent of the population of GR4. In addition, 12.8 per cent of the population of GR4 did not search for a job but would have liked to work during the last calendar week of the interview; 9 per cent of GR4 is economically active but work in a informal sector; and 6.5 per cent of GR4 did not seek a job and would not have liked to work but if a suitable job had been offered or circumstances had allowed, he or she would have been able to start work or a business in the last calendar week of the interview. From these results, we can clearly see that not been economically active is a part of the majority of the population of GR4. Actively economically active population rate can be considered quite few around 27 per cent; however, 33 per cent of this economically active population performs a job in an informal sector. The numbers we obtain for GR4 are quite at similar level with those of GR3. In this regard, unemployment including discouraged people for searching for a job or for having willing of working, and employment in informal sector flourish as two crucial problems which can generate other social problems such as violence, instability across the country as well as increase poverty severity.

As we have mentioned, the profiles of geographic situation are more complex than those of GRANT, and EMP. With the main GS profiles, we reach to capture around 80 per cent of population of GR4 across the 3 major patterns. First, we will look at the profiles, then we will check with the cross-table. Across the 3 patterns, we obtain the following results at GR4 level: 18 per cent live in tribal areas of Limpopo; 9 per cent live in tribal areas of KwaZulu-Natal; 8 per cent live in tribal areas of Mpumalanga; 6.8 per cent live in tribal areas of Eastern Cape; 6 per cent live in urban formal areas of Gauteng; 5.6 per cent live in urban formal areas of Free State; 5.5 per cent live in urban formal areas of Western Cape; 5.3 per cent live in tribal areas of North West; 4 per cent live in urban formal areas of North West; 4 per cent live in urban formal areas of Northern Cape; 2.5 per cent live in urban formal areas of KwaZulu-Natal. Based on the results, we can see that the per centage of people living in tribal areas compared to urban formal is higher. Notably, in case of GR3, it was inverse-per centage of living in urban formal areas is higher than those of tribal areas.

Among GR4, (i) half of them live in tribal areas, and 40 per cent of them live in urban formal areas; (ii) 19 per cent of them live in Limpopo, 15 per cent of them live in KwaZulu-Natal; 11.

8 per cent of them live in Mpumalanga; 11.5 per cent of them live in Eastern Cape; and 11 per cent of them live in North West. Compared to GR3, there is less population of GR4 live in the province of Gauteng. In terms of population of GR4 living in urban formal areas, Western Cape, Free State, and Gauteng have the highest scores; on the other hand, clearly Limpopo is the province where the most of population of GR4 lives in tribal areas, and in Western Cape, there is no population living in tribal areas (and in Gauteng, Northern Cape, and Free State, the ratio of GR4 people living in tribal areas is substantially low). In this sense, the results we obtain for Limpopo and Western Cape are similar to the case of GR3. In terms of urban informal areas, Gauteng and KwaZulu-Natal have the highest scores as in GR3. And, regarding rural formal areas, Free State and North West have the highest population of GR4. Compared to GR3, the province of Free State emerges more important among population of GR4, notably for living in urban formal areas. We can see that there is a shift from Western, Northern Cape to Free State and North West concerning living in rural formal areas. Furthermore, in order to reduce poverty among GR4, the 12 locations (see the table below) appearing in the 1st pattern can be put on priority.

Table 72. Sub-group profiles of GR4 based on the CAP113 model, data: South Africa GHS 2011

Pattern number (Freq in % in full data)	GRANT [496]profile (Freq in % in full data)	EMP[497] profile (Freq in % in full data)	GS profile (Freq in % in full data)	INDC profile (Freq in % in full data)
1 (76)	33333333 (79.4), 21222222 (15.5), 12222222 (4.4)	60 (34.6), 11 (18.9), 40 (13.8), 20 (12.6), 12 (10), 50 (7.6), 70 (2)	94 (23.4), 54 (12), 84 (10.6), 24 (8.9), 64 (7), 41 (5.3), 71 (4.9), 61 (3.6), 31 (3.4), 51 (2.3), 11 (2.15), 81 (2)	112 (13), 211 (12), 111 (12), 212 (10.6), 113 (8.4), 114 (8.2), 116 (7.5), 213 (6.5), 115 (6.45), 214 (4.8), 215 (4), 216 (3.7)
2 (12)	33333333 (82.5), 21222222 (12),	60 (26.6), 20 (23), 11 (22.7), 40 (13), 12 (6), 50 (6)	11 (30.7), 21 (12.5), 71 (11.5), 61 (11.4), 41 (8.6), 31 (6.3), 51 (6)	111 (8.9), 211 (8.6), 112 (7.6), 212 (7.5), 113 (6), 114 (5.9), 213 (5)
3 (6)	33333333 (79.3), 21222222 (12.3)	60 (32), 11 (20), 20 (17), 40 (12), 12 (10)	11 (24), 71 (13.8), 31 (12.9), 41 (10.3), 21 (9.4)	-

[496] 20 individuals were dropped from the data due to missing values regarding all variables of *GRANT* dimension.

[497] 340 individuals were dropped from the data due to missing values regarding all variables of *EMP* dimension.

Total info/GR4 population coverage in %	(76*0.993) + (12*0.945) + (6*0.916) = **92**	(76*0.995) + (12*0.973) + (6*0.91) = **93**	(76*0.855) + (12*0.87) + (6*0.704) = **79.6**	(76*0.97) + (12*0.495) = **80**

For the same reasons mentioned in the case of GR3 3rd model analysis, we will here also produce 2 classification trees where the target response variable is those of the model CAP113 (GR1 vs. GR4) and all variables of GS, INDC, and HHC will be included as regressors. The parameters are the same as in the model CAP113. It is important to note that 877 observations were deleted from the data, corresponding to 5 per cent of the data[498]. In other words, 326 observations were deleted from GR1 and 551 observations were deleted from GR4. The ratio of sample of GR1 to GR4 remains approximately the same after deleting some missing values. The new sample comprises of 7995 observations from GR1 and 8669 observations from GR4 which gives a data of 16664 observations in total (GR1 is 48 per cent and GR4 is 52 per cent of that data). We obtain the following tree named CAP113DEC pruned tree:

```
Classification tree:
rpart(formula = CAP113 ~ GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] AH       DWELS    hh_child HS       INDG     POP      POPH     PRO
[9] SET

Root node error: 5330/11109 = 0.47979

n= 11109

         CP nsplit rel error  xerror      xstd
1 0.5088180      0   1.00000 1.00000 0.0098793
2 0.1121951      1   0.49118 0.49118 0.0083927
3 0.0127580      2   0.37899 0.37692 0.0076111
4 0.0095059      3   0.36623 0.35591 0.0074413
5 0.0058161      8   0.31839 0.32214 0.0071482
6 0.0037523     10   0.30675 0.31257 0.0070604
7 0.0037523     12   0.29925 0.31276 0.0070621
```

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR4
       GR1 2286  478
       GR4  379 2412

               Accuracy : 0.8457
                 95% CI : (0.836, 0.8551)
    No Information Rate : 0.5203
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.6914

 Mcnemar's Test P-Value : 0.0008151
```

[498] The number of missing values in parenthesis for variables is as follows: DISG (21), OLDG (4), WARG (1), AIDG (1), SRG (3), ES1 (439), SEC (67), DWELS (27), INDG (314).

```
            Sensitivity : 0.8578
            Specificity : 0.8346
         Pos Pred Value : 0.8271
         Neg Pred Value : 0.8642
             Prevalence : 0.4797
         Detection Rate : 0.4115
   Detection Prevalence : 0.4976
      Balanced Accuracy : 0.8462

       'Positive' Class : GR1
```

From the CAP113DEC pruned tree, we can observe that the 3 following patterns construct 83 per cent of the population of GR1: (i) living in urban formal areas in a no-subsidized or no receiving RDP household and the head of household is of origin Indian/Asian/White (38 per cent of GR1); (ii) living in urban formal areas in a no-subsidized or no receiving RDP household with maximum 3 children in the provinces of Western Cape, KwaZulu-Natal, Gauteng, Limpopo, and the head of household is of origin African/Black/Colored (35 per cent of GR1); (iii) living in urban formal areas in a no-indigent, no-subsidized or no receiving RDP household with maximum 4 children in the provinces of Eastern Cape, Free State, North West, Mpumalanga, and the head of household is of origin African/Black/Colored and the age of the head of the household is between 25 and 34 (9.5 per cent of GR1). We can see that GR1 is more likely living in urban formal areas without receiving subsidies, or RDP compared to GR4. On the other hand, with the following 6 patterns we arrive to capture 84 per cent of GR4: (i) African/Black/Colored/Indian/Asian people living in urban informal or tribal, or rural formal areas (59.5 per cent of GR4); (ii) living in a RDP or state subsidized dwelling in urban formal areas in the provinces of Eastern Cape, Northern Cape, Free State, North West, Limpopo (9 per cent of GR4); (iii) living in a RDP or state subsidized dwelling in urban formal areas in the provinces of Western Cape, KwaZulu-Natal, Gauteng, Mpumalanga (4.6 per cent of GR4); (iv) living in an indigent, RDP or state subsidized dwelling in urban formal areas in the provinces of Eastern Cape, Northern Cape, Free State, North West, Mpumalanga and the head f household is of origin African/Black or Colored (3.3. per cent of GR4); (v) living in an indigent, RDP or state subsidized dwelling with maximum 4 children in urban formal areas in the provinces of Eastern Cape, Northern Cape, Free State, North West, Mpumalanga and the head of household is of origin African/Black/Colored and the age group of head of the household is less than 18, or 18-24, or 25-34, or 65+ (4 per cent of GR4); and (vi) living in an indigent, no RDP or no state subsidized dwelling with 5 or more children in urban formal areas in the provinces of Eastern Cape, Free State, North West, Mpumalanga and the head of household is of origin African/Black/Colored and the age of the head of the household is between 35 and 64 (4 per cent of GR4). The first pattern emerging in identifying GR4 is the most important one which in a most influential way distinguishes GR4 compared to GR1. So that, racial and geographical segregation should be prominently considered in determining how to reduce poverty among GR4 and notably reducing inequality between GR1 and GR4. In addition, we can also see that GR4 households are more likely to live in an RDP or state subsidized dwelling compared to GR1.

Concerning the predictive accuracy of this model, the overall accuracy is around 85 per cent and sensitivity and specificity rates are quite close so that misclassification rates for both GR1 and GR4 are quite similar.

In another model called CAP113DEC2. pruned, we added the variables of GRANT, and EMP as the regressors. We obtain the following results:

```
Classification tree:
rpart(formula = CAP113 ~ DISG + OLDG + WARG + FOSG + DEPG + AIDG +
    SRG + ES1 + SEC + GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] DWELS    hh_child HS       POP      POPH     PRO      SEC      SET

Root node error: 5330/11109 = 0.47979

n= 11109

         CP nsplit rel error  xerror      xstd
1 0.5088180      0   1.00000 1.00000 0.0098793
2 0.1121951      1   0.49118 0.49118 0.0083927
3 0.0156973      2   0.37899 0.37542 0.0075992
4 0.0127580      5   0.33189 0.33058 0.0072239
5 0.0037523      6   0.31914 0.31501 0.0070830
6 0.0033771      8   0.31163 0.31576 0.0070899
7 0.0031895      9   0.30826 0.31276 0.0070621
8 0.0030957     10   0.30507 0.31276 0.0070621
```

CAP113DEC pruned tree

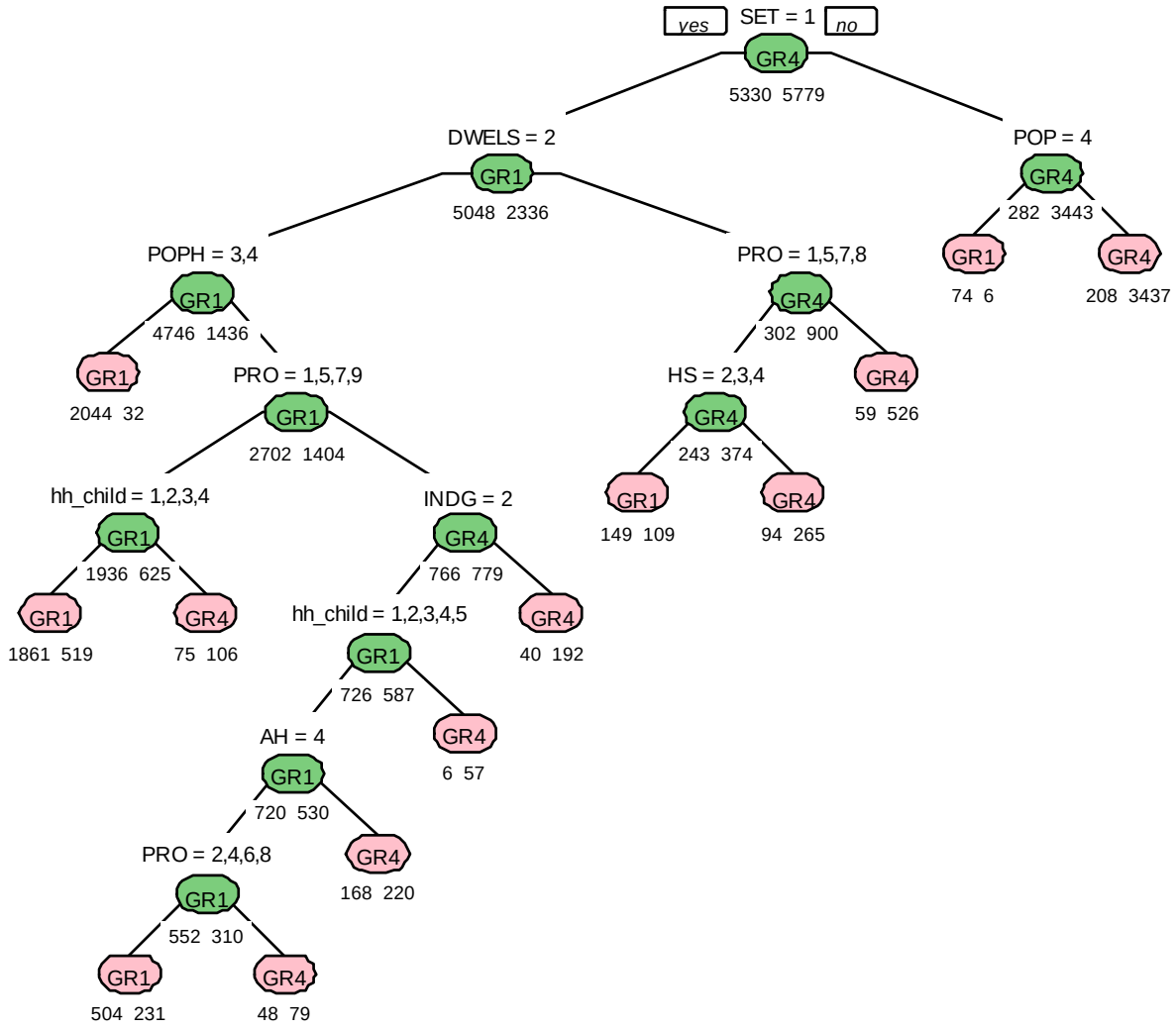


Figure 48. CAP113DEC pruned tree, data: South Africa GHS 2011

CAP113DEC2 pruned tree

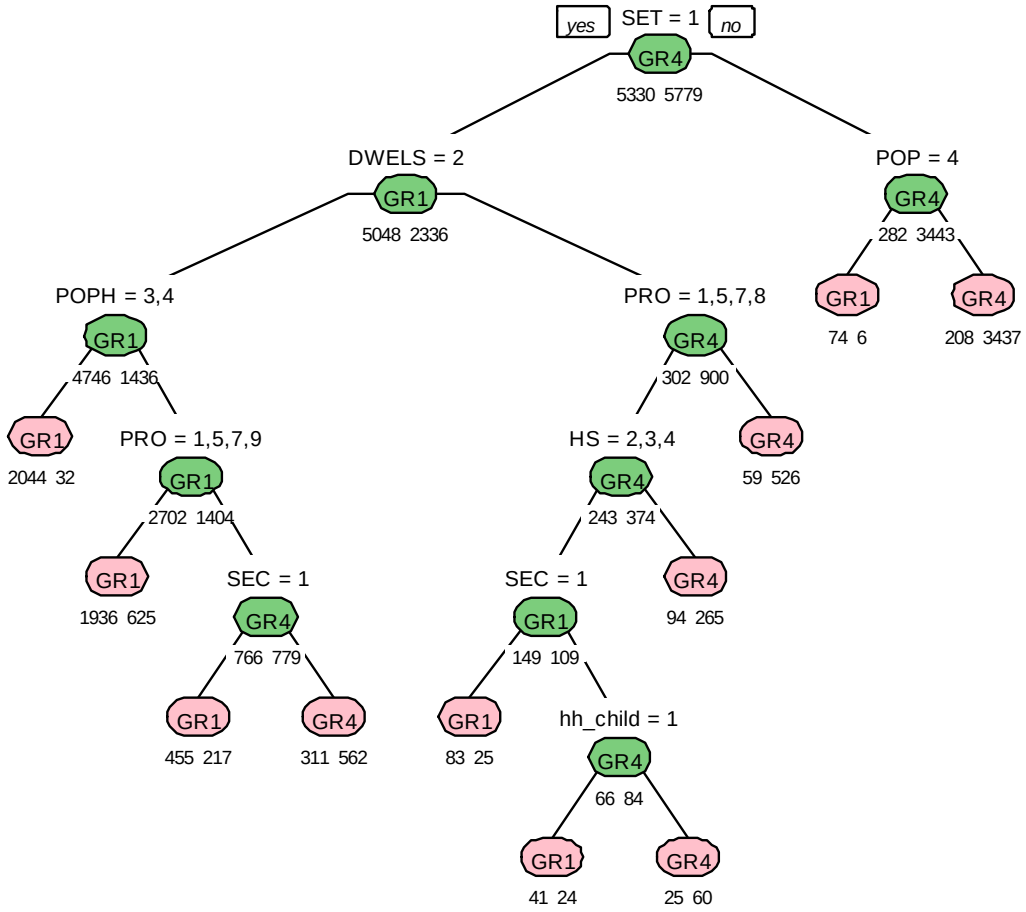


Figure 49. CAP113DEC2 tree, data: South Africa GHS2011

From CAP113DEC2 pruned tree, we can see that the 3 following patterns capture 83 per cent of GR1: (i) living in a non RDP or a non-state subsidized dwelling in urban formal areas and the head of the household is of origin White/Indian/Asian (38 per cent of GR1); (ii) living in a non RDP or a non-state subsidized dwelling in urban formal areas in the provinces of Western Cape, KwaZulu-Natal, Gauteng, Limpopo, and the head of the household is of origin African/Black/Colored (36 per cent of GR1); (iii) working in a formal sector, and living in a non RDP or a non-state subsidized dwelling in urban formal areas in the provinces of Eastern Cape, Northern Cape, Free State, North West, Mpumalanga, and the head of the household is of origin African/Black/Colored (8.5 per cent of GR1). As in the case of CAP113DEC, here also, we can observe that GR1 is mainly characterized by living in a non-RDP or a non-state subsidized dwelling in urban formal areas. In addition, working in a formal sector appears for 8.5 per cent of GR1 in the tree. The GRANT features and individual characteristics (INDC) do not emerge in the tree. On the other hand, with the 4 following patterns, 83 per cent of GR4 are detected: (i) living in urban informal, tribal, rural formal areas and of being origin African/Black/Colored/Indian/Asian (59.5 per cent of GR1); (ii) working in informal sector and living in a non RDP or a non-state subsidized dwelling in urban formal areas of the provinces

of Eastern Cape, Northern Cape, Free State, North West, Mpumalanga and being origin of African/Black/Colored (9.7 per cent of GR4); (iii) living in a RDP or a state subsidized dwelling in urban formal areas in the provinces of Eastern Cape, Northern Cape, Free State, North West (9 per cent of GR4); and (iv) living in a RDP or a subsidized dwelling with 1 or more than 5 persons in urban formal areas in the regions of Western Cape, KwaZulu-Natal, Gauteng, Mpumalanga. These results show that compared to GR1, a considerably important part of GR4 live in a non-urban formal area and they are not White. Similar to the results of CAP113DEC, here also geographical and racial segregation come to light. In opposition to formal sector appeared in the pattern of GR1, working in informal sector flourishes in the 2nd pattern of GR4 affecting almost 10 per cent of that population.

Concerning the predictive accuracy of the model, as the confusion statistics show below, the overall accuracy is 85 per cent and we capture quite similar amount of information for both GR1 and GR4.

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR4
       GR1 2318  476
       GR4  347 2414

               Accuracy : 0.8518
                 95% CI : (0.8422, 0.8611)
    No Information Rate : 0.5203
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7038

 Mcnemar's Test P-Value : 8.127e-06

            Sensitivity : 0.8698
            Specificity : 0.8353
         Pos Pred Value : 0.8296
         Neg Pred Value : 0.8743
             Prevalence : 0.4797
         Detection Rate : 0.4173
   Detection Prevalence : 0.5030
      Balanced Accuracy : 0.8525

       'Positive' Class : GR1
```

GR5 sub-group profiles based on CAP114 model

As we have seen previously, we could be able to capture 97 per cent of GR5 based on the 2 patterns emerged in the CAP114 classification tree analysis. First pattern only itself detects 93 per cent, and the second gathers 4 per cent of the GR5. As a reminder, we display only profiles with at least 50 observations in the data. The GRANT, and EMP profiles shown in the table below concern at least 95 per cent of the entire population of GR5. For individual characteristic (INDC), we reach to capture 90 per cent of GR5 (which is much higher compared to GR3 and GR4). So that, in terms of INDC, GR5 has more homogenous profiles compared to GR3 and GR4. Moreover, for geographical situation (GS), poorer performance is emerged compared to GR3 and GR4, which means that GS profiles are more heterogenous in case of GR5.

Concerning *GRANT*, 78 per cent of the entire population of GR5 do not receive any social grant; 12.5 per cent of them receive only old age grant, and 4.65 per cent of them receive only

disability grant. These results are very similar compared to what we obtained previously for GR3 and GR4. We can say that the way that GR3, GR4, and GR5 distinguish from GR1 is quite similar in terms of grant, based on the most frequent capabilities patterns extracted in CAP112, CAP113, and CAP114 respectively. In other words, the grant profiles associated with the most frequented capabilities patterns of GR3, GR4, and GR5 remain similar, with basically two most received types of grant which are old age and disability grants given that a substantial part of them do not receive any social grant.

When we look at EMP profiles across the two patterns, we can observe that 32 per cent of the entire population of GR5 did not seek a job and he or she would not have liked to work and even if a suitable job had been offered or circumstances had allowed, and he or she would not have been able to start work or a business in the last calendar week of the interview. This profile emerges as the most frequented one among GR3, GR4, GR5 profiles with similar per centage around 30. In addition, (i) 17 per cent of GR5 did not search but would have liked to work during the last calendar week; (ii) 15 per cent of GR5 is employed in formal sector working for a wage, a commission or a salary, or engaged in business activities or temporarily absent but will return to work or business, or work without remuneration; (iii) 14 per cent of GR5 sought a job during the last 4 weeks and available or ready to start to work (unemployment rate); (iv) 10.5 per cent of GR5 is economically active (employed) and working in informal sector; (v) 6 per cent of GR5 did not seek a job and would not have liked to work but if a suitable had been offered or circumstances had allowed, he or she would have been able to start work or a business in the last calendar week of the interview. Thus, the rate of economically active population is low, which is around 25.5 per cent and around 40 per cent of that population works in informal sector.

When it comes to individual characteristics (INDC), with 12 profiles shown below in the table, we reach 90 per cent of GR5. The most frequented profile is being female, African/Black origin, and between 18 and 24 yrs. old, which concerns 11. 7 per cent of GR5. After that, the following profiles come respectively with per centages indicated in parenthesis: male & African/Black & 18-24 yrs. (10.8); female & African/Black & 25-34 yrs. (10.7).; male & African/Black & 25-34 yrs. (9.95); female & African/Black & 35-44 yrs. (8.9); female & African/Black & 45-54 yrs.(7.7); male & African/Black & 35-44 yrs. (6.5); female & African/Black & 65+ yrs. (6); female & African/Black & 55-64 yrs. (5.6); male & African/Black & 45-54 yrs. (5.1); male & African/Black & 55-64 yrs. (3.5); male & African/Black & 65+ yrs. (3). Based on the 2 patterns, the INDC profile results show that (i) 90 per cent of GR5, the poorest group in the country, is composed of African/Black people; (ii) 40 per cent of GR5 is male population vs. 60 per cent of them is female population; (iii) the group of 18-34 yrs. forms 44 per cent vs. 35+ forms 46 per cent of GR5. Poverty reduction programs should take in consideration racial segregation but also youth population is very important for future generations in the country. Notably, higher economic activity rate among youth population and across country should be targeted to combat poverty.

Concerning geographical situation (GS), the most frequent 2 profiles are: living in tribal areas in Eastern Cape (17 per cent of GR5); and living in tribal areas in KwaZulu-Natal (15 per cent of GR5). After that, living in tribal areas in Limpopo, tribal areas in North West, tribal areas in Mpumalanga arrive with 8.3, 6.5, and 5.3 per centages respectively. Living in tribal areas constructs therefore 37 per centage of GR5 based on the information extracted through the 2 patterns. On the other hand, living in urban formal areas concern 21 per cent of GR5. Per

centages of the rest of the profiles are quite low, under 4 per cent each of them. If we look at the cross-table of GS among GR5 population, we can see that 57.5 per cent of them live in tribal areas, and 24 per cent of them live in urban formal areas. In terms of provinces, Eastern Cape and KwaZulu-Natal have the highest rates of GR5 population. Notably, living conditions of people living in tribal areas of these 2 provinces should be urgently taken in consideration for ameliorating living standards of the poorest group of the country.

Table 73. Sub-group profiles of GR5 based on the CAP114 model, data: South Africa GHS 2011

Pattern number (Freq in % in full data)	GRANT [499]profile (Freq in % in full data)	EMP[500] profile (Freq in % in full data)	GS profile (Freq in % in full data)	INDC profile (Freq in % in full data)
1 (93)	33333333 (80.5), 21222222 (13.5), 12222222 (5)	60 (33.5), 40 (17.7), 11 (15.5), 20 (14.25), 12 (11.3), 50 (6), 70 (1.56)	24 (18.3), 54 (16), 94 (9), 64 (7), 84 (5.7), 41 (4.3), 61 (3.7), 72 (3.5), 31 (3.2), 71 (2.6), 11 (2.5), 21 (2.3), 22 (2.2), 44 (2)	111 (12), 211 (11.7), 112 (11.5), 212 (10.7), 113 (9.6), 114 (8.3), 213 (7), 116 (6.5), 115 (6), 214 (5.5), 215 (3.8), 216 (3.3)
2 (4)	33333333 (85),	20 (28.6), 60 (27), 11 (14.55), 40 (14.55)	31 (18)	111 (13)
Total info/GR5 population coverage in %	(93*0.99) + (4*0.85) = **95.5**	(93*0.998) + (4*0.847) = **96**	(93*0.823) + (4*0.18) = **77**	(93*0.959) + (4*0.13) = **90**

We will now proceed with 2 classification trees where the target response variable is those of the model CAP114 (GR1 vs. GR5) and all variables of GS, INDC, and HHC will be included as regressors. The parameters are the same as in the model CAP114. It is important to note that 906 observations were deleted from the data, corresponding to 4.9 per cent of the data[501]. In other words, 326 observations were deleted from GR1 and 580 observations were deleted from GR4. The ratio of sample of GR1 to GR5 remains approximately the same after deleting some

[499] 25 individuals were dropped from the data due to missing values regarding all variables of *GRANT* dimension.

[500] 376 individuals were dropped from the data due to missing values regarding all variables of *EMP* dimension.

[501] The number of missing values in parenthesis for variables is as follows: DISG (20), OLDG (6), WARG (1), AIDG (1), SRG (1), ES1 (482), SEC (65), DWELS (25), INDG (304).

missing values. The new sample comprises of 7995 observations from GR1 and 9699 observations from G54 which gives a data of 17694 observations in total (GR1 is 45 per cent and GR5 is 55 per cent of that data). We obtain the following tree named CAP114DEC pruned tree:

```
Classification tree:
rpart(formula = CAP114 ~ GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] DWELS    hh_child INDG     POP      POPH     PRO      SET

Root node error: 5330/11796 = 0.45185

n= 11796

         CP nsplit rel error  xerror      xstd
1 0.6472795      0   1.00000 1.00000 0.0101411
2 0.0306754      1   0.35272 0.35272 0.0074585
3 0.0135084      3   0.29137 0.30056 0.0069809
4 0.0072233      4   0.27786 0.25929 0.0065534
5 0.0054409      6   0.26341 0.25422 0.0064975
6 0.0050657      9   0.24709 0.24991 0.0064492
7 0.0034709     10   0.24203 0.24672 0.0064131
8 0.0034709     12   0.23508 0.24296 0.0063702
```

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR5
       GR1 2427  438
       GR5  238 2795

               Accuracy : 0.8854
                 95% CI : (0.877, 0.8934)
    No Information Rate : 0.5482
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7701

 Mcnemar's Test P-Value : 1.951e-14

            Sensitivity : 0.9107
            Specificity : 0.8645
         Pos Pred Value : 0.8471
         Neg Pred Value : 0.9215
             Prevalence : 0.4518
         Detection Rate : 0.4115
   Detection Prevalence : 0.4858
      Balanced Accuracy : 0.8876

       'Positive' Class : GR1
```

From CAP114DEC tree, we can see that with the following 3 patterns, 88 per cent of GR1 can be captured: (i) living in a no RDP or non-state subsidized dwelling in urban formal areas in Western Cape, Eastern Cape, KwaZulu-Natal, Gauteng, Mpumalanga, Limpopo and the head of the household is African/Black/Colored (43 per cent of GR1); (ii) living in urban formal areas and the head of household is Indian/Asian/White (39 per cent of GR1); and (iii) living in a no indigent, no RDP or non-state subsidized dwelling with maximum 4 children in urban formal areas of Free State, North West and the head of the household is African/Black/Colored

(6 per cent of GR1). On the other hand, 83 per cent of GR5 can be detected by the following 3 patterns: (i) living in urban informal, tribal, rural formal areas and being African/Black/Colored/Indian/Asian (75 per cent of GR5); (ii) living in a RDP or state-subsidized dwelling in urban formal areas of North Cape, Free State, North West and the head of household is African/Black/Colored (5 per cent of GR5); and (iii) living in a no RDP or state-subsidized dwelling in urban formal areas of North Cape, Free State, North West, and the head of the household is African/Black/Colored (2.25 per cent of GR5). As we can see, the poorest group distinguishes from the least poor group notably in terms of racial and geographical segregation. An African/Black/Colored/Indian/Asian living in a non-urban formal area are most likely at risk of being in GR5 compared to GR1. Regarding confusion matrix statistics the overall accuracy of the model is 88.5 per cent and misclassification rates are 8 per cent or GR1 and 13 per cent for GR5 in the training data.

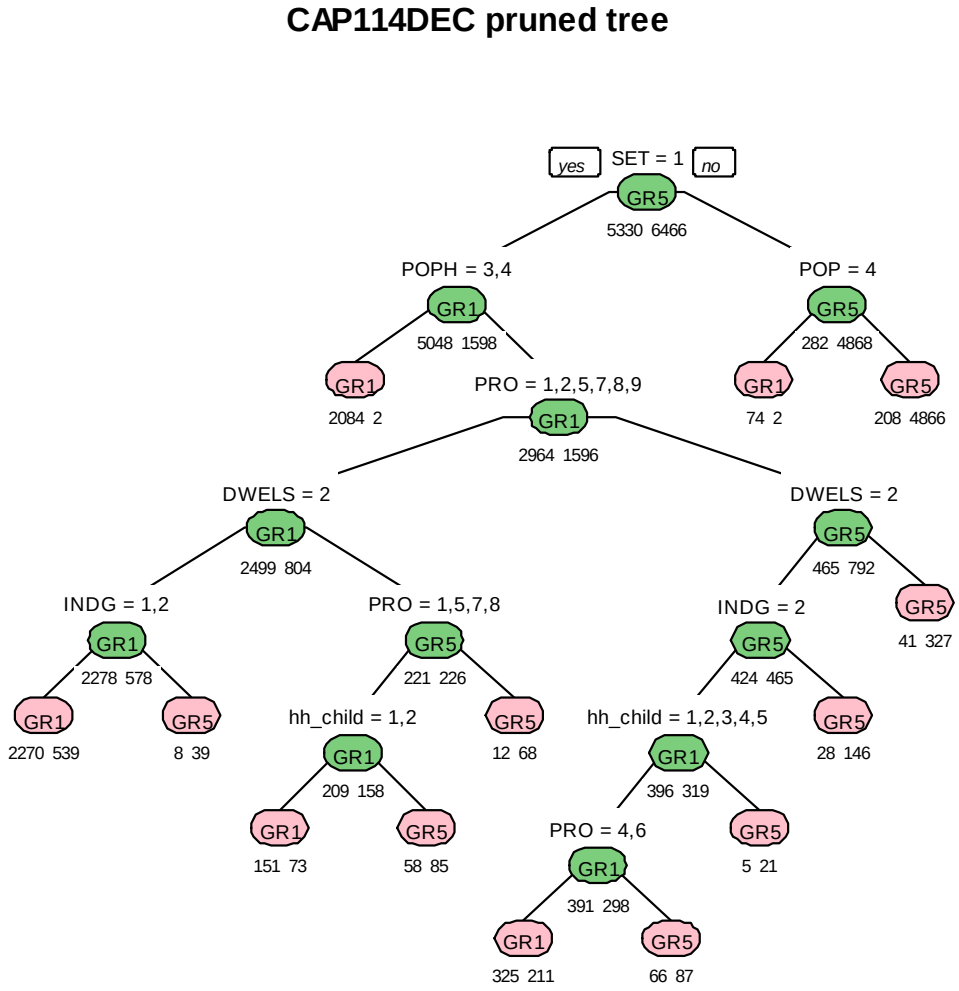


Figure 50. CAP114DEC tree, data: South Africa GHS 2011

In another model called CAP114DEC2. pruned, we added the variables of GRANT, and EMP as the regressors. We obtain the following results:

```
Classification tree:
rpart(formula = CAP114 ~ DISG + OLDG + WARG + FOSG + DEPG + AIDG +
    SRG + ES1 + SEC + GEN + POP + AGE + DWELS + HS + hh_child +
    GH + AH + POPH + INDG + PRO + SET, data = train.set, method = "class",
    parms = list(split = "gini"), control = temp)

Variables actually used in tree construction:
[1] DWELS INDG  POP   POPH  PRO   SEC   SET

Root node error: 5330/11796 = 0.45185

n= 11796

         CP nsplit rel error  xerror      xstd
1 0.6472795      0   1.00000 1.00000 0.0101411
2 0.0306754      1   0.35272 0.35272 0.0074585
3 0.0264540      3   0.29137 0.28274 0.0068022
4 0.0135084      4   0.26492 0.26266 0.0065903
5 0.0077861      5   0.25141 0.24916 0.0064407
6 0.0041276      7   0.23583 0.23752 0.0063072
7 0.0033771      8   0.23171 0.23377 0.0062631
8 0.0025000      9   0.22833 0.22908 0.0062073
```

From the CAP114DEC2 below, we can see that 84 per cent of GR1 is detected through the 3 following patterns: (i) living in urban formal areas and the head of the household is Indian/Asian/White (39 per cent of GR1); (ii) working in formal sector, living in urban formal areas of Western Cape, Eastern Cape, KwaZulu-Natal, Gauteng, Mpumalanga, Limpopo, and the head of the household is African/Black/Colored (26 per cent of GR1); and (iii) working in informal sector, living in urban formal areas of Western Cape, Eastern Cape, KwaZulu-Natal, Gauteng, Mpumalanga, Limpopo, and the head of the household is African/Black/Colored (19 per cent of GR1). On the other hand, 85 per cent of GR5 can be captured via the following 2 patterns: (i) being African/Black/Colored/Indian/Asian, and living in urban informal, tribal, rural formal areas (75 per cent of GR5); and (ii) working in informal sector, living in urban formal areas of North Cape, Free State, North West, and the head of the household is African/Black/Colored (10 per cent of GR5). We arrive here also at the same main result: geographical and racial segregation play crucial role in elevating quality of life of the poorest in the country. African/Black/Colored/Indian/Asian people living in a non-urban formal area are at most risk of being poorest. Third pattern of GR1 and the 2nd pattern of GR5 have the same characteristics in terms of working in informal sector, settling, and the population group of the head of the household; nevertheless, regarding provinces there is difference. Notably, GR5 people in this pattern are appearing living in Northern Cape, Free State, and North West provinces. Moreover, 19 per cent of GR1 which is less poor group experience working in informal sector. Issues concerning employment sector should be also in the agenda of poverty reduction programs in the country for improving quality and notably safety at work. Regarding quality of the model, the predictive accuracy is around 90 per cent, and sensitivity and specificity rates are also very close.

CAP114DEC2 tree

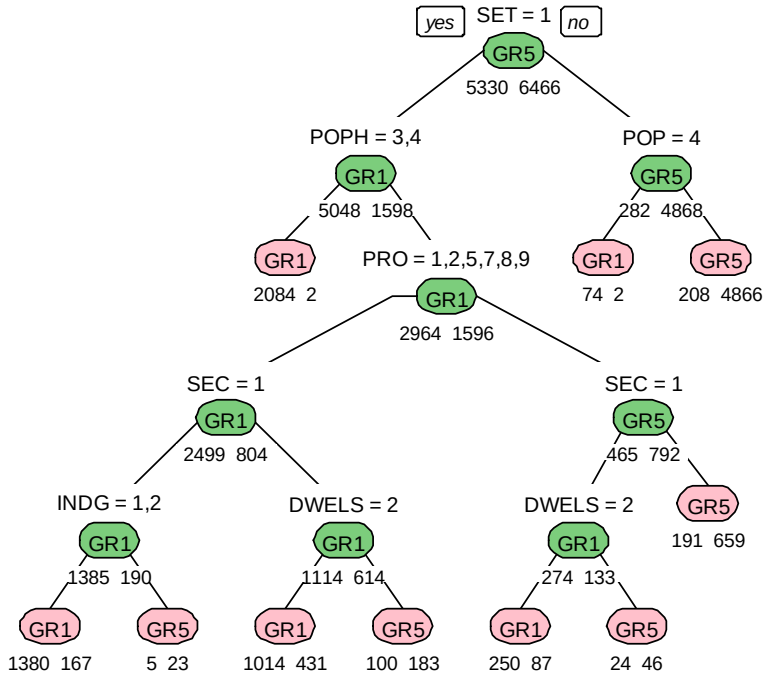


Figure 51. CAP114DEC2 tree, data: South Africa GHS 2011

```
Confusion Matrix and Statistics

          Reference
Prediction  GR1  GR5
       GR1 2395  341
       GR5  270 2892

               Accuracy : 0.8964
                 95% CI : (0.8883, 0.9041)
    No Information Rate : 0.5482
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.7914

 Mcnemar's Test P-Value : 0.004627

            Sensitivity : 0.8987
            Specificity : 0.8945
         Pos Pred Value : 0.8754
         Neg Pred Value : 0.9146
             Prevalence : 0.4518
         Detection Rate : 0.4061
   Detection Prevalence : 0.4639
      Balanced Accuracy : 0.8966

       'Positive' Class : GR1
```

5.5 Fourth step (step4 analysis): detecting multidimensional poverty profiles of specific groups

In this section, the focus will be on children (0-17 yrs. old)[502], disabled, elderly, unemployed, the most prosperous (no deprivation), poverty group 1 (DEP_GRS1), poverty group 2 (DEP_GRS2) populations as well as some descriptive statistics regarding some dwelling and health related variables which were not used in constructing overall deprivation index. It is important to note that (i) the non-weighted numbers (i.e., no estimation at national level by using weighting function) are presented in the following descriptive statistics; and (ii) computations are based on the same data used for building overall deprivation index and poverty analysis above; in other words, the data is without missing values.

Children

We start off by children's situation in terms of poverty groups, school, and nutrition. There are in total 31496 individuals living in a household with at least 1 child and headed by a person aged 18 yrs. or more. These individuals can be classified in 5 groups: (i) 32.6 per cent of them live in a household with 1 child; (ii) 29 per cent live in a household with 2 children; (iii) 18 per cent live in a household with 3 children; (iv) 10 per cent live in a household with 4 children; and (v) 10.3 per cent live in a household with at least 5 children. If we look at the distribution of number of children in household (hh_child variable) by 5 poverty groups (DEP_GRS variable), we find that: regarding individuals living in a household with 1 child: 23 per cent of them are in DEP_GRS3; 20 per cent are in DEP_GR1; 19.5 per cent are in DEP_GR2; 18.7 per cent are in DEP_GRS4; and 18.6 per cent are in DEP_GRS5. Based on the results, we can observe that from individuals living in a household with 1 child to those living in a household with 3 children, the most frequented situation is being in DEP_GR3 given that the latter two ones experience GR3, GR4, GR5, GR2, and GR1 respectively whilst individuals living in a household with 1 child are in GR3, GR1, GR2, GR4, and GR5 respectively. The poverty situation exacerbates for individuals living in a household with at least 4 children: they are in GR5, GR4, GR3, GR2, and GR1 respectively. In other words, individuals living in a household with 1 child are living mostly in better conditions compared to others, and individuals living in households with 2 or 3 children can be seen as in the middle being mostly in GR3[503].

---

[502] Please see the convention on the rights of the child in the link below (accessed 3 November 2019): https://www.ohchr.org/en/professionalinterest/pages/crc.aspx

[503] Here, we cannot provide statistics of children by deprivation groups because deprivation groups are based on 18+ individual level overall deprivation index whilst children are affiliated with a household. And, within the same household members may be affiliated with a different deprivation group due to some individual level variables such as educational attainment, medical coverage. To know directly children are living in which type of household in terms of deprivation, the overall deprivation index should be computed at household level and not at individual level. For a comprehensive child poverty study, a specific survey focusing in particular on the needs and living standards of children should be used. Another alternative is, constructing deprivation groups based on only the household level variables, and then discovering children's profiles. Some poverty studies, in order to include at the same time both children and adults, use household level variables and transform information at individual level to household level. For instance, highest educational attainment concerns each adult but at household level, this information can be used to know whether there is at least one person deprived of standard obligatory educational attainment. In such case, every member in the household will have a same score for educational attainment whereas not everyone in the household is deprived regarding that variable. So that, there are some inconvenient to do poverty analysis at household level whilst some indicators are at individual level. On the other hand, for a global view on poverty at national level

In the data we use (no missing values for overall deprivation index), there are 27,372 children. 80 per cent of them never went hungry, 11 per cent of them went sometimes hungry, 5.7 per cent of them went seldom hungry, 2 per cent of them went often hungry, and 0.7 per cent always went hungry because there was not enough food in the household (in the past 12 months of the interview time).

In terms of school attendance, there are in total 749 children aged 5-17 yrs. and are not attending school/educational institution at the time of the interview[504]. The main reason of not attending any educational institution are: too old/young (28.7 per cent), no money for fees (13.75 per cent), education is useless or not interesting (10.4 per cent), unable to perform at school (6.5 per cent), disability (5.7 per cent), illness (5.3 per cent). Regarding the time taken by child to get the educational institution he or she attends, 43 per cent of school attending 19,102 children take 15-30 minutes, 41 per cent of them take less than 15 minutes, 12 per cent of them take 31-60 minutes, 1.7 per cent of them take 61-90 minutes, and 0.28 per cent of them take more than 90 minutes. So that, 14 per cent of them take more than 30 minutes to get to the educational institution he or she attends. When it comes to problems encountered at educational institution, 5.5 per cent of school attending children (5-17 yrs.) lack of books, 2.5 per cent receive poor quality of teaching, 2.2 per cent lack of teachers, 4 per cent are going to school which has facilities in bad condition, for 4.6 per cent of them school fees are too high, 4.85 per cent are in classes too large/too many learners, 2.75 per cent are in a school where teacher are often absent, and 1.5 per cent are in school where teachers were involved in strike. Moreover, some children are reported to have experienced also violence corporal punishment or verbal abuse at school between January and June 2011[505]: 3,153 children have had corporal punishment by teacher, 119 children have had physical violence by teacher, 122 children have encountered verbal abuse (being insulted, teased or harassed) by teacher, 238 children have encountered verbal abuse (being insulted, teased or harassed) by other learners, and 226 children had been exposed to physical abuse (being hit or punched) by another learner.

Moreover, in terms of child support grant (CHSG), 58.5 per cent of children receive the grant. For instance, one of two children living in a household containing 2 children receive the grant. Among unique child households, 45 per cent of children receive the grant. Among households with at least 3 children, between 60 and 70 per cent of children receive the grant.

Disability

For this short disability sub-group analysis, we will use 4 different indicators, which are DIS1, DISUN, DISSEV, and DIS2. The first indicator is based on the question "Is.... limited in his or her daily activities, at home, at work or at school, because of a long-term physical, sensory, hearing, intellectual, or psychological condition, lasting six months or more?" in the survey. If

including both children and adults, household level computation may be considered an advantage. However, we suggest separating adult vs. child poverty examination as children's case deserve a more particular attention given that their needs are not necessarily the same as those of adults, or their life activities including school, social life differ in some respects from those of adults. A separate adult and child poverty analysis leads to a substantial volume of work. As we empirically focus on 4 countries in particular, then such volume of work will be greatly excessive. This is why, given the purpose of this thesis, we preferred to focus on adults and not include children in main poverty analysis.

[504] which corresponds to 3.76 per cent of 19.907 children aged 5-17 yrs. Moreover, there are 7,465 children aged 0-4 yrs.

[505] Due to the period mentioned in the question, the target population does not equal those of other school attendance questions; this is why, we give rather the effective numbers than per centages.

the answer 'yes' is given by the interviewee, the he or she is considered disabled regarding that question. The second indicator is computed regarding the United Nations criteria mentioned in the survey questionnaire; thus the person is considered disabled if he or she has reported having some difficulties regarding at least 2 of 6 following items: seeing (even with glasses if he/she wears them), hearing (even with a hearing aid, if he/she wears one), walking a kilometer or climbing a flight of steps, remembering and concentrating, with self-care, such as washing or dressing, in communicating in his or her usual language including sign language (understanding others and being understood by others). The third indicator corresponds to severe disability which means that if a person has reported having a lot of difficulty or unable to do regarding at least 1 of 6 items mentioned in DISUN indicator. And, the last indicator is composed of 4 categories: (1) severe disability (regarding DISSEV); (2) disability regarding UN criteria (DISUN) and they are not severely disable; (3) having some difficulty in at least 1 of 6 items and they are not severely disable; (4) having no difficulty in any of 6 items.

At first step, we will cross-tabulate each of these 4 disability indicators with poverty group variable DEP_GRS. Then, we will check whether disabled people receive disability grant or not. The variable DIS1 takes 1 if the person is disable, 2 otherwise, and 99 for missing values regarding DIS1 variable. The variable DISUN takes 0 if the person is not disable, 1 otherwise. The variable DISSEV takes 0 if the person is not disable, 1 otherwise. The variable DIS2 has 4 categories as mentioned above, and 99 for missing values regarding DIS2 variable[506].

Regarding DIS1, 1 per cent of the population of 18+ are disabled and are part of the poorest group, or 31.5 per cent of disabled 18+ people regarding DIS1 criteria are in GR5, or we can also say that 4.8 per cent of GR5 are disabled. But also, in total 3.4 per cent of 18+ population are disable regarding DIS1 criteria; (ii) regarding DISUN definition, 0.8 per cent of the population of 18+ are disabled and are part of the poorest group, or 31.6 per cent of disabled 18+ people regarding DISUN criteria are in GR5, or 3.65 per cent of GR5 are disabled. The numbers of DIS1 and DISUN are quite close, and 2.6 per cent of 18+ population are disabled regarding the UN criteria; (iii) regarding severe disability (DISSEV), the numbers are very slightly higher. For instance, 5.2 per cent of GR5 are severely disabled, or 1.16 per cent of the entire population of 18+ are disabled and part of GR5, or 30 per cent of severely disabled 18+ people are in GR5. And 3.9 per cent of 18+ population are severely disable regarding DISSEV criteria; (iv) regarding DIS2 indicator, 84 per cent of the entire population of 18+ do not encounter any disability related issues. The third category has the highest shares in 18+ population regarding each category of deprivation. 8.8 per cent of 18+ population are in the third category of DIS2, 3.9 per cent are severely disable, and 2 per cent are disable regarding UN criteria. Notably, people in third category of DIS2 but also in GR5 of deprivation construct 2 per cent of the 18+ population, or in other words 9 per cent of GR5 experience third category of DIS2 (i.e., having some difficulty in at least 1 of 6 items and they are not severely disable). Moreover, 8.3 per cent of GR3 are disable in terms of third category of DIS2. Compared to poverty groups GR3, GR4, and GR5, there are less disable people in general in poverty groups GR1, and GR2.

In addition, among people disable regarding DIS1 criteria, 50.5 per cent of them receive disability grant; among people disable regarding UN criteria, 10 per cent of them receive disability grant; among severely disable people, 22.6 per cent of them receive disability grant;

[506] No missing values regarding DEP_GRS variable.

and 6.5 per cent of people experiencing third category of DIS2 receive disability grant. Interestingly, 2 per cent of people not experiencing any disability regarding DIS2 indicator (4th category) receive disability grant. This last result may due to data inconsistency or another administrative reason. It is very important to increase disability grant for who suffer severely disability. We already saw the various profiles and poverty configurations of deprivations groups, as a function of them poverty reduction policies will certainly increase for who are disability position. However, for more specific tailored programs, specific surveys can be done or detailed profiles and capabilities patterns analysis can be done to find out whether these people have opportunity of reaching adequate educational attainment or of performing adequate jobs for them, or whether disability grant or their personal income are enough for them to lead a quality life given their specific needs, or to find who is disable and who does not receive disability grant the administrative records should be consulted further.

Elderly

We consider elderly a person aged at least 65 or more. 10 per cent of 18+ population corresponds to elderly. When we plot poverty groups in which elderly population are affiliated, we obtain the numbers below. According to th, we can say that 23.36 per cent of elderly population live in conditions of derivation group GR3. The per centages are quite close for GR4 and GR5 as well. Concerning old age grant (OLDG), 80.7 per cent of elderly population receive the grant. For increasing quality of life elderly population, the capabilities patterns discovered for GR3, GR4, and GR5 should be considered. To find out, where they live exactly (profiles of geographical situation are already given in 2nd model analysis), administrative records should be addressed.

Deprivation groups (DEP_GRS)	Freq.	Percent	Cum.
1	859	18.70	18.70
2	686	14.93	33.63
3	1,073	23.36	56.99
4	975	21.22	78.21
5	1,001	21.79	100.00
Total	4,594	100.00	

Employment, unemployment

Here we will focus on economic activity and unemployment, the variable ES1 which is composed of 7 modalities. The first group corresponds to economically active population, the second is unemployed people (seeking a job and available to start to work), the third to sixth groups represent different sorts of not economically active population, and the last one concerns missing values on availability of the person to start to work even he or she sought a job during the reference period indicated in the questionnaire. We already saw economic activity related profiles of people from poverty groups GR3, GR4, and GR5. It is also important to look at whether other less deprived groups GR1 and GR2 experience also unemployment.

We can observe that the share of economically active population in the market is about 39 per cent among 18+ population. Such per centage can be considered quite frightening given that

almost 6 of 10 people do not work. Based on the international standards, for instance, the national statistical office of Switzerland gives the unemployment definition based on the International Labour organization criterion "The Unemployment Statistics (ILO-based) provide unemployment figures based on the International Labour Organisation's definitions. The unemployed are defined as permanent residents who do not have a job, who are looking for work and who can start work within a short period of time"[507], both seeking a job and being available to start to work should be fulfilled to be considered as unemployed. According to that definition, only 12 per cent of 18+ population are actively seeking a job in the market for starting as soon as possible once they find a suitable job for them. When we look at the economically active figures, we can observe that 59 per cent of GR1, 47 per cent of GR2, 37 per cent of GR3, 30 per cent of GR4, and 26 per cent of GR5 are actually working. There is a clear downward tendency of economically active population rate when we go from the less poor group GR1 to the poorest group GR5. Given the conditions in which notably GR4 and GR5 suffer, considerably low per centage of employment surfaces as one of the biggest jeopardies which can trap poor people to the poorest situation not only in terms of material, or financial, or access to some essential elements of life such as sanitation facilities, drinking water, food but also in terms of emotional, intellectual, and physical collapse. Being active in the labour market does not bring only money but also it is a kind of socialization for making new friends, getting new vision, holding hope for future, and getting motivation not only for surviving but also for catching new opportunities in life and living for real. Even for GR1 which is composed of more wealthy population with probability more access to labour market, or diverse jobs, and potential formations and education, employment rate is slightly above than fifty per cent- slightly less than one of two people in GR1 are unemployed. On the other hand, unemployment rates give another context to this picture: 14 per cent of GR5, 13.5 per cent of GR4, 13 per cent of GR3, 12.4 per cent of GR2, and 6.3 per cent of GR1 are seeking a job and they are available or ready to start to work. If we add up employment and unemployment rates for GR5, the rest 60 per cent are not working due to some specific reasons or they are discouraged for seeking a job. Notably, the category of "he or she did not seek a job and he or she would not have liked to work and even if a suitable job had been offered or circumstances had allowed, he or she would not have been able to start work or a business in the last calendar week" emerges with relatively outstanding per centages: 23.5 per cent for GR1, 25 per cent for GR2, 27.6 per cent for GR3, 31.4 per cent for GR4, and 31.8 per cent for GR5. These per centages give a total of 28 per cent among 18+ population. Different reasons are possible for this category, but it is important to know what is behind notably for GR3, GR4, and GR5 because these groups suffer in a significant way poverty compared to GR1 and GR2, and they need an economic activity for improving their living standards otherwise they risk also being dependent on social grants such as RDP, subsidized dwelling, or becoming indigent household. In addition, the category 4 "he or she did not search but would have liked to work during the last calendar week" mirrors with relatively important per centages among GR3, GR4, and GR5: 12 per cent for GR3, 13 per cent for GR4, and 16.8 per cent for GR5. When the figures of that group are summed up with those of 5 category "he or she did not seek a job and would not have liked to work but if a suitable job had been offered or circumstances had allowed, he or she would have been able to start work or a business in the last calendar week", we reach 7.9 per cent for GR1, 11 per cent for GR2, 17.3 per cent for GR3, 20 per cent for GR4, and 22.3 per cent for GR5. The categories

[507] Please refer to the link below for further information (accessed 4 November 2019) : https://www.bfs.admin.ch/bfs/en/home/statistics/work-income/surveys/els-ilo.html

4 and 5 reflect the situation of discouraged workers, or in other words people who want to be active in fact in the labour market, but they are in somewhere discouraged for seeking a job.

In terms of age group breakdown, when we look at economic activity status for 18-34 yrs. age group, we can see that the category 6 play important role, but also the sum of groups 4 and 5. The rate of economically active population corresponds to 23.3 per cent for GR5, 26.3 per cent for GR4, 32 per cent for GR3, 40 per cent for GR2, and 52 per cent for GR1. There is 28.7 per centage point difference between the least poor and the poorest group. And, these figures remain under the total percentages presented in previous paragraph, meaning that youth people face more unemployment in large sense including discouraged workers compared to people aged between 35 and 64. Given that the share of youth population is higher than 35-64 yrs. group, it becomes important to reduce unemployment and increase new job opportunities for younger generation for the future of the country.

In terms of sector of work, there is a clear shift from less deprived groups GR1, GR2, GR3 to the poorest groups GR4, and GR5. Whilst working in informal sector is about 4 per cent among GR1, 12 per cent among GR2, 23 per cent among GR3, the rate increases to 30.5 per cent among GR4, and 41 per cent among GR5. Informal sector is known for hazardous, not socially secure, and not stable jobs. These figures make think seriously about the structure of labour market, especially type of available jobs in the market. In poverty reduction programs, amongst the effort for increasing employment rate, the possible types of potential works should be also considered because new informal sector jobs may bring other social troubles such as augmentation of gangs, or monopoly of some groups, and create an insecure environment in the labour market.

The most prosperous group (GR0)

In the data, among 18+ people, there are 1029 individuals living with no deprivation regarding all items of our overall deprivation index (this group corresponds to 2 per cent of 18+ population data with no missing values regarding the components of deprivation index)-called GR0. We could not do classification trees for capturing how distinguish GR1 from GR0 and GR2 from GR0 due to imbalanced data. However, for further research, some resampling methods can be used for obtaining a balanced data and apply classification trees for a good prediction. Here, we will simply be contented with descriptive statistics. Their profiles in terms of employment status (EMP), geographical situation (GS), individual characteristics (INDC), and highest educational attainment (EDU) will be explored. This will help us to understand where this most prosperous group lives in the country, whether they have a job in a formal sector or not, whether they are from relatively richer white population or not, and whether they have superior educational attainment or not.

In terms of EMP, the most frequent profile is working in a formal sector (79 per cent of GR0), and then the profile of "did not seek a job and he or she would not have liked to work and even if a suitable job had been offered or circumstances had allowed, and he or she would not have been able to start work or a business in the last calendar week" forms 15.5 per cent of GR0. These two profiles together capture 94.6 per cent of GR0.

Regarding GS, 31 per cent of GR0 lives in urban formal areas of Gauteng, 29 per cent lives in urban formal areas of KwaZulu-Natal, 20 per cent lives in urban formal areas of Western Cape, 6.4 per cent lives in urban formal areas of Eastern Cape, and 5.3 per cent lives in urban formal areas of North West. These 5 profiles bring information on 91.64 per cent of GR0. Given that

the prosperous cities such as Johannesburg, Durban, and Cape Town are in the provinces of Gauteng, KwaZulu-Natal, and Western Cape respectively, it might be no surprise the fact that the most well-off 18+ people of the country live in urban formal areas of these provinces.

When it comes to INDC, there is quite heterogeneity. With only 8 profiles, we arrive to capture 46 per cent of GR0. But, if we just look at in terms of population group, we observe that 56 per cent of GR0 is White, 25.5 per cent is African/Black, 12.3 percent is Indian/Asian, and 6.6 per cent is Colored.

In addition, in terms of highest educational attainment, 66 per cent of GR0 have grade 12/high school diploma, 25 per cent have university degree, 6.6 per cent have higher or post higher diploma, 1.65 per cent have NTC4-NTC6, and 0.78 per cent have NTC3. With the first three types of achievements, 97.6 per cent of educational attainment information is gathered for GR0. Given that most frequented highest educational attainment among 18+ population is grade 12/high school diploma (30 per cent of 18+ population), the it is followed by grade 10-11 (23 per cent), lower secondary education (15.6 per cent), illiterate (9.5 per cent), grade 1-6 (9 per cent), and primary education diploma (6 per cent), we can say that educational attainment of GR0 is significantly higher across the country. Among GR5, the most frequented highest educational attainment is grade 10-11 (23.5 per cent) followed by lower secondary education (20.7 per cent), illiterate (19.6 per cent), grade 1-6 (15.7 per cent), and grade 12/high school diploma (9.8 per cent). For GR4 and GR3 also, the situation is quite similar with slight differences. On the other hand, for GR1, 57.5 per cent of them have a grade 12/high school diploma, and 15 per cent has grade 10-11. For GR2, 40.7 per cent have a grade 12/high school diploma, 24 per cent have a grade 10-11, and 13.8 per cent have a lower secondary education. The illiterate level is mostly seen among GR5 and in opposite only 0.6 per cent of GR1 and 0.3 per cent of GR2 are illiterate. Concerning university diploma, only 2.6 per cent have achieved this level. This achievement rate can be considered very few for a country in upper-middle-income group and given that almost no one in GR4 and GR5 have university degree, 0.9 per cent of GR3 have this achievement, and interestingly 3.25 per cent of GR2 have experienced such high-level education whilst only 0.6 per cent of GR1 have had.

It is important to note prosperous situation of GR0 especially in terms of economic activity and formal employment, high level of education notably university degree, and living in urban formal areas compared to the most deprived groups such GR3, GR4, and GR5 in the country. But still, it is crucial to point out that educational achievement of even prosperous groups such as GR1, and GR2 are quite low as mostly grade 12 is achieved rather than university degree.

GR1, GR2[508]

Similar to GR0, for GR1 and GR2 also, we will explore here their profiles in terms of economic activity status, geographical situation, individual characteristics, and educational attainment respectively.

With the following 6 profiles of economic activity status, we reach to capture 99 per cent of GR1, and 98 per cent of GR2: 58 per cent of GR1, and 42 per cent of GR2 are economically active and work in formal sector; 24 per cent of GR1 and 26 per cent of GR2 did not seek a job

[508] Here, we don't use their most relevant patterns obtained through the classification trees because the tree analysis is particularly important for GR3, GR4, and GR5 to understand their position vis à vis the least poor groups GR1 and GR2, which shows also inequality between relatively prosperous and relatively poor groups.

and he or she would not have liked to work and even if a suitable job had been offered or circumstances had allowed, and he or she would not have been able to start work or a business in the last calendar week of the interview; 6.5 per cent of GR1 and 13 per cent of GR2 are unemployed (seeking a job and available or ready to start to work); 5.3 per cent of GR1 and 8 per cent of GR2 did not search but would have liked to work during the last calendar week of the in interview; 2.8 per cent of GR1 and 3.6 per cent of GR2 did not seek a job and would not have liked to work but if a suitable job had been offered or circumstances had allowed, he or she would have been able to start work or a business in the last calendar week of the interview; and 2.4 per cent of GR1 and 5.8 per cent of GR2 are economically active working in informal sector.

In terms of geographical situation, 26. 3 per cent of GR1 and 19 per cent of GR2 live in urban formal areas of Gauteng; 22 per cent of GR1 and 17.8 per cent of GR2 live in urban formal areas of Western Cape; 18 per cent of GR1 and 11 per cent of GR2 live in urban formal areas of KwaZulu-Natal; 7.5 per cent of GR1 and 8.6 per cent of GR2 live in urban formal areas of Free State, 7.2 per cent of GR1 and 7.6 per cent of GR2 live in urban formal areas of Eastern Cape; 4.5 per cent of GR1 and 4.6 per cent of GR2 live in urban formal areas of North West; 4.4 per cent of GR1 and 5.35 per cent of GR2 live in urban formal areas of Mpumalanga; 3 per cent of GR1 and 1.55 per cent of GR2 live in urban formal areas of Limpopo; 2.5 per cent of GR1 live in tribal areas of Limpopo, Mpumalanga, and KwaZulu-Natal; 11.95 per cent of GR2 live in tribal areas of Limpopo, Mpumalanga, North West, and KwaZulu-Natal; 2.9 per cent of GR2 live in urban informal areas of Gauteng, and KwaZulu-Natal.

When it comes to individual characteristics, profiles of both GR1 and GR2 are very heterogenous which makes difficult to obtain synthetic information. Concerning population group, 44 per cent of GR1 and 68 per cent of GR2 are African/Black; 30.6 per cent of GR1 and 7.78 per cent of GR2 are White; 16 per cent of GR1 and 20.2 per cent of GR2 are Colored; and 9.5 per cent of GR1 and 3.8 per cent of GR2 are Indian/Asian. For the highest educational attainment, 57.5 per cent of GR1 and 41 per cent of GR2 have grade 12/high school diploma; 14.9 per cent of GR1 and 24 per cent of GR2 have grade 10-11, 10 per cent of GR1 and 2.9 per cent of GR2 have university degree; 6 per cent of GR1 and 14 per cent of GR2 have lower secondary education; 6 per cent of GR2 have grade 1-6; 4.8 per cent of GR2 have primary education diploma; and 3.25 per cent of GR2 are illiterate.

Additional variables

In GHS questionnaire, concerning health dimension, the following 3 questions were asked: (i) What is the main reason, why,,,,,did not consult any health worker concerning her/his illnesses/injuies during the past month? (HWC); (ii) Has …been informed by a medical practitioner or nurse that he or she suffers from any of the chronic illnesses or conditions including asthma, diabetes, cancer, HIV and AIDS, hypertension/high blood pressure, arthritis and he or she takes the medication for this? (CILM_1,…,CILM_6); (iii) What is time to travel to the health institution? (TIMEH2). These 3 health indicators were not used in construction of overall deprivation index as their target groups are more specific (subset of 18+ population), and so that in classification tree analysis. We will explore distribution of these 3 indicators among the five poverty groups, and by medical coverage instead.

Concerning HWC which concerns only people did suffer from any illness or injuries during the past month, the per centages of not consulted because it is too expensive, or not consulted

because it is too far, or not consulted because of fear of stigmatization emerge very low (under 2 per cent). 77 per cent of them consulted; 14.8 per cent did not consult because the person did self-medicated/treated her/his self, and 5.3 per cent did not consult because it was not necessary/the problem was not serious enough. Among people suffering the 6 types of chronic illnesses or conditions mentioned in CILM_1,…,CILM_6, 10 per cent of people suffering asthma, 4 per cent of people suffering diabetes, 17.65 per cent of people suffering cancer, 10 per cent of people suffering HIV and AIDS, 4.3 per cent of people suffering hypertension/high blood pressure, and 12 per cent of people suffering arthritis does not take medication. 90 per cent of people suffering these illnesses are not able to take medication but notably they are not covered by medical coverage. 32 per cent of this last group are in poverty group GR5, 25 per cent are in GR4, 20 per cent are in GR3, 12.3 per cent in GR2, and 11 per cent are in GR1. The share of these illness people suffering of both no medical coverage and no medication is 1.6 per cent of population 18+, which can be considered very low however, it is important to take in consideration that even chronic sick people in prosperous groups such as GR1 and GR2 suffer capabilities deficit even if it is more sever for GR3, GR4, and notably GR5.

Regarding time to travel to the health institution, across all poverty groups, 16.6 per cent of 18+ people take more than 30 minutes and facility is the nearest of its kind to the dwelling; and 80.3 per cent take less than 29 minutes. 63 per cent of GR5, 74.8 per cent of GR4, 83.6 per cent of GR3, 90.6 per cent of GR2, 94.4 per cent of GR1 take less than 29 minutes to reach health institution. In this regard, we can observe that for the poorest groups, the per centages are relatively lower notably for GR5. It is important to note that this indicator does not inform about the quality of service in the institution. 80 per cent overall statistics of reaching in less than 29 minutes to the health institution is an important element increasing capability of the person regarding health dimension; but on the other hand, medical coverage is low across all groups of deprivation.55 per cent of GR1, 18.65 per cent of GR2, 6.7 per cent of GR3, 2.4 per cent of GR2, and only 0.6 per cent of GR5 have medical coverage. For GR1, the situation is better and in opposite for GR5 it is terrible and even for GR2 is less than 5 per cent, and there is a clear diminishment of medical coverage from GR1 to GR5. Access to adequate health services is one of the core elements of surviving. In this vein, distance from home to health institution, social insurance related to medical coverage of citizens, adequate quality of services in the institution, and availability and accessibility to health institutions for both general and specific treatments are primordial. In terms of former item, statistics show relatively good tendency at national level, but lack of medical coverage should be taken seriously as our classification tree analysis have already shown.

In questionnaire, there are also 2 specific questions related to water supply interruption by municipality: “Has your municipal water supply been interrupted at any time during the last 12 months?” (WATI), and “Thinking about the interruptions in your municipal water supply over the last 12 months, was any specific interruption longer than two days?” (WATI_D). These 2 indicators were not included in overall deprivation index as not everyone receives water supply from their municipality. 40 per cent of 18+ population across 5 poverty groups has suffered of water supply interruption. More specifically, 28 per cent of GR1; 34 per cent of GR2, 42.3 per cent of GR3, 47.4 per cent of GR4, 44.2 per cent of GR5 had encountered this situation. In addition, 30 per cent of GR5, 17 per cent of GR4, 12 per cent of vs. 6,7 per cent of GR2, and 2.8 per cent of GR1 do not receive water from municipality. We can observe that given that notably GR4 and GR5 suffer inadequate access to drinking water, it becomes important water interruption and municipality service indicators. Only 26 per cent of GR5 and 36 per cent of

GR4 have received with no interruption water from municipality. In other words, the capabilities of 2.5 of 10 persons among GR5 and 3.6 per cent of 10 persons among GR4 are reduced in terms of access to water. In addition, water supply to 31 per cent of GR5, 32 per cent of GR4, 24.4 per cent of GR3, 13.6 per cent of GR2, 7 per cent of GR1 by municipality is interrupted longer than 2 days. This former water interruption indicator shows severity of deprivation regarding adequate access to water. Interruption longer than 2 days is a serious problem which should not be happen in normal conditions if we think that water is one of the fundamental and natural needs of each single human being. The fact that even if amongst the less poor groups such as GR1, GR2 the rate of water interruption more than 2 days is on average 10 per cent.

We already used ELC variable in overall deprivation index; however, there is also a more specific variable, namely ELCCUT which focuses on electricity cutoff including no connections to the mains electricity supply, electricity cutoff, free basic electricity and electricity cutoff, and electricity cutoff for another reason. Regarding that variable, 41.15 per cent of GR5, 11.65 per cent of GR4, 6 per cent of GR3, and 7.6 per cent of GR2, and 9.3 per cent of GR1 have encountered electricity cutoff during the past 30 days for nonpayment. Notably, 10 per cent across five poverty groups have experienced no connections to the mains electricity supply which concerns 38.5 per cent of GR5, 6.5 per cent of GR4, and 1.5 per cent of GR3. As we can see GR5 experience significantly a deficit in electricity supply compared to other groups. Regarding the variable ELC which is based on the question “Does this household presently have a connection to the mains electricity supply?”, 66 per cent of 18+ individuals across the five poverty groups have answered yes, 10.35 per cent have said no, and 24 per cent have access to free basic electricity. When we look at its distribution among GR5, we can observe that 39 per cent of them does not have electricity, 19 per cent have free basic electricity, and 43 per cent have electricity. Access to electricity rate increases to 93.5 per cent for GR4 (including free basic electricity). So that, electricity supply programs target on GR5 should be elaborated at national level. However, when we compare this result with what we obtained in GR5 dimensional profiles analysis, we can say that deprivation in electricity supply does not show up as one of the most influential top 5 or top 10 variables. The situation on electricity supply for GR5 is relatively better compared to capabilities failure in medical coverage, sanitation (toilet) facilities, educational attainment, and so on.