

Exploring heterogeneity among Swiss households and
modeling their mobility demand with the Agent-based
computational economics (ACE) approach

PhD thesis submitted to the Faculty of Economics and Business

Institute of Economic Research

University of Neuchâtel

For the degree of PhD in Economics

by

Alperen BEKTAS

Approved by the dissertation committee:

Prof. Mehdi Farsi, University of Neuchâtel, thesis director

Prof. René Schumann, HES-SO Valais, thesis director

Prof. Hannes Weigt, University of Basel

Prof. Sylvain Weber, HEG Genève

Defended on 11. 05. 2022

IMPRIMATUR POUR LA THÈSE

Exploring heterogeneity among Swiss households and modeling
their mobility demand with the agent-based computational
economics (ACE) approach

Alperen BEKTAS

UNIVERSITÉ DE NEUCHÂTEL
FACULTÉ DES SCIENCES ÉCONOMIQUES

La Faculté des sciences économiques,
sur le rapport des membres du jury

Prof. Mehdi Farsi (co-directeur de thèse, Université de Neuchâtel)
Prof. René Schumann (co-directeur de thèse, HES-SO Valais)
Prof. Milad Zarin (président du jury, Université de Neuchâtel)
Prof. Sylvain Weber (HEG Genève)
Prof. Hannes Weigt (Université de Bâle)

autorise l'impression de la présente thèse.

Neuchâtel, le 13 septembre 2022

Le doyen

Valéry Bezençon



Abstract

The thesis follows a research narrative that is constructed through chapters and their connections. The thesis first examines the heterogeneity among Swiss households regarding their mobility behavior and characteristics. It first analyzes the characteristics of habit-driven and cost-sensitive mobility users and scrutinizes the variables that increase or decrease the likelihood of being habit-driven or cost-sensitive. Then, it conducts a user segmentation study. It segments Swiss households based on their mobility behavior and characteristics. It presents the Swiss mobility segments, which are used later for economic modeling and policy recommendation. After that, the thesis builds and validates an agent-based computational economics model that incorporates the observed heterogeneity. The agents in the model can be heterogeneous regarding their socio-economic characteristics, as well as mobility behavior and preferences e.g., agents belonging to the same segment have the same utility function. Finally, the thesis uses the built model for a socio-economic study, in which the goal is to define scenarios for the Swiss mobility system and analyze each scenario with the built model. The scenarios are analyzed at the macro-level; total mobility demand and CO₂ emissions per scenario have been calculated. Besides, agents' reactions to the specific scenarios have been analyzed at the micro-level, since the model has the capability of simulating the micro-behavior of each agent separately. Overall, the thesis provides readers with empirical insights and methodological novelties that potentially improve the existing knowledge and practices in the literature.

Keywords: agent-based computational economics (ACE), heterogeneous agent models, decarbonization, heterogeneity, consumer segmentation, scenario analysis, mobility behavior

Résumé

La thèse suit un récit de recherche qui est construit à travers des chapitres et leurs connexions. La thèse examine d'abord l'hétérogénéité des ménages suisses en ce qui concerne leurs caractéristiques et leur comportement de mobilité. Elle analyse d'abord les caractéristiques des utilisateurs de la mobilité motivés par l'habitude et sensibles aux coûts et examine les variables qui augmentent ou diminuent la probabilité d'être motivé par l'habitude ou sensible aux coûts. Ensuite, elle réalise une étude de segmentation des utilisateurs. Elle segmente les ménages suisses en fonction de leurs caractéristiques (p. ex. socio-économiques et démographiques) et de leur comportement de mobilité. Elle présente les segments de mobilité suisses, qui sont ensuite utilisés pour la modélisation économique et la recommandation de politiques. Après cela, la thèse construit et valide un modèle économique computationnel basé sur les agents qui incorpore l'hétérogénéité observée. Les agents du modèle peuvent être hétérogènes en ce qui concerne leurs caractéristiques socio-économiques, ainsi que leur comportement et leurs préférences en matière de mobilité ; par exemple, les agents appartenant à un même segment ont la même fonction d'utilité. Enfin, la thèse utilise le modèle construit pour une étude socio-économique, dans laquelle le but est de définir des scénarios pour le système de mobilité suisse et d'analyser chaque scénario avec le modèle construit. Les scénarios sont analysés au niveau macroéconomique ; la demande totale de mobilité et les émissions de CO_2 par scénario ont été calculées. En outre, les réactions des agents aux scénarios spécifiques ont été analysées au niveau micro, puisque le modèle a la capacité de simuler le micro-comportement de chaque agent séparément. Dans l'ensemble, la thèse fournit aux lecteurs des aperçus empiriques

et des nouveautés méthodologiques qui améliorent potentiellement les connaissances et les pratiques existantes dans la littérature.

Keywords: économie computationnelle basée sur les agents (ACE), modèles d'agents hétérogènes, décarbonisation, hétérogénéité, segmentation des consommateurs, analyse de scénarios, comportement de mobilité

Contents

Contents	iv
Acknowledgements	ix
List of Figures	x
List of Tables	xii
List of Abbreviations	xiv
1 Introduction	2
1.1 Motivation	2
1.2 An overview of heterogeneity in economics	11
1.2.1 Heterogeneity and the representative agent concept	11
1.2.2 Heterogeneity at the group level	14
1.2.3 Economic models with heterogeneous agents	17
Dynamic Stochastic General Equilibrium (DSGE) Models	18
Agent-Based Computational Economics (ACE) Models	22
1.3 Modeling mobility demand in Switzerland	28
1.4 Contributions and research questions	31
1.4.1 Rationale to use ACE in the context of the thesis	36
1.5 Structure of the thesis	37

2	Habit-driven or cost-sensitive? Exploring the origin of heterogeneous motives of Swiss households for mobility consumption	38
2.1	Introduction	40
2.2	Related Work	43
2.2.1	Habit as determinant of mode-choice	44
2.2.2	Habit and policy response	46
2.2.3	Heterogeneity among mobility users	46
2.3	Data	47
2.3.1	Sample Selection	48
2.4	Methodology	49
2.4.1	Multiple Logistic Regression Analysis	49
2.4.2	Model Specification	50
2.4.3	Descriptive Statistics	53
2.5	Results	57
2.6	Discussion	59
2.7	Limitations and Future Work	60
2.8	Conclusion	61
2.9	Appendices	62
	Appendix A	62
3	Market Segmentation: Obtaining Swiss mobility consumer segments to analyze heterogeneity at the group level	64
3.1	Introduction	65
3.2	Literature Review	71
3.3	The overall concept of a market segmentation procedure	74
	The 1st Step	75

The 2nd Step	80
3.4 Detailed Configuration and Descriptive Statistics	81
3.5 Results	86
3.5.1 Optimal number of clusters and Gower distance weights	86
3.5.2 The Swiss mobility consumer segments	88
3.6 Discussion	92
3.7 Limitations	93
3.8 Conclusion	94
3.9 Appendices	95
4 Building an ACE demand model and validating it with a new meso-	
level empirical validation method	101
4.1 Introduction	104
4.2 Validation Methods	108
4.2.1 Theoretical Background of Validation of ACE models	108
4.2.2 Related Works Regarding Empirical Validation	112
4.2.3 The Overall Concept of the Meso-level Validation Method	116
4.3 Case Study	122
4.3.1 The Simulation Framework	123
General Features	123
Overview of Agent's Design	124
Agent's Decision Making Mechanism	125
4.3.2 Model Building	129
4.3.3 Variable Selection for the Case Study	131
4.4 Empirical Validation Results	132
4.5 Discussion	135

4.5.1 Discussion of the potential application of the meso-level approach to validation upon further models	135
4.6 Conclusion	137
4.6.1 Summary on model building	137
4.6.2 Summary on model validation	139
4.6.3 Future Work	140
4.7 Appendices	141
Appendix A	145
5 Analyzing Future Scenarios & Decarbonization Pathways for the Swiss	
Mobility Sector with an ACE model	147
5.1 Introduction	149
5.2 The selection of active and passive scenario variables	151
5.3 Macro-level Analysis	153
5.3.1 Model Coupling Structure	154
5.3.2 Trajectories for the active variables	155
5.3.3 The space of scenarios	159
5.4 Micro-level Analysis	160
5.5 Results	161
5.5.1 Results from the Macro-Analysis	161
Estimated mobility demands across scenarios	161
Estimated Carbon-dioxide emissions per scenario	163
5.5.2 Results from the Micro-Analysis	167
5.6 Conclusion	171
6 Conclusion	174
6.1 General Conclusion	174

6.2 Final Remarks - Policy Discussion	178
6.3 Future Work	179

Acknowledgements

Throughout the writing of this dissertation I have received a great deal of support and assistance. I would first like to express my sincere gratitude to my supervisors, Professor Dr. Mehdi Farsi and Professor Dr. René Schumann for their guidance throughout this journey. Their expertise was invaluable in formulating the research questions and methodology. Their insightful feedback pushed me to sharpen my thinking and brought my work to a higher level.

I would like to extend my sincere thanks to my colleagues at the HES-SO for their excellent collaboration. I would particularly like to single out my co-authors Valentino Piana and Khoa Nguyen at the HES-SO; I want to thank you for your collaboration in particular.

I would like to thank my current colleagues at the Bern University of Applied Sciences for their support and help. I would like to thank especially my director Prof. Dr. Matthias Stürmer, for his outstanding support and motivation. I would also like to extend my deepest gratitude to my dear colleague Anja Wüst for her encouragement whenever I've needed it.

Very special thanks also for the financial support of the Swiss Federal Office of Energy and from the Swiss Competence Center for Energy Research (SCCER CREST) funded by the Commission for Technology and Innovation.

And finally, thanks to my parents, brother, and numerous friends who endured this long process with me, always offering support and love.

List of Figures

1.1	A spectrum of heterogeneity in economic modeling	16
1.2	Basic structure of DSGE models (Sbordone et al., 2010)	20
2.1	Income level distributions of the habit-driven and cost-sensitive private car users	54
2.2	Degree of urbanization of the habit-driven and cost-sensitive private car users (1- Cities, 2-Smaller cities and suburbs, 3-Rural areas)	55
2.3	Age distributions the habit-driven and cost-sensitive users	55
2.4	Daily distances of the habit-driven and cost-sensitive PC users	56
3.1	The Overall Concept of the Segmentation Procedure (contribution of the colors to the steps; blue: the first, red: the second, black: contributes to both)	76
3.2	Number of clusters and the corresponding ASW values	86
3.3	Segment 1	96
3.4	Segment 2	96
3.5	Segment 3	97
3.6	Segment 4	97
3.7	Segment 5	97
3.8	Segment 6	98
3.9	Segment 7	98
3.10	Segment 8	98

3.11 Segment 9	99
3.12 Segment 10	99
3.13 Segment 11	99
3.14 Segment 12	100
3.15 Segment 13	100
4.1 A general procedure to validate ACE models (Klügl, 2008)	110
4.2 The Overall Concept of the proposed empirical validation method	117
4.3 Overview of Agent's Design (Nguyen and Schumann, 2019)	124
4.4 Triandis TIB model (Triandis, 1979)	126
4.5 Agent's decision-making mechanism with mapping to TIB (Bek- tas et al., 2018)	127
4.6 ASW values with different cluster numbers	132
4.7 Density distribution of all possible validation scores	134
4.8 Density distribution of all possible validation scores with 300 quan- tum size (instead of 150 in Figure 4.7)	145
5.1 Each coordinate on the x-axis is a scenario, and y-axis is corre- sponding Pkm in 2050	162
5.2 Ordered scenarios and their OCCUP trajectories	163
5.3 Ordered scenarios and their LOCREG trajectories	164
5.4 Ordered scenarios and their ONLINE and HOMEOFF trajectories	165
5.5 Ordered scenarios and their PTEVO trajectories	166
5.6 Ordered scenarios and their mobility price MOBPRICE and TAX- FUEL trajectories	166
5.7 Modal Split with (red) and without (blue) LOCREG	169
5.8 Modal Split with (red) and without (blue) PTEVO	171

List of Tables

2.1	Income Levels in Swiss francs according to the MTMC data	53
2.2	Degree of urbanisation of the residence categories according to the MTMC data	54
2.3	Regression Results (Only odds ratios of significant variables are demonstrated)	57
2.4	The Hosmer-Lemeshow goodness of fit test results	58
2.5	Pearson's Chi-squared Test Results	62
2.6	One Way ANOVA Test Results (*F-Values are in the parenthesis)	62
2.7	Pearsons' Correlation Test Results	63
2.8	VIF values for both models	63
3.1	The Optimized Gower Distance weights per variable	87
3.2	Some characteristics of the Swiss mobility consumer Segments	88
4.1	Quantity of Artificial Agents and Real Instances in the clusters and corresponding scores	133
4.2	Quantity of Artificial Agents and Real Instances	142
4.3	The scores and their judgments by statistical distributions	142
4.4	The number of real (R) and artificial (A) individuals in the clusters according to the assumed configurations and situations	144
4.5	The scores generated by the indicator	144
4.6	Percentage values	145

5.1	Passive Scenario Variables and their categories (Schumann, Weigt, and Kober, 2020)	152
5.2	Active Scenario Variables and their acronyms (Schumann, Weigt, and Kober, 2020)	153
5.3	The space of scenarios as a matrix (the numbers represent the trajectories) (Piana, Del Duce, and Schumann, 2019)	159
5.4	The effect of LOCREG at the segment-level	170
5.5	The effect of PTEVO at the segment-level	170

List of Abbreviations

ACE	Agent-based computational economics
ASW	Average Silhouette Width
BedDeM	Behavior Driven Demand Model
BAFU	Bundesamt für Umwelt
CO₂	Carbon-dioxide
DEGURBA	Degree of Urbanization of the Residence
DSGE	Dynamic Stochastic Equilibrium Models
ECB	The European Central Bank
EEA	European Environment Agency
EV	Electric Vehicle
FSO	The Swiss Statistical Office
GHG	Greenhouse Gas
GSL-div	General Subtracted L-divergence
HF	Half-fare
ICE	Internal Combustion Engine
IMF	The International Money Fund
IPCC	Intergovernmental Panel on Climate Change
mDGP	Model Data Generation Process
MSM	Method of simulated moments
MTMC	Mobility and Transport Micro-census
PAM	Partition around medoids

PC	Private Car
Pkm	Personal kilometers traveled
PT	Public Transportation
rwDGP	Real World Data Generation Process
SMD	Simulated minimum distance
SVAR	Structural Vector Autoregressive
TBP	Theory of Planned Behavior
TIB	Theory of Interpersonal Behavior
TRA	Theory of Reasoned Action
VIF	Variance Inflation Factor
Vkm	Vehicle kilometers traveled

Chapter 1

Introduction

1.1 Motivation

Given the urgency of mitigating the economic and environmental impacts of climate change, decarbonization of the mobility sector constitutes an important issue. The sector produces about 16% of the global Greenhouse Gas (GHG) emissions (Ritchie and Roser, 2020). In Switzerland, this share is even higher; the sector accounts for around 40% of the domestic GHG emissions and about 48% of the domestic carbon-dioxide (CO₂) emissions (Boulouchos et al., 2021). The country has committed to achieving net-zero carbon emissions by 2050 (Bundesrat, 2021). Reaching this target requires a substantial transformation of the entire Swiss economy towards a vast reduction of GHG emissions. While other economic sectors such as industrial production and construction could manage to reduce their GHG emissions in the period of 1990-2019, the emissions from the mobility sector grew during that period due to domestic demand growth (BAFU, 2019). Considering the sector's share in the total emissions and increasing emissions trend, more attention should be paid to it.

Developments in the technology side support countries to transform different economic sectors in line with their emission reduction commitments. These

developments improve the efficiency of energy use in production and consumption. So, they contribute to mitigating GHG emissions from economic activities. The mobility sector has also profited from these developments; for instance, the average CO₂ emissions per car in Switzerland decreased by 3.6% on a yearly average basis between 2007 and 2015 (Cox, 2018). The fleet is partially renewed each year due to newly registered cars and becomes more efficient relative to the previous year. The increasing efficiency of the fleet reflects positively on mobility emissions. Economic policies targeting the supply side of the mobility sector by promoting efficiency have been implemented to speed up the transformation in the fleet towards lower emissions. However, technological developments at the supply side alone will not lead to deep decarbonization in the mobility sector. As Switzerland's example shows, the mobility sector's emissions continued to grow in the last two decades due to the demand growth despite a more efficient fleet. Emission reductions thanks to the efficiency improvements in the supply side (e.g., more efficient fleet) fell behind the emission increase due to demand growth. Changing individual mobility behavior, therefore, plays a pivotal role for a deep decarbonization in the mobility sector. As the IPCC (2018) report indicates, changes in individual mobility behavior can lead to substantial aggregate demand shifts away from high-emitting modes as well as significant CO₂ emission reductions. Therefore, it is required to design and implement economic policies targeting the demand side of the sector to stimulate individual behavior change.

Economic agents on the demand side of the private (personal) mobility sector¹ can be observably heterogeneous, which should be kept in mind while formulating policies to change their mobility behavior. Agents can have differences in factors such as preferences, constraints, beliefs, expectations, or overall characteristics that are relevant and known to them while making a particular economic decision. (Browning, Carro, et al., 2007). As a consequence of the decision-making-related differences, agents can respond to policies differently. That is, a group of them can respond to policy and change their behavior while others do not. Or the magnitude of responses (e.g., consumption reduction) can be diverse. Therefore, heterogeneity among agents should be analyzed and considered while formulating decarbonization policies. Besides, like real agents, agents in mobility demand models should be heterogeneous; thus, models can capture different responses to policies, similar to the real world. Assuming everybody in the population is identical, i.e., a single representative agent for the whole demand side might be sufficient to analyze different economic problems for which inclusion of heterogeneity might bring little benefits in return for high computational complexity (Krusell and Smith, 1998). However, specifically, when modeling mobility demand, models with heterogeneous agents can better represent individual-level behavior responses to policy and consequent aggregate demand shifts. Thus, these models can better assess the effectiveness of decarbonization policies to change individual mobility behavior. The main rationale for that is because mobility systems are complex adaptive systems by

¹Private mobility refers to personal or individual use of mobility that is not open to the public at large (Queiroz, Roque, and Moura, 2020). Individuals can freely choose the time and route of the mobility activity and carry it out via various modes of mobility such as private car, bike, walking, motorbike, etc. Throughout the thesis, economic agents on the demands side of the mobility sector refer to individuals who make decisions to carry out their private mobility activities (Cellina et al., 2013)

their nature. It means that they involve many heterogeneous interacting units (Holland, 2006). Models with heterogeneous agents can better depict the characteristics of complex adaptive systems since they can capture the heterogeneity among units and their interactions more appropriately (Hommes, 2021). Detailed information about heterogeneous agent models and their appropriateness to model mobility systems will be provided throughout the chapter.

The thesis makes four main contributions in Chapters 2, 3, 4, and 5 by answering research questions in the broad scope of analyzing and modeling the heterogeneity among economic agents. The first main contribution of the thesis is to the literature on individual consumption behavior. Chapter 2 explores the origin of behavioral heterogeneity in mobility consumption of two specific mobility user groups. Previous research presents empirical evidence that the existence of habit hinders individual behavior change (Brette, Lazaric, and Silva, 2017; Gsottbauer and Bergh, 2011). Economic agents motivated by habit are less sensitive to economic stimuli than cost-sensitive ones (Maréchal, 2010; Idris, Habib, and Shalaby, 2015; Gärling and Axhausen, 2003). The chapter rationalizes differentiating mobility consumption behavior of habit-driven and cost-sensitive Swiss individuals by conducting an empirical analysis to explore heterogeneity in their characteristics. Chapter 2 intends to answer the following research question:

- What characteristics increase (or decrease) the likelihood of being habit-driven and cost-sensitive mobility user?

They are analyzed as two distinct behavioral segments, and the traits that increases (or decreases) the likelihood of being a member of these segments are analysed. The study extends the body of knowledge on mobility consumption

behavior and can provide policy-makers with new insights into the heterogeneity of the two behavioral segments.

Chapter 2 shows that there is behavioral heterogeneity² among Swiss households and analyzes the structural heterogeneity³ to examine the origin of the behavioral heterogeneity. That is, it observes that there are two groups with different motives for mobility behavior and analyzes the characteristics of those groups to rationalize their differentiating consumption motives. However, the chapter does not quantify the behavioral heterogeneity. In other words, it does not assess how many of these such groups we have. Therefore, Chapter 3 conducts an extensive market segmentation study to obtain the amount of behavioral heterogeneity among Swiss households. So, the chapter contributes to the economic literature on market segmentation and the decarbonization of the Swiss mobility sector. The chapter answers the following question.

- What are the segments of Swiss households based on their mobility consumption behavior and characteristics?

Chapter 3 presents a new structured segmentation procedure to answer the question that improves the existing practice and knowledge. The procedure and its methodological novelties are discussed in the chapter in detail. The procedure is applied to the Swiss data to explore heterogeneity among Swiss households at the group (segment) level. Thanks to the segments, we can explore the mobility behavior patterns of Swiss households and also the quantity of behavioral heterogeneity among them. The application yields the Swiss mobility segments that are internally relatively homogeneous and different from each other.

²Behavioral heterogeneity refers to heterogeneity in mobility behavior throughout the thesis.

³Structural heterogeneity refers to heterogeneity in socio-economic and demographic characteristics (e.g., income level, age, place of living, etc.) throughout the thesis

That is, each segment consists of households with similar mobility behavior and characteristics. The chapter is one of the first market segmentation studies for the demand side of the Swiss mobility sector. The segments can be used for two purposes. Firstly, they can be considered in policy formulation to change individual mobility behavior. Different segments can be targeted by tailored segment-level policies that would yield a better aggregate result than assuming everybody is homogeneous and targeting them as a whole with a single "one size fits all" policy (Anable, 2005; Morton, Anable, and Nelson, 2017; Aksen, Bailey, and Castro, 2015). Secondly, the segments can be used in economic models. They help us to set the degree of heterogeneity in economic models. For instance, instead of using a single representative agent for the whole population, segment-level multiple representative agents can be created and used. That reduces the bias in economic analysis since assuming units in the segments homogeneous is more advisable and realistic than assuming everybody in the population homogeneous.

The third main contribution of the thesis is to the economic literature on economic models with heterogeneous agents. After observing behavioral and structural heterogeneity on the demand side of the Swiss mobility sector, the thesis builds and validates an economic model with heterogeneous agents to model mobility demands of Swiss households. The model incorporates both behavioral and structural heterogeneity among agents. Chapter 4 answers the following question:

- How can an economic model with heterogeneous agents that model mobility demand in Switzerland be built and validated?

Since it is suitable for economic systems with many heterogeneous and independent decision-makers, agent-based computational economics (ACE) is used as the modeling approach. The rationale for using the ACE approach will be discussed in the following sections in detail. This approach enables the capture of heterogeneity at the individual level. Chapter 4 builds an ACE model that simulates and produces the micro-behavior of each agent separately. The model incorporates the findings of the Chapters 2 and 3; thus, it involves behavioral and structural heterogeneity among agents. That is, agents in the model can be heterogeneous in overall socio-economic characteristics as well as their behavior and preferences. Agent's utility function is configured according to their segments, which were obtained in the previous chapter. Thus, agents in the model become heterogeneous in terms of their utility functions. Thanks to the heterogeneity that the model involves, agents can respond to the changes in their economic environment including those caused by policies differently. The model produces individual mobility demands as core output, which can be aggregated to obtain macro patterns such as total mobility demand and total CO₂ emissions. The model enables the evaluation of economic policies targeting individual behavior change. Aggregate demand shifts and consequent CO₂ emission reductions due to the policy implementation can be produced. Moreover, the chapter presents a new validation method for ACE models. As mentioned in many influential ACE studies, developing validation methods is an essential line of research in ACE research since ACE models do not have standard validation methods that hamper the credibility of their outputs (Tsfatsion, [2021](#); Fagiolo et al., [2019](#); Windrum, Fagiolo, and Moneta, [2007](#)). The developed validation method is applied to the built model to assess its output in Chapter 4. Detailed information about the model, its novelties, and the validation method

will be provided throughout the thesis. More discussion about the contributions and clear definitions of research questions will be provided at the end of the introduction chapter.

Chapter 5 discusses a socio-economic study, a part of which is carried out through the model that was built and validated in Chapter 4. That is, the chapter presents a case study with the model. The case study consist of macro and micro level analyses. As for the macro-level analysis, the chapter discusses first the future mobility scenarios, which consist of different trajectories of mobility dimensions whose changes are relevant for the evolution of the Swiss mobility sector. The model estimates for each scenario mobility demand and CO₂ emission trajectories until 2050 that is the year by which the Swiss Federal Council has committed to reducing country's net carbon emissions to zero (Bundesrat, 2021). As for the micro-level analysis, the chapter formulates policies targeting specific user segments specifically. The model simulates micro-behavior with and without the policies. The chapter identifies the agents changing their mobility behavior with the policy and analyzes their segments. The goal is to observe whether the policies affect specific segments as envisaged in the policy formulation phase. Besides, the micro-level analysis compares agents' reactivity to the changes in their economic environment. It has been discussed that agents in ACE models can be heterogeneous, so they can respond to economic stimuli differently (or do not respond at all). Two scenarios that differ only in prices are simulated with the model, and the micro-behavior of agents is explored. The goal is to observe whether some agents react to price changes while others don't due to heterogeneity. That is difficult to capture with the models with a representative agent. The chapter answers the following questions.

- How to use an economic model to estimate future decarbonization and mobility demand trajectories for scenarios?
- How to formulate segment-level decarbonization policies and test them with an economic model?

The rest of the introduction chapter follows a general-to-specific orientation and is structured as follows: the following section presents a broad overview of heterogeneity in economics. It discusses the evaluation of economic modeling approaches from the representative agent concept to the heterogeneous agents concept. Afterward, the types of heterogeneity (e.g., heterogeneity at the individual or group level) and the methods to analyze them are discussed. The section also explains the economic modeling approaches with heterogeneous agents, their advantages and limitations, and their application areas in economic research. As of Section 1.3, the chapter gets more specific and discusses the problems in the Swiss mobility sector and the need to model mobility demand with a heterogeneous agent model. Section 1.3 discusses the rationale for using the ACE approach in the context of this thesis. Finally, Section 1.4 presents the research questions and their contributions to the economics literature. The section shows how answering each research question improves the body of knowledge in a different line of research in economics.

1.2 An overview of heterogeneity in economics

1.2.1 Heterogeneity and the representative agent concept

In economics, the term heterogeneity refers to dissimilarities between the units being studied (Kirman, [2006](#)). The dissimilarities might exist in different factors such as in preferences, constraints, beliefs, expectations, and overall characteristics, which are relevant and known to economic agents when making a particular decision (Browning, Carro, et al., [2007](#)). As a consequence of heterogeneity, economic agents might respond to economic stimuli differently, which should be considered for the interpretation of empirical evidence and the construction of economic policies (Pennings, Garcia, and Irwin, [2005](#); Parker et al., [2013](#)). Economists also use the term latent heterogeneity, which refers to heterogeneity in decision-making relevant factors such as ability or emotions that are known to economic agents but not explicitly to researchers (Gerhard, Gladstone, and Hoffmann, [2018](#); Browning, Carro, et al., [2007](#); Prada and Urzúa, [2017](#)).

For more than a century, economists have used the "representative agent" concept fruitfully (Wickens, [2012](#)). The concept was first adopted by Alfred Marshall in 1890 in his influential study, *Principle of economics*, from the idea of the average man (L'Homme Moyen) of Adolphe Quetelet (Mosselmans, [2005](#)). Marshall initially used the term "representative" for firms and individuals, which economists have generalized as the representative agent (Hartley and Hartley, [2002](#)). The concept implies that a representative (or average) agent (e.g., household, firm, etc.) could stand in for any other in the economy. It is a form of the reductionist approach assuming that all concepts applicable to the macro sphere (i.e., the economy or its financial system) are fully reduced to concepts

and knowledge for the lower-level domain of the representative agent (Colander et al., 2008). The idea is to understand macro phenomena through the cumulative actions of a typical decision-maker for each different type considered (e.g., consumer or firm). These typical decision-makers are called representative agents and maximize their own welfare (e.g., utility function for consumers or profits for firms). The concept makes economic analysis more straightforward and less complex because it is simpler to analyze the decisions of a representative unit instead of studying decisions of many units at the same time. An economic model is called a representative agent model if all agents having the same type are identical. In such a model, all agents act in a sense that their cumulative actions (e.g., choices) are equivalent to the actions of one agent maximizing her welfare. For instance, if the representative agent would increase its consumption by 5 percent in response to a change in the economy, it is assumed that the aggregate consumption would also increase by 5 percent in response to the same change (Hartley and Hartley, 2002). So, aggregation, which is summing up the behavior of individuals to derive the behavior of an economy, is relatively simpler in representative agent models.

Models built upon the representative agent concept are widely used in different research areas of economics because they make the analysis more tractable (Wickens, 2012). Nonetheless, economists have become more aware of the importance of heterogeneity in analyzing economic phenomena and they have had more sophisticated methods to handle it. Especially after the increasing availability of micro-data and developments in numerical methods and computational power, many studies emerged that expand the frontiers of economics by introducing heterogeneity among economic agents (Hamill and Gilbert, 2016). A reason for this tendency is that collected data in any sub-discipline of economics

demonstrates that the diversity among agents is not negligible. Agents might have different preferences, characteristics, expectations, or social networks for interaction, which can determine their economic actions. As Kirman (2006) states that heterogeneity among agents is obvious, and it plays a role in all aspects of the economy. So, it seems unreasonable to assume it is unimportant or ignore it. Heckman underlines the role of heterogeneity in his Nobel Lecture that

"The most important discovery was the evidence on the pervasiveness of heterogeneity and diversity in economic life. When a full analysis of heterogeneity in responses was made, a variety of candidate averages emerged to describe the "average" person, and the long-standing edifice of the representative consumer was shown to lack empirical support. This changed the way economists think about econometric models and policy evaluation. (Heckman, 2001) (p. 674)"

After acknowledging the role of heterogeneity, economists attempted to develop models with heterogeneous agents. Actually, in the '70s, economists like Lucas, Sargent, and some others already attempted to develop macro models grounded in optimal individual-level decision-making (Evans and Honkapohja, 2005). Their initial intent was to use heterogeneous agents whose decisions result in aggregate outcomes (Hurtado, 2014). However, their initial models relied on the representative agent concept since economics in the '70s lacked the tools to build dynamic models with heterogeneous agents (Heathcote, Storesletten, and Violante, 2009). Nevertheless, since the second half of the '80s, increasing computational power and improvements in numerical methods enabled the

building of heterogeneous agent models in which agents differ from one another (Hamill and Gilbert, 2016; Heathcote, Storesletten, and Violante, 2009). These models can be broadly categorized into two groups; Dynamic Stochastic General Equilibrium (DSGE) models and Agent-based Computational Economics (ACE) models (Giri, 2017). An essential difference among them is that DSGE models are designed to model the economy as a whole. So, they are specific to address macroeconomic problems. ACE models can also model the economy as a whole. Thus, they are the natural competitors and alternatives of DSGE models in macroeconomics (Turrell, 2016). Nonetheless, ACE models can also be sector or system-specific. They are thus more widely used to model decentralized markets, energy systems, industrial organisation, and mobility systems. A more detailed discussion about ACE and DSGE approaches, their applications in economics, their technical properties, the scope of research, advantages and limitations, will be presented in the following sections. Besides, it will be explained why ACE is more suitable than DSGE for this thesis.

1.2.2 Heterogeneity at the group level

Assuming all agents are identical (i.e., the representative agent concept) might cause theoretically controversial and biased results in applied economics (Kirman, 2006). A way to overcome that bias is to model the behavior of each individual separately. As it will be discussed in the next sections more detailedly, some evolutionary modeling approaches (e.g., ACE models) allow for modeling micro-behavior of each agent separately (Dawid and Pyka, 2018). Thus, one can estimate, in principle, the idiosyncratic response of each agent to economic

stimuli and find the aggregate response accordingly. Although modeling heterogeneity at the individual level is appealing and would be theoretically feasible, it is data-demanding. Besides, depending on the number of agents in the model, the concept might demand some computational power and time. So, for those cases, modeling individual level of heterogeneity might not be so feasible. For such situations, economists might be more interested in the concept of heterogeneity at the group level. The concept hinges on grouping individuals with similar preferences, characteristics, or behavior together and generating quasi-homogeneous⁴ clusters (i.e., segments). So, the whole population (i.e., the sum of the observed units) can be divided into groups in which heterogeneity is minimized since individuals belonging to the same group are similar. Thus, the behavior of each group can be modeled instead of each individual separately, which reduces the quantity of the agents modeled. That is, instead of a single representative agent for the whole population, one representative agent for each group is used (Kleven and Kreiner, 2006; Wangsness, Proost, and Rødseth, 2020). In this way, different responses of groups, which contain similar individuals, to economic stimuli can be considered in economic analysis. Modeling heterogeneity at the group level can be viewed as a middle way between modeling heterogeneity at the individual level and assuming all individuals are identical/homogeneous (see Figure 1.1). So, it is a way not to dismiss heterogeneity as unimportant or to assume it away in the presence of time, computation, or data constraints.

Analyzing heterogeneity at the group level is also relevant for the formulation of economic policy. The effectiveness and efficiency of a policy depend

⁴The term quasi-homogeneous is used to indicate that units in the same cluster have very few variations. So, clusters with very similar units are internally almost homogeneous.

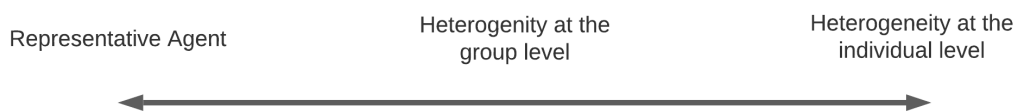


FIGURE 1.1: A spectrum of heterogeneity in economic modeling

on the extent of its impact on the target population. It is unlikely to consider a policy effective if very few individuals in a large target population respond to it. One should consider the heterogeneity in the target population's characteristics and its potential effect on the policy response in policy formulation. Following the "one size fits all" approach to policy formulation may backfire with some sub-groups in the population (Axsen, Bailey, and Castro, 2015). In an ideal but extreme case, each individual in the population would be targeted by a tailored policy based on his preferences, needs, resources, characteristics, etc. However, that might not be feasible with large population sizes. In such cases, heterogeneity among individuals can be analyzed instead at the group level by splitting individuals into internally relatively homogeneous subgroups (i.e., segments) based on their similarities in different dimensions (Ferraro and Miranda, 2013). Then, tailored sub-group-level policies can be formulated and applied that are more likely to generate a higher policy response (Anable, 2005; Alonso-González et al., 2020; Soto, Cantillo, and Arellana, 2021).

A way to analyze heterogeneity at the group level is using cluster analysis, which is a multivariate quantitative method for grouping individuals with similar characteristics into the same clusters. The method hinges on using similarities (or dissimilarities) among observation units (e.g., households, firms, countries, regions, etc.) to group them into clusters, in which internal homogeneity prevails. The similarities are determined according to the observed variables in

the data. For instance, if the task is observing heterogeneity among households at the group level, observed variables constituting the characteristics of households such as income, household size, place of living, consumption patterns, etc., could be the ones used in the clustering method. Individuals allocated to a particular cluster should be as homogeneous (similar) to each other as possible and differ as much as possible from the ones in different clusters. Thus, intra-group variance is minimized that is the main objective of the method. Chapter 3 provides a practical exercise by grouping Swiss households according to their mobility consumption behavior and characteristics.

1.2.3 Economic models with heterogeneous agents

Traditional macroeconomic models (i.e., large-scale models) analyzed macroeconomics with an abstraction from microeconomics until the middle of the '70s (Hamill and Gilbert, 2016). Such models are made up of a set of equations hinging on the correlation between time series from the national accounts. Lucas Jr (1976) published an influential paper called Lucas critique, in which he stated that policy effects could only be estimated by involving the behavior of individuals (e.g., households, firms, etc.); therefore, existing models lacking the micro-level decisions could not evaluate the policy. That paved the way for the development of models that integrate macroeconomics with microeconomics through providing micro-foundations for macroeconomics (Wickens, 2012). These models are called dynamic stochastic general equilibrium (DSGE) models, which were initially built upon the representative agent concept for the sake of simplicity. Then, many DSGE models with heterogeneous agents were developed and widely used for policy evaluation (Krusell and Smith, 1998). Therefore,

DSGE models are the first category of models with heterogeneous agents to be addressed in this section. The second category is the Agent-based Computation Economics (ACE) models, which appeared in '80 and were proposed as an alternative to DSGE models. Similar to DSGE models, ACE models explain macro phenomena with micro decisions by involving many autonomous, heterogeneous, and adaptive agents that do not necessarily need to be rational (Farmer and Foley, 2009). In the following sections, DSGE and ACE models are explained in detail. The rationale to the ACE approach in the context of this thesis will be discussed after presenting the research questions in Section 1.4.

Dynamic Stochastic General Equilibrium (DSGE) Models

Dynamic Stochastic General Equilibrium (DSGE) models are micro-founded macroeconomic tools to be used in policy analysis. These models describe the behavior of the economy as a whole by analyzing the decisions of many micro-entities. A basic DSGE model consists of three interrelated components; demand (i.e., households), supply (i.e., firms), and policy side (i.e., the government) (see Figure 1.2). Each component is represented by agents whose behavior is specified through a set of equations as in Figure 1.2. DSGE models follow the top-down approach, unlike ACE models, as will be discussed in the next section. Agents in these models (e.g., households, firms, policy-side) have complete information about the overall economy, which they use to optimize their private welfare. Besides, agents are assumed to have unlimited computational capacity. As a consequence of having complete information and unlimited computing capacity, agents in these models are assumed to behave rationally (Hamill and

Gilbert, 2016; Giri, 2017⁵. For instance, agents in the demand component are assumed to maximize their expected utility functions over consumption, while those in the supply component are assumed to maximize their profits given some restrictions on technology or other inputs that they use. DSGE models are built upon system-level rules and assumptions, e.g., the system will reach an equilibrium state eventually. In other words, they specify how the overall system will behave rather than how micro-entities (i.e., agents) behave, unlike ACE models. Agents act according to the system level rules, That is, macro-level pre-defined rules (assumptions) determine the micro-behavior and decisions of agents. Aggregating the decision of each agent makes it possible to find prices that clear the market.

The characteristics of DSGE models are defined by their name:

- Dynamic: time is considered in the models, expectations about future effect current decisions
- Stochastic: the models evolve according to unexpected shocks, which can be identified via the economic structure imposed
- General: the models analyze entire economy (system) as a whole
- Equilibrium: agents interact in markets that clear every period

The basic structure of the DSGE models for macroeconomics is illustrated in Figure 1.2. The demand component specifies the real activity (Y) as an equation of the ex-ante interest rate ($i - \pi^e$, i.e., nominal rate minus expected inflation) and of expectations on the future real activity (Y^e). It captures the idea that

⁵Some few studies intended to introduce the bounded rationality concept in DSGE models although it is not traditional to them (Massaro, 2013; Zhao, Lu, and Feng, 2019)

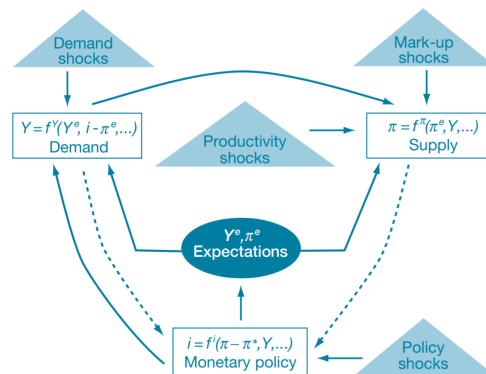


FIGURE 1.2: Basic structure of DSGE models (Sbordone et al., 2010)

individuals and firms would save rather than consume or invest when real interest rates are temporarily high. The emerging real activity (Y) is the primary input for the determination of the inflation rate (π) as well as the expectation of the future inflation (π^e), as the arrow between the demand component and supply component in Figure 1.2 shows. It captures the idea that when the level of activity is high, firms increase wages to convince employees to work longer hours. Increased wages put pressure on prices that leads to higher inflation. The outputs of demand and supply components feed into the policy component as illustrated with the dashed lines in Figure 1.2. The equation of that component sets the interest rates according to the inflation rate and the level of activity. The equation reflects the idea that the short-term interest rate is increased when the economy is overheating (i.e., high Y) and also when inflation rises. The policy component then affects the real activity and through it interest rate by adjusting the nominal interest rates. The equations in three blocks derive from micro-foundations: they describe the behavior of economic agents (i.e., households, firms, policy makers).

The term DSGE was for the first time used by (Kydland and Prescott, 1982)

in their influential paper in which they presented a dynamic model built upon optimizing agents, rational expectations, and market clearing that could generate data that resembled observed variables to some extent. Since then, the usage of DSGE models in the literature made a significant increase, and some assumptions of the early models were relaxed in extended models (Sbordone et al., 2010). They are mainly used for forecasting, counterfactual monetary and fiscal policy analysis, scenario analysis, and assessing the relative strength of shocks contributing the cyclic fluctuations. They are thus the primary tools of central banks (Turrell, 2016), although they got serious critics after the great recession experience (Hamill and Gilbert, 2016). Nonetheless, the usage of DSGE models to forecast the effect of policies out of monetary and fiscal policies is also growing in the literature. For instance, Kiuila (2018) builds a DSGE model to forecast the effect of energy transition policies on the entire economy of Poland. Similarly, Carattini, Heutel, and Melkadze (2021) develop a DSGE model to analyze the impact of climate policies (e.g., carbon tax) on the whole economy in the US. In another study, Zhang and Yang (2019) forecast the effect of subsidy policies in the tourism sector in small-size economies.

The early DSGE models were initially developed upon the representative agent concept, meaning that aggregate consumer behavior is aptly captured by consideration of a representative consumer, and a single representative firm similarly captures aggregate firm behavior. The concept is still used in many models for the sake of simplicity because DSGE models with heterogeneous agents are complicated to solve even with high computational power. Actually, solving DSGE models is still an important research question (Christiano, Eichenbaum, and Trabandt, 2018). Some economists like (Krusell and Smith, 1998) state that the inclusion of heterogeneity in DSGE models brings huge computational costs

and few benefits. On the other hand, many economists using DSGE models support the opposite opinion that DSGE models should include heterogeneity; otherwise, they can't speak to inequality (i.e., there is one person in the economy). Besides, with one person, these models can not conduct a welfare analysis (i.e., who wins and who loses when any change occurs in the economy?). That makes models unable to address questions of unequal growth or differential income risk exposure during a recession. To address these concerns, many DSGE models with heterogeneous agents were developed and presented in recent studies (Christiano, Eichenbaum, and Trabandt, 2018; Levine, 2020). These models can involve heterogeneity among agents in different dimensions such as consumer income and wealth, assets, age, productivity, education, appetite for risk, etc. DSGE models with heterogeneous agents can discuss the distributional effect of policies, have a richer representation of household portfolios, and have more data on which to specify parameters. Nonetheless, it should be noted that the literature on DSGE models with heterogeneous agents is still young (Christiano, Eichenbaum, and Trabandt, 2018; Giri, 2017).

Agent-Based Computational Economics (ACE) Models

As the prevalent definition given by (Tsfatsion and Judd, 2006) states;

agent-based computational economics (ACE) is the computational modeling of economic processes (including whole economies) as open-ended dynamic systems of interacting agents.

ACE models aim for modeling economies as complex, dynamic, and evolving

systems⁶. Instead of the reductionist approach, these models often follow the generativist approach⁷ whose purpose is analyzing macro-level regularities by local interactions of autonomous, heterogeneous agents (Cogliano and Jiang, 2016). Agents in these models can be economic agents (e.g., households, firms, policy-makers, etc.) and also other social and environmental entities such as vehicles, products, banks, social groups. Since agents can be heterogeneous, they might differ in their utility functions, objectives, resources, constraints, etc. Each agent is an encapsulated piece of software that contains microdata for characteristics and behavioral rules for actions. The sum of all agents (including their environment) composes a virtual universe⁸ that enables investigation into how interactions between these agents add up to form the patterns seen in the real world (Hamill and Gilbert, 2016)

ACE models are initialized by constructing an economy with an initial population of agents. The ACE modeler defines the initial state of the economy by specifying the initial attributes of all agents. An agent's attributes might involve its characteristics, utility function, cost function, behavioral rules (including modes of communication and learning), or internally stored information about itself and other agents. The system then evolves over time endogenously without any external intervention⁹. The evolution of the economy can

⁶A system is called a complex adaptive system if it consists of heterogeneous interactive units (Holland, 2006). A system is called an emerging system if units are adaptive and can have the learning ability (Gallegati, Palestrini, and Russo, 2017)

⁷The reductionist approach is widely used in economics that assumes all concepts applicable to the macro level are fully reduced to concepts and knowledge for the lower-level domain of the representative agent. The generativist approach explains the emerging macro-level regularities through micro-level interactions of heterogeneous agents (Colander et al., 2008; Epstein, 1999)

⁸The terms "virtual universe," "artificial world," or "artificial laboratories" are often used to describe ACE models (Duffy, 2006; Tesfatsion and Judd, 2006; Farmer and Foley, 2009; Hamill and Gilbert, 2016)

⁹ACE models might have external variables describing the characteristics of agents' environment. The trajectories of such variables might be exogenously given to the model.

be observed in discrete time steps $T = 1, 2, 3, \dots$. Time steps can be hours, days, quarters, or years. Agents can interact in each time step, learn from these interactions, adapt their expectations, and make decisions according to their objectives. In each time step, agents can collect information about the current and past state of a subset of other agents, those they interact with. They can use that information as well as the information, which might be limited, they can obtain about the state of the whole economy for their actions. The actions of agents can change the environment, which might include physical entities (e.g., infrastructure) and institutions (e.g., central banks, regulatory systems, etc.). In the next time step, agents observe the altered environment and take actions again. All events that subsequently occur in the model are solely driven by agents' interactions and decisions.

Although ACE modeling is relatively new in economics, its applications in different research areas of economics have surged recently. One of these research areas is macroeconomics, where ACE models are natural competitors of DSGE models. ACE models can analyze the economy as a whole, although that is not necessarily their only scope unlike DSGE models. The use of ACE models in macroeconomics gained momentum both in academia and among central banks, especially after the great recession (2007-2009), which posed a serious challenge for the top-down economic models, such as DSGE models (Trichet, 2010). Many major central banks, such as the Bank of England, constructed their own in-house ACE models to project the effect of monetary and fiscal policies (Turrell, 2016). The International Money Fund (IMF) also developed its own macro-financial ACE model called ABBA for banking system (Chan-Lau, 2017). As IMF economists stated, reliance on agent-based formulation, rather than on traditional top-down models, make it possible to account for heterogeneity and

adaptive behavior in the banking sector. ACE models could attract also the attention of the European Central Bank (ECB) just one year after the recession (2010) as the president of ECB Jean-Claude Trichet spelt out "agent based modelling dispenses with the optimisation assumption and allows for more complex interactions between agents. Such approaches are worthy of our attention." (Trichet, 2010). In parallel with the developments on the side of central banks, the application of ACE methodology in macroeconomic literature also increased significantly. For instance, Gatti et al. (2005) analyze the role of monetary policy including agents' learning; Lamperti et al. (2020) investigate the macroeconomic effects of climate policies with an ACE model; Delli Gatti and Reissl (2020) develop an ACE model to analyze the macroeconomic complications of Covid-19; Alexandre and Lima (2020) analyze the interaction between monetary policy and prudential regulation in an ACE model; Dosi et al. (2015) explore the effect of both fiscal and monetary policies in evolving complex economies; Giri (2017) compare ACE models and DSGE models based on their way of evaluating monetary policies; Dawid et al. (2012) develop an ACE macro model with spatial structure for macroeconomic policy analysis. More recent applications and novelties will be discussed in Chapter 4 in detail.

ACE models are not specific to macroeconomics, although it is an important research and application area for them. The literature using ACE modeling in different research areas of economics is growing as well. Since ACE models can contain many autonomous interacting agents, they are particularly suitable to analyze decentralized systems/markets. For instance, electric power markets are largely decentralized and rank among the most complex of all markets operated at present (Weidlich and Veit, 2008). Many economists suggest ACE models as appropriate tools to model electricity markets as complex systems with

many interacting decision-makers (Ringler, Keles, and Fichtner, 2016; Lopes and Coelho, 2018; Veit, Weidlich, and Krafft, 2009; Tesfatsion, 2018). Likewise, transportation systems are complex decentralized systems containing many decision-makers with very heterogeneous and different roles, needs, and aims. Moreover, individual decisions in these systems are partly determined by behavioral factors, for which there is considerable literature in behavioral economics. ACE models can incorporate tenets from classical economics and insights from behavioral economics to formulate individual behavior as well as their interaction structure. So, they are appropriate and frequently used to address questions from the transport domain such as transport demand modeling (Bernhardt, 2007). They can serve as a virtual world with autonomous, heterogeneous, and interacting agents in which economic policies in the scope of decarbonization can be analyzed. Considering the importance and urgency of decarbonizing transport systems, ACE models can serve an important function by enabling a realistic evaluation of policies. Other research areas of economics in which the application of ACE models is increasing would be financial economics (Pruna, Polukarov, and Jennings, 2020; LeBaron, 2012; Franke, 2009), industrial organisation (Chang, 2015; Dawid, 2006), technological change and development economics (Antonelli, 2011), agricultural economics (Berger, 2001).

ACE models have certain limitations and get some critiques. These limitations and critiques might explain why they are still in the frontiers of economics despite their appeals. The first and the most crucial critique is their empirical validation. They do not have standard validation techniques as DSGE models, which undermine the credibility of their results (Fagiolo et al., 2019). ACE modelers borrow the validation and estimation techniques from DSGE models; however, due to the structural differences, these validation techniques are not

sufficient to assess ACE models' output. Although recent studies (e.g., (Lamperti, 2018b; Guerini and Moneta, 2017; Barde, 2020)) provided specific validation methods for ACE models, there is still a need for more sophisticated methods. As many studies stated (Tsefatsion, 2021; Fagiolo et al., 2019), validation is an important research area for ACE economists. Chapter 4 also provides a new empirical validation method for ACE models. The second critique is the "black box" critique. ACE models are, in the end, software programs to model complex adaptive systems. They are thus perceived as black boxes where structural dynamics driving results are blurred. That is mainly because of the differences between the reductionist approach traditionally used in economics and the generativist approach that ACE models follow. Many recent studies attempted to address the black box critique through counterfactual analyses (Petrovic et al., 2017; Gualdi et al., 2015). They have shown that even the causal mechanisms of a very complicated ACE can be detected through counterfactual analyses (Napoleto, 2018). Another critique towards ACE models is that their behavioral rules are sometimes based on ad-hoc assumptions of human behavior without enough theoretical foundation, which is referred to as a "wilderness". However, many recent ACE models specify behavioral rules by using empirical or experimental evidence about real agents' behavior. In addition, ACE models can incorporate insights from behavioral economics; thus, they can specify behavioral rules based on behavioral economic and cognitive theories that can be calibrated with empirical methods. The final critique towards ACE models is that some ACE models highly depend on microdata to initialize agents and the environment. ACE modelers, therefore, developed synthetic micro-data generation tools, which generate realistic microdata over macro aggregates that statistical offices often release. So, even micro-data is not available, such tools can be

used to create realistic microdata over aggregates, and the model can be initialized in that way.

1.3 Modeling mobility demand in Switzerland

The previous sections drew a general framework of the overall thesis by discussing heterogeneity in economics, heterogeneity at the individual and group level, and economic modeling approaches with heterogeneous agents. The discussion so far was mainly on the overall economy rather than a specific sector. This section dives into a more specific look at the Swiss mobility sector since the overall thesis specifically explores the heterogeneity among the decision-makers of the sector and presents a validated ACE model to model the heterogeneous mobility demand of these decision-makers separately. By aggregating individual demands, the model can produce macro-patterns such as total mobility demand and total CO_2 emissions. The section briefly discusses the motivation behind focusing on the mobility sector and its relevance for the overall Swiss economy, mainly within the scope of decarbonization and climate change.

Mobility plays an essential role in the well-being of society as it enables to efficiently deploy economic activities, exchange goods, and move individuals according to their needs. The mobility sector in Switzerland provides jobs for approximately three hundred thousand people (Boulouchos et al., 2021). The private companies in the sector make around 95 billion Swiss francs of revenue. Besides, the sector is interrelated with other major sectors since it provides services and infrastructure for them (e.g., moving inputs of other sectors). So, it is an important sector for the overall Swiss economy. On the other hand, as of 2018, it accounts for about 40% of Switzerland's domestic Green House Gas

(GHG) emissions, 48% of CO₂ emissions, and about 36% of the final energy consumption (BAFU, 2019; Boulouchos et al., 2021). Given the importance of climate change and mitigating its economic and environmental damages, emissions from any sector of the economy, including mobility, must be substantially reduced. While many sectors of the Swiss economy could manage to reduce their CO₂ equivalent emissions from 1990 to 2019 (e.g., industrial production approx. -13% and construction approx. -34%), the emissions from the mobility sector increased in the same period as a result of growth in mobility demand despite the 20% reduction target (BAFU, 2019). Considering the ambitious commitments of the country in the scope of the Paris Agreement, which is getting net zero emissions by 2050 (Bundesrat, 2021), the decarbonization of the mobility sector needs to be paid more attention in the years to come. Taking into account its share in the total emissions, the mobility sector requires a substantial transformation towards a massive reduction of CO₂ emissions to fulfill the commitments of the country.

The developments in the technology side help countries to transform different sectors of the economy towards lower emissions. The mobility sector also profits from the developments; for instance, in Switzerland, the average CO₂ emission per car decreased by around 1.5% until 2006 and about 3.6% per year between 2007 and 2015 (Cox, 2018). Thanks to more efficient vehicles entering the market, the fleet renews itself and becomes more efficient relative to the previous year. Considering the increasing adoption rates of electric vehicles, the yearly pro car CO₂ emission reduction percentage will likely be higher in the years to come. Promoting the technology and incentivizing households to invest in more efficient vehicles to speed up the transformation in the fleet is a dimension of the strategy for decarbonization (Edenhofer et al., 2014; EEA,

2015). For instance, the Swiss Federal Office of Energy announced in 2017 in its detailed CO₂ strategy report that they plan to reduce the CO₂ emissions per car to 95 grams of CO₂ per km in 2020 by putting into force new regulations (BFE, 2021). As of 2021, the average level of emissions from cars registered in Switzerland for the first time may not exceed 118 grams of CO₂ per kilometer, which will more likely improve the average 95 grams of CO₂ per km targets for 2020 (BFE, 2021). However, as many governmental reports and academic studies indicate, technological developments will not alone lead to a deep decarbonization in the mobility sector (EEA, 2015). As Switzerland's example shows, the emissions from the mobility sector continued to increase due to the growth in demand despite a more efficient fleet in the last two decades. That is, emission reductions thanks to the efficiency improvements in the supply side fell behind the emission increases due to the demand growth.

Another dimension of the decarbonization strategy is incentivizing individuals to change their unsustainable mobility behavior (Edenhofer et al., 2014; EEA, 2015; Boulouchos et al., 2021). That is, shifting individuals towards more efficient and environmentally compatible modes and reducing per-capita demand of unsustainable mobility modes (e.g, fossil fuel vehicles). As the IPCC report in 2018 indicates, changes in individual mobility behavior can lead to substantial aggregate demand shifts towards more sustainable modes and significant CO₂ emission reductions consequently (IPCC, 2018). Therefore, several pricing and non-pricing climate policies to change individual mobility behavior such as carbon taxes, campaigns to increase awareness, nudges, etc. have been implemented in many countries, including Switzerland. The policies to change individual mobility behavior consumption primarily target the demand side, although they might also influence the supply side in long term (e.g., change

in fleet). The potential effects of the policies on aggregate mobility demand and consequent emission changes should be assessed ex-ante, which allows for identifying and prioritizing the potentially most effective policies. We need, therefore, models to evaluate the effectiveness of climate policies and explore potential scenarios with different CO_2 emission pathways. The models enable us to project potential demand shifts among mobility modes, growth in aggregate demand, and changes in CO_2 emissions in a time horizon, as a consequence of policy implementations. Thus, they can inform us about how far we are from the long-term commitments. Chapter 5 presents a socio-economic analysis in which a number of future scenarios for the Swiss mobility sector are evaluated by an ACE model. The results can signal the policy makers regarding how far the estimated CO_2 emissions or total mobility demand from the targeted ones (i.e., commitments). Thus, the chapter, as well as the overall thesis, contribute to the efforts in decarbonizing the Swiss mobility sector.

1.4 Contributions and research questions

The overall thesis consists of six chapters. The first and last chapters are introduction and conclusion chapters; the former describes the overall idea of the thesis, and the latter sums up the works done throughout the thesis. Chapters 2, 3, 4 and 5 follow a specific line of research in economics and answers particular research questions. Besides, these chapters answer technical questions to contribute to the methodological aspect. The findings and methodological novelties of the chapters fill the knowledge gap in economics and provide materials (e.g., the model) for other researchers to be used in their research.

Chapter 2 empirically analyzes the heterogeneity in the characteristics of habit-driven and cost aware individuals to rationalize their different motives for mobility consumption behavior. Previous studies in behavioral economics and evolutionary economics has already explored the role of habit in economic decisions (Rozen, [2010](#); Abel, [1990](#); Maréchal, [2010](#); Barnes, Gartland, and Stack, [2004](#); Royer, Stehr, and Sydnor, [2015](#); Roberto and Kawachi, [2015](#)). However, we have little knowledge about the characteristics (e.g., socio-economic and demographic variables) that increases or decreases the likelihood of being a habit-driven or cost-sensitive mobility user. Improving the body of knowledge can help us to understand why a group of individuals respond to economic stimuli while another group does not. The chapter answers the following research question. The evaluation of the answer will be presented in the conclusion chapter.

- What characteristics increase (or decrease) the likelihood of being habit-driven and cost-sensitive mobility user?

Chapter 2 demonstrates two groups with different mobility consumption motives, which is called behavioral heterogeneity throughout the thesis. It then analyzed the structural heterogeneity to explain (understand) the behavioral heterogeneity. However, it does not quantify the behavioral heterogeneity. That is, it does not provide information about how many such groups we have. Chapter 3 analyzes that issue: it examines the quantity of behavioral heterogeneity by conducting a consumer segmentation study. The overall goal of the chapter is to obtain the market segments of the Swiss households whose mobility consumption behavior and characteristics are similar. The chapter uses a large dataset Mobility and Transport Microcensus (ARE/BfS, [2017](#)) with a rich set of information about the characteristics (e.g., demographics), stated preferences,

and mobility consumption behavior (e.g., mode choices for daily routine trips). Based on the information in the dataset, the Swiss households are split into different segments in which internal homogeneity prevails. The chapter answers the following research and technical questions:

- What are the segments of Swiss households based on their mobility consumption behavior and characteristics?

Technical questions:

- How to measure and improve the segment quality?

Chapter 4 exploits the insights from the first two chapters and uses them as a foundation. It builds and validates an ACE model to model mobility demand in Switzerland. The rationale for using the ACE approach will be discussed in the following sub-section. The model incorporates heterogeneity at the individual level and enables modeling the micro-behavior of each individual separately. The model captures both behavioral and structural heterogeneity among agents. It produces individual mobility demands that can be aggregated to find macro-patterns, such as CO₂ emissions from the mobility sector and total mobility demand. The chapter brings novelties (improvements) to the ACE literature. First, the micro-foundations in the model are formulated according to a standard behavioral decision-making theory. The theory describes the drivers of real agents' mobility behavior. The model uses the theory to construct its agents' time-inseparable utility functions, which can be heterogeneous. Agents' utility functions are configured according to their segment information, which we obtained in the previous chapter. Thus, agents belonging to different behavioral

segments have different utility functions, enabling the model to capture behavioral heterogeneity among agents. Agents evaluate existing options according to their utility function, which is based on the theory, and rank the options accordingly. They can choose the first option with the highest-ranking, which brings the highest utility. The model avoids the "wilderness" critique since it uses a standard theory for agents' decision-making instead of making ad-hoc assumptions of human behavior without enough theoretical foundation. That also improves the body of knowledge of integrating insights from behavioral economics to tenets in traditional economics. Second, the chapter presents a novel empirical validation method for ACE models. As discussed in many ACE studies (Fagiolo et al., 2019; Barde, 2020; Lamperti, 2018b; Guerini and Moneta, 2017), ACE models do not have standard empirical validation methods. This is one of the main critiques towards them. So, developing new empirical validation methods is an important research line in ACE literature (Tsfatsion, 2021). This chapter is one of the earliest to use machine learning for the empirical validation of ACE models, which was suggested in influential papers in ACE literature (Fagiolo et al., 2019). The chapter validates the built mobility demand model empirically. Finally, it presents a validated model to test the potential impact of climate policies and project potential CO₂ pathways in Switzerland. The chapter answers the following research question.

- How can an economic model with heterogeneous agents that model mobility demand in Switzerland be built and validated?

Technical question:

- How to use machine learning to assess the output of ACE models for empirical validation?

Chapter 5 demonstrates a use case of the built model for a socio-economic study. The use case consist of the macro and micro-level analyses. As for the macro-level analysis, the chapter first analyzes different dimensions whose changes determine the evolution of mobility in Switzerland. The dimensions are broadly in the scope of socio-economic dynamics (e.g., population development, economic growth, energy prices, etc.), new lifestyle patterns (e.g., aging population, home office and online shopping behavior), Swiss mobility characteristics, policy implementations (e.g., taxation on fuel), and technological developments (e.g., penetration of autonomous taxis). With the combination of different assumptions for the trajectories of the dimensions, future mobility scenarios are created. The model simulates each scenario and estimates mobility demand and CO₂ emission trajectories until 2050 that is the year by which the Swiss Federal Council has committed to be climate neutral (Bundesrat, [2021](#)). As for the micro-level analysis, the chapter formulates segment level policies and test them with the model. The behavior of agents with and without policy is compared to understand whether the policies affect the targeted segments as ex-ante envisaged. The micro-level analysis also shows different reactiveness of agents to economic stimuli thanks to the heterogeneity involved in the model. The chapter answers the following questions.

- How to use an economic model to estimate future decarbonization and mobility demand trajectories?
- How to formulate segment-level decarbonization polices and test them with an economic model?

To sum up, the overall thesis first aims to explore the heterogeneity among Swiss households in many dimensions, such as mobility consumption behavior, and socioeconomic characteristics. It then builds and validates an economic model incorporating the observed heterogeneity. Finally, it uses the validated model to evaluate different future mobility scenarios for the Swiss mobility sector, which involve policy implementations and socio-economic dynamics. Overall, the thesis closes knowledge gaps in different lines of research of economics by answering the research questions and contributes to the decarbonization of the mobility sector in Switzerland.

1.4.1 Rationale to use ACE in the context of the thesis

First of all, this thesis aims to analyze heterogeneity on the demand side of a particular sector, the mobility sector, as discussed in the previous section and builds and validates a heterogeneous agent model specific to modeling demand in that sector. Thus, the thesis presents the empirical findings to be used for decarbonization policy formulation and the validated economic model to be used for policy assessment (evaluation) efforts. So, modeling the economy as a whole is not in the scope of the thesis. From that aspect, the ACE approach is more suitable than DSGE since it enables the modeling of specific sectors or systems without handling the economy as a whole. That is the main rationale for using ACE for this thesis. Besides, this thesis intent to build a model that incorporates the insights from behavioral economics in addition to tenets of traditional economics, as suggested in many studies, such as (Stiglitz, 2018). That is however difficult to do with DSGE models.

1.5 Structure of the thesis

After the Introduction Chapter, Chapters 2, 3, 4, and 5 are presented consecutively. The thesis ends with the concluding chapter that evaluates how these research questions are answered throughout the thesis. It discusses the difficulties and limitations of each chapter separately. It also draws a frame for future work in which the thesis findings can be used, and the methods used in the thesis can be further improved.

Chapter 2

Habit-driven or cost-sensitive? Exploring the origin of heterogeneous motives of Swiss households for mobility consumption¹

Abstract

Individuals become less sensitive to economic stimuli if they use mobility services or modes out of habit, implying that habit-driven mobility users differ from cost-sensitive ones regarding their reaction to traditional price-based decarbonization policies. However, there is little knowledge about the characteristics of these two behavioral user segments with different motives for mobility consumption. Analyzing their characteristics and identifying the variables (socio-economic and demographic) that increase (or decrease) the likelihood of being habit-driven or cost-sensitive can help us to understand the origin of the heterogeneity in their mobility behavior, which will be called

¹A paper based on this chapter has been submitted to the Case Studies on Transport Policy. A. Bektas is the corresponding author of the paper with the co-authors V. Piana and R. Schumann. The paper is currently under review.

behavioral heterogeneity throughout the chapter. That is the main goal and contributions of this chapter, which expands the body of knowledge on mobility consumption patterns and behavior. For the analysis, we utilized the large-scale Swiss Mobility and Transport Micro-census data and singled out individuals who stated their mobility consumption reasons (motives) for their regular trips as habit and as cost. Thus, we created two groups to statistically analyze their characteristics based on the following variables: income level, degree of urbanization of the residence, household size, age, and daily travel distance. As for the methodology, multiple logistic regression was carried out to see which variables significantly raise the likelihood of being habit-driven and cost-sensitive. The results show that certain levels of the categorical variables income level and degree of urbanization of the residence increase the likelihood of being a habit-driven private car user. Besides, the likelihood of being a habit-driven private car user decreases with the increasing daily traveled kilometers. Finally, with the increasing age the likelihood of choosing public transportation out of habit increases. The overall findings demonstrate that there are two behavioral mobility user segments and the distribution of certain explanatory variables is heterogeneous among them, so these variables increase (or decrease) the likelihood of being a member of one of these segments.

Keywords: consumer behavior, heterogeneity, mobility demand, consumer segments, decarbonisation

JEL Classification: D91, D12, R40, C25

2.1 Introduction

Understanding individual consumption behavior patterns is essential for many economic questions, particularly those with policy implications. It is important for decision-makers to understand how consumers would react towards the changes in their economic environment called shocks, including those produced by economic policies. Conventionally, economists handle consumers' reaction to the changes caused by economic policy from the rational behavior point of view under the assumption that individuals are well-informed, self-interested, and make choices based on what delivers them the most significant material benefit. Nevertheless, some sub-fields of economics such as behavioral economics, neuroeconomics, or evolutionary economics find rational behavior assumption is unrealistic and controversial with the empirical and clinical evidence. They intend to learn from different disciplines such as cognitive science, psychology, and neuroscience to understand and explain better how real economic agents (consumers) behave. That try to better understand consumer behavior has yielded a large number of literature that extend the frontiers of economics. A line of research in that literature is related to analyzing the effect of habit on individual consumption. This chapter also follows that line of economic research.

$$\sum_{t=0}^{\infty} \beta_t U(C_t) \tag{2.1}$$

$$\sum_{t=0}^{\infty} \beta_t U(C_t - \theta C_{t-1}) \tag{2.2}$$

Many economists intend to analyze the role of habit in consumption behavior both empirically and theoretically. For the latter, many macro (Jermann,

[1998; Boldrin and Canova, 2001] and micro (Meghir and Weber, 1996; Abel, 1990; Campbell and Cochrane, 1999) economic models have been developed that allow for habit formation, which is a process related to the frequency of the consumption of a good or service in the past. With habit formation, current utility depends not only on current consumption but also on a habit stock formed by lagged consumption (Dynan, 2000). To be more concrete, let's the function in Equation 2.1 denotes a simple utility function without habit formation, in which C_t represents consumption at time t , U represents the utility function and $\beta \in [0, 1]$ represent the subjective discount factor. Then the utility function enabling habit formation becomes as in Equation 2.2 where $\theta \in [0 - 1]$ represents the strength of habit and introduces non-separability of preferences over time (Abel, 1990). The models with the habit formation relax the consumption independence, which refers to that tastes in one period are unaffected by consumption in another (Crawford, 2010). The fourth chapter of the thesis presents a demand model with a time-inseparable utility function enabling habit formation. This chapter rather focuses on empirically analyzing the role of habit in consumption behavior. It especially examines the source of habit-driven consumption in the mobility sector.

Analyzing and changing individual unsustainable consumption behavior is crucial in the sense of the urgency of reducing GHG emissions to mitigate the economic and environmental impacts of climate change. The mobility sector is one of the most significant contributors to the global GHG emissions by about 16%. The sector accounts for around 23% of the global CO_2 emissions (Ritchie and Roser, 2020). In Switzerland, where this chapter focuses on, the contribution of the mobility sector is even higher, as mentioned in the previous chapter. Taking into consideration that Switzerland has committed to achieving net

carbon zero emissions by 2050 (Bundesrat, 2021), changing individual mobility consumption needs to be paid more attention. Previous empirical studies revealed that the existence of habit hinders changing unsustainable mobility behavior (Gärling and Axhausen, 2003). Especially private car usage, which needs to be reduced massively, is mainly habit-driven (Bamberg, Ajzen, and Schmidt, 2003). Consumers with habit-driven behavior tend to stick to their routines, resulting in low reactivity to policies with behavior change targets. These consumers react less to changes in their economic environment than the cost-sensitive ones. That is, as Thaler says, they react to price stimuli not in the ways assumed by economists (Thaler and Ganser, 2015), unlikely to cost-sensitive ones. That might explain the lower reactivity of the mobility sector towards traditional price-based decarbonization policies than its ex-ante estimations (Maréchal, 2010; Jenkins, 2014).

Although there are signals from previous studies that habit-driven and cost-sensitive mobility users have different responses to the changes in their economic environment, few studies analyze the characteristics of these two groups that can be the source of this different behavior and responses (Schwanen, Banister, and Anable, 2012; Idris, Habib, and Shalaby, 2015). We have little information about the variables that increase or decrease the likelihood of being habit-driven mobility user². The present chapter aims for expanding the body of knowledge by analyzing socio-economic variables and their relation with being a habit-driven or cost-sensitive mobility consumer in Switzerland. In other words, the chapter examines the structural heterogeneity in the characteristics of habit-driven and cost-sensitive mobility users to understand the heterogeneity

²The terms mobility consumer and mobility user are used interchangeably throughout the chapter.

in their behavior (i.e., behavioral heterogeneity).

The rest of the chapter is structured as follows. In Section 2.2, we conduct a literature review. Then, we explain the data set that we used for the analysis. In Section 2.4, the used methodology is introduced. Findings of the study follow it. We interpret the findings in the Discussion section. The current limitations and future work plans are explained in Section 2.7. The chapter ends with the Conclusion section.

2.2 Related Work

We split relevant studies into three groups that are discussed in the following subsections. The studies in the first group provide empirical evidence regarding the existence and role of habit in mobility consumption behavior. We discuss these studies because they provide a foundation to demonstrate that habit is a relevant determinant of mobility behavior; thus, individuals can declare (state) it as the reason (motive) for their mobility consumption (as in our data that we explain in the next section). The studies in the second group present empirical evidence showing that habit-driven consumers are less reactive to the changes in their economic environment. They also imply that the existence of habit can lead to bounded rational consumption behavior. The studies in the second group motivated the present chapter that habit-driven and cost-sensitive consumers can have different responses to price-based decarbonization policies; so, we aim to investigate their characteristics to explore the source of heterogeneity. The studies in the third group analyze heterogeneity in different user groups' behavior and characteristics. We discuss such studies to demonstrate that exploring the heterogeneity among users/consumers is a relevant topic for mobility and other

sub-fields of study in economics, such as household consumption, energy consumption, and individual preferences.

2.2.1 Habit as determinant of mode-choice

Starting with the influential paper of Verplanken et al. (1994), there are increasing pieces of empirical evidence regarding the existence of habits in mobility consumption behavior (Klößner and Matthies, 2004; Gardner, 2009; Bamberg, Ajzen, and Schmidt, 2003). Friedrichsmeier, Matthies, and Klößner (2013) examine how to measure mobility habits. The paper suggests that the frequency of past behavior can be used for determining the strength of habit³, which is in line with other economic and transport research (Abel, 1990; Carroll, Overland, and Weil, 2000; Rozen, 2010; Triandis, 1979). Another way to determine habit is self-reported (stated) behavior, as discussed in (Verplanken, Myrbakk, and Rudi, 2005). Verplanken et al. (1994) investigate the trade-off between habit and attitude for mobility consumption behavior. The paper introduces a survey with around two hundred participants and applies a script-based approach to measure habit. The study's findings align with the standard mobility consumption theory of Triandis (1979) that habit and attitude are negatively correlated. The study clearly states that habit is an important factor for determining individual mobility consumption behavior. Şimşekoğlu, Nordfjærn, and Rundmo (2015) clusters mobility users in Norway through survey data to examine the role of attitudes and habits on mobility consumption behavior. The study presents two clusters: private car users and public transportation/soft mobility users. The

³The strength of the habit here refers to the extent of the impact of past consumption on current utility.

results demonstrate that users have been influenced in their mobility consumption by their habits. Chen and Chao (2011) state that reducing private car (PC) usage and encouraging people to use public transportation and soft mobility modes are fundamental policy goals to cut mobility emissions. The paper asserts that habit-driven mobility consumption exists, hindering individuals' intentions to switch from car to public transportation. Hoffmann et al. (2017) carry out a meta-analysis to understand the cognitive mechanism behind the mobility consumption behavior. The article reviews 43 relevant studies and states that habit strongly predicts private car consumption. Nonetheless, it might weaken when existing circumstances change. For instance, moving location can weaken habits, which is in line with the previous research of (Verplanken and Wood, 2006; Bamberg, Ajzen, and Schmidt, 2003). In the same vein, Moser, Blumer, and Hille (2018) states that mobility consumption is highly habit-driven, and it can be reduced via disrupting habit-driven choice. The study conducted a longitudinal survey in Switzerland to observe the change of habit strength. Private car users were given free e-bikes to try. After a year, it was observed that some users' car habits were weakened, and they started to use the e-bike. In another research, Idris, Habib, and Shalaby (2015) demonstrate that traditional attributes such as cost and time have lower importance than habits for mobility consumption behavior. Schubert et al. (2022) employs structural equation modeling as the methodology to analyze the underlying mechanisms of the mobility mode-choice behavior. Habit is included in the model as a latent variable. The results revealed that habits show the most significant impact on the mobility behavior.

2.2.2 Habit and policy response

Maréchal (2009) provides empirical evidence that the existences of habits can explain the low effectiveness of traditional price-based climate policies than their ex-ante estimations. The paper also mentions that the existence of habits can move individual choices out of the optimal option that rational theory suggests (i.e., bounded-rationality), similarly to (Simon, 1972). The paper states that unsustainable habits should be broken for the climate targets, which can be done via disruption, as Friedrichsmeier, Matthies, and Klöckner (2013) and Moser, Blumer, and Hille (2018) suggest. Gsottbauer and Bergh (2011) assert that some individuals do not react to price signals and do not take into account costs because decisions are influenced by various biases such as habits. Idris, Habib, and Shalaby (2015) provide empirical evidence that habits lead to the domination of a specific choice even in cases where the rational choice favors another option. The findings are in line with (Di and Liu, 2016; Sharmeen and Timmermans, 2014).

2.2.3 Heterogeneity among mobility users

Zhou et al. (2020) investigate the heterogeneity in mobility users' characteristics and its influence on car sharing. The results show that well-educated individuals are more likely to use car-sharing services, while individuals with more daily trips (e.g., daily commuters) are less likely to use them. Hausteine (2012) explores the heterogeneity in the behavior of the elderly mobility users. The paper implements segmentation research to group mobility users based on their mobility consumption behavior and characteristics. It presents the heterogeneity among the users at the group level by analyzing the characteristics of segments. Ma et

al. (2015) analyze the heterogeneity in mobility users' behavior in Luxembourg and examine the influence of socio-economic characteristics on the mobility consumption behavior. The results show that private car users perform fewer daily distances, and cost plays a lesser role in their private car consumption. On the other hand, train users to perform longer daily distances. De Ona et al. (2015) explore the heterogeneity in the perceptions of railway users in Italy. The results show that the perceptions about the quality of service are heterogeneous among different user segments.

Although there are related contributions, analyzing the characteristics of habit-driven and cost-sensitive mobility users is still lacking that we perceive as a research gap to fill in the present chapter. Particularly in Switzerland, where this chapter focuses, there is no such study. The analysis can enable us to explore the origin of behavioral heterogeneity, potentially expanding the body of mobility mode choice behavior.

2.3 Data

The Swiss Federal Office of Spatial Development (ARE) conducts quinquennial surveys to observe the Swiss population's characteristics. The surveys are released with the name Mobility and Transport Micro-census (MTMC) (ARE/BfS, 2017). The surveys contain information about Swiss households in various dimensions such as their characteristics, mobility activities, resources (e.g., private car, travelcards, parking-place), stated preferences and motives, etc. The MTMC is used as the primary data source throughout this chapter. There are already studies in the literature that employ the MTMC data for the analysis like (Baumgartner et al., 2019; Becker and Axhausen, 2018; Janzen and Axhausen, 2018),

but none of them addresses the issues discussed in this chapter.

The latest release of the MTMC data in 2015 consists of 57090 Swiss households along with information describing their characteristics in various dimensions (ARE/BfS, 2017). The data also contains mobility activities (i.e., trips) of the households on the interview day. The trips can have different properties such as the primary mode of mobility, trip distance, departure time, etc. They consist of stages. That is, a person can perform a trip through multiple stages. For instance, a commuter respondent can have a trip from the domicile to the work-space with three stages; from domicile to the nearest train station, from the nearest train station to the train station near to the work-space, and from the second station to the work-space. So, he might perform a single trip with three separate stages. Each trip stage in the data is covered by a single means of mobility. The data contains a specific table for all the stages and various information in different dimensions describing the characteristics of stages such as distance, the chosen mode, departure, duration, location, etc.

2.3.1 Sample Selection

For each trip stage performed via private car (PC) and public transportation (PT), the respondents who declared their primary mode-of transport as PC and PT were asked the following questions:

- What is the main reason to use private car?
- What is the main reason to use public transportation?

The questions were asked to the respondents who had at least a trip on the interview day and used private car or public transportation to perform the trips. So, the respondents who use soft mobility modes (e.g., walking, bike, scooter,

etc.) have not received these questions. Besides, the respondents with no trips on the interview day have not received these questions. In total about three thousand two hundred respondents have received and answered these questions. We filtered the respondents who stated habit and cost as the main reasons for all their PT and PC trip stages for the analysis. They constitute about 13% of the overall sample (440 households). Thus, we created two distinct user segments to compare their characteristics, enabling us to explore their differentiating motives for mobility consumption. The respondents who stated other reasons such as time, comfort, easiest option, etc. for their PT and PC use have not been considered in the final sample. To sum up, the final sample consists of Swiss households who have PT and PC trips on the interview day and declared habit or cost as their mobility mode-choice reasons. So, they are a relatively small part of the overall survey, which will be discussed in the limitation section.

2.4 Methodology

2.4.1 Multiple Logistic Regression Analysis

As for the methodology, we carried out the multiple logistic regression analysis. It allow us to analyze the significance of the ex-ante chosen variables on the likelihood of being a habit-driven or cost-sensitive mobility user with statistical significance (Hidrue et al., 2011). It is a widely used methodology in case there exist multiple explanatory variables, which can be ordinal, nominal, or continuous, and a single response variable with two discrete (i.e., dichotomous) outcomes (Hosmer Jr, Lemeshow, and Sturdivant, 2013). The methodology was

carried out in various studies in the economics literature that aim to answer similar research questions to the present chapter but in different sub-fields of study in economics, such as (Yakovlev, 2018; Wen and Cheng, 2013; Eriksson et al., 2016; Huang and Thakor, 2013; O'Brien and Jones, 1999).

2.4.2 Model Specification

As mentioned in Section 2.4.1, we used two specific questions from the MTMC data. The first question asks about the reason to use private car, and the second one asks about the reason to use public transportation. The first question is asked to the respondents who have trips by private car on the interview day, and the second one to those who have trips with public transport. That is, the questions are independent. So, we obtained two independent samples after filtering the respondents who stated habit and cost as the answer to these questions. Then we built two distinct multiple logistic regression models (Model 1 and Model 2, whose results will be introduced in the following section). Model 1 is built with the sample who state their private car use reasons as habit or cost. Model 2 is built with a different sample who stated their public transportation use reasons as habit or cost. So, the response variable of the models is the stated reasons with two classes (categories), habit and cost for PC and PT users, respectively. The explanatory variables are the same for both models, which are listed below;

- Income Level
- Degree of Urbanization of the Residence (DEGURBA)
- Household Size

- Age
- Daily Distance

The dichotomous response variable enables us to create two segments with different mobility consumption motives and compare them based on a set of explanatory variables (Hosmer Jr, Lemeshow, and Sturdivant, 2013). Explanatory variables depict the characteristics of respondents. Since the MTMC data contains a vast number of variables in different dimensions, we singled out a list of them (i.e., ex-ante choose) as they are frequently used in related studies comparing different user segments and exploring the origin of habit-driven private mobility behavior (Li et al., 2016; Elhorst, Madre, and Pirotte, 2020; Osborne, 2005; Cherchi and Cirillo, 2014). For the sake of clarity, we note again that the first model is built with the respondents who answered the first question mentioned in the previous section (i.e., the reason of choosing private car). The second model is built with those answered the second question (i.e., the reason of choosing public transport). The models enable us to analyze whether certain explanatory variables increase or decrease the likelihood of being in habit-driven or cost-sensitive PC or PT user.

$$P = \frac{e^{\beta_0 + \beta_1 X_{IncomeLevel} + \beta_2 X_{DEGURBA} + \beta_3 X_{HouseholdSize} + \beta_4 X_{Age} + \beta_5 X_{DailyDistance}}}{1 + e^{\beta_0 + \beta_1 X_{IncomeLevel} + \beta_2 X_{DEGURBA} + \beta_3 X_{HouseholdSize} + \beta_4 X_{Age} + \beta_5 X_{DailyDistance}}} \quad (2.3)$$

The Eq. 2.3 illustrates the mathematical form of the logistic regression. P is the expected probability that the outcome events (i.e., being habit driven PC user

for Model 1 and PT user for Model 2) is present; X values are distinct explanatory variables, and β_0 through β_5 are the corresponding regression coefficients.

We controlled the multicollinearity, which occurs when explanatory variables in the model are correlated. It was controlled because it encumbers the analysis by reducing the estimated coefficients' precision, which weakens the statistical power of the model (Graham, 2003; Farrar and Glauber, 1967). In other words, it undermines the statistical significant of an explanatory variable, which hampers to get accurate modeling results. Since we have a mixed-type explanatory variable set (i.e., both categorical and continuous), we applied different statistical tests to control the correlation among explanatory variables (multicollinearity). To control categorical-categorical peers, Pearson's chi-square test was employed (Urdan, 2016). The frequency of each category for one categorical variable is compared across the categories of the other variable. For the continuous-categorical control, one way ANOVA test was applied (Lee, Lee, Lee, et al., 2000). In-group (category) and intra-group variances are computed and compared. For the explanatory continuous variables, the variance inflation factor (VIF) was calculated (Faraway, 2002). The VIF measures the multicollinearity for the regression models but only with continuous explanatory variables. Thus, we built regression models only with continuous explanatory variables and obtained their VIF value. All the results showed us that there is no significant multicollinearity among explanatory variables. The results are illustrated in the Appendix section of this chapter.

2.4.3 Descriptive Statistics

The first explanatory variable in the list is income level that consists of nine levels (see Table 2.1). As Figure 2.1 shows, the proportion of habit-driven private car users in higher income levels is higher than cost-sensitive ones'. We analyze in the following section with the regression analysis whether a certain income level increases the likelihood of being a habit-driven and cost-sensitive mobility user.

Income Level	Interval
1	Below 2.000
2	2000 to 4000
3	4001 to 6000
4	6001 to 8000
5	8001 to 10000
6	10001 to 12000
7	12001 to 14000
8	14001 to 16000
9	More than CHF 16000

TABLE 2.1: Income Levels in Swiss francs according to the MTMC data

The second explanatory variable to be used in the analysis is the degree of urbanization of the residence (DEGURBA). The variable contains categorical values for the urbanization level of the living addresses of the households (see Table 2.2). The urbanization levels are defined according to the standard definition of Eurostat. Residential areas can provide different PT services according to their urbanization levels (e.g., city centers have tram, train, and bus while rural areas have mostly only bus). Likewise, they can have different levels of congestion as well as PT frequency based on their urbanization characteristics. As Figure 2.2 illustrates the proportion of habit-driven PC users in city centers is lower

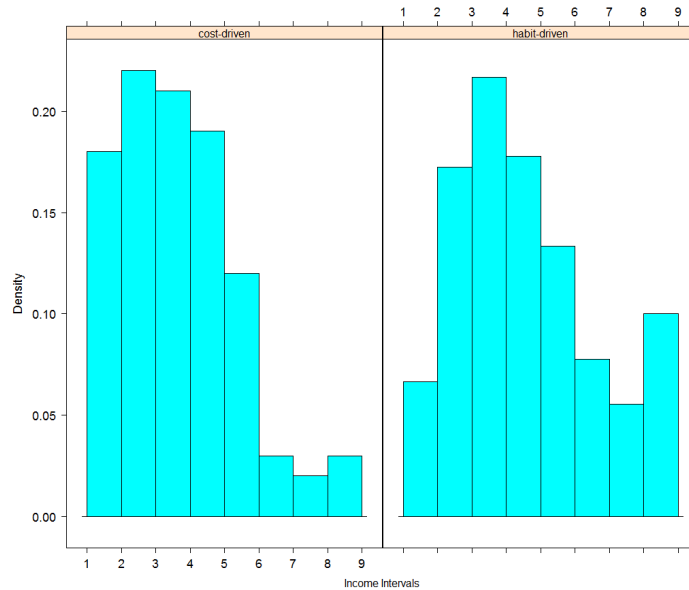


FIGURE 2.1: Income level distributions of the habit-driven and cost-sensitive private car users

than cost-sensitive ones'. Smaller cities and suburbs is the most frequent level in both groups. The proportion of cost-sensitive users in rural areas is lower than habit-driven ones'. In the following section we analyze whether any level of the degree of urbanization increase or decrease the likelihood of being habit-driven or cost-driven mobility user.

DEGURBA number	Category
1	Cities (densely populated area)
2	Smaller cities and suburbs (areas with average population density)
3	Rural areas (sparsely populated areas)

TABLE 2.2: Degree of urbanisation of the residence categories according to the MTMC data

The third explanatory variable is the household size that is a continuous variable. Increasing household size might impact the motive of mobility consumption behavior. The mean household size of habit-driven private-car users is

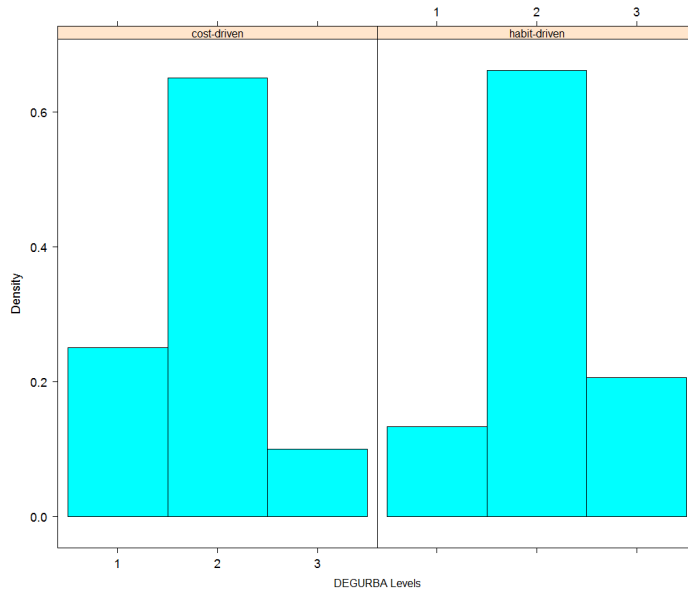


FIGURE 2.2: Degree of urbanization of the habit-driven and cost-sensitive private car users (1- Cities, 2-Smaller cities and suburbs, 3-Rural areas)

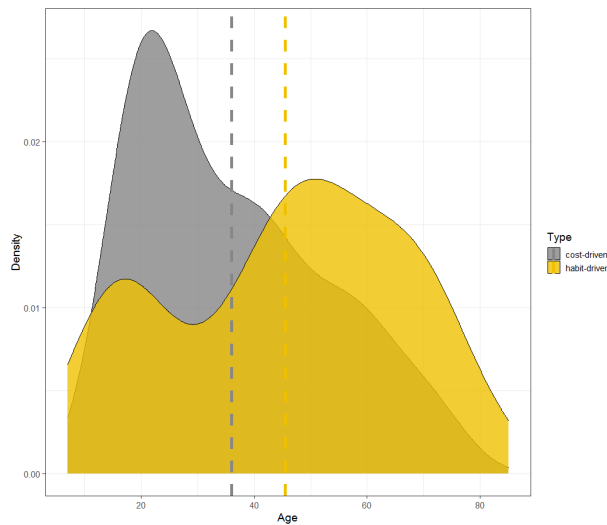


FIGURE 2.3: Age distributions of the habit-driven and cost-sensitive users

about 2.89 while it is slightly higher for cost-sensitive ones, 3.04. In the next section we analyze whether this difference is statistically significant. The fourth explanatory variable is age, which we used to examine whether the young and

the old have different reasons for their mobility consumption behavior. In other words, that variable is used to understand whether different age groups are associated with two user segments. Figure 2.3 show the age distributions of habit-driven and cost-sensitive PT users. As shown in the figure, the mean age of habit-driven PT users is higher than cost-sensitive ones'. Similar to other continuous variables, we will analyze whether this difference is statistically significant in the next section. The final explanatory variable is the daily traveled distance (i.e., kilometers). That is also a continuous variable that we used to analyze whether individuals having short and long daily distances have different motives for their mobility consumption behavior. As Figure 2.4 shows, cost-sensitive PC users declare more daily traveled distance than habit-driven ones.

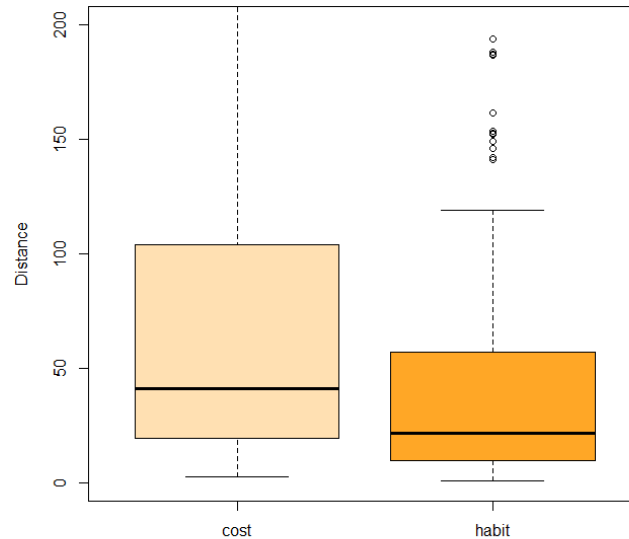


FIGURE 2.4: Daily distances of the habit-driven and cost-sensitive PC users

2.5 Results

The section discusses the results from Model 1 and Model 2. Please consider that there is one more category for the categorical variables listed in the regression summary below in Table 2.3. This means that the unlisted category is the reference one, as explained in (Hosmer Jr, Lemeshow, and Sturdivant, 2013).

Variable	Model 1			Model 2		
	Estimated Coefficient	p-value	Odds ratio	Estimated Coefficient	p-value	Odds ratio
2000-4000 CHF	2.559704	0.00387**	12.93	1.323342	0.323342	-
4001-6000 CHF	0.452159	0.56774	-	-0.191667	0.828853	-
6001-8000 CHF	-0.251424	0.73520	-	0.354415	0.648798	-
8001-10000 CHF	0.430768	0.51828	-	-0.049760	0.945751	-
10001-12000 CHF	-0.548924	0.36201	-	-0.679343	0.316214	-
12001-14000 CHF	0.742035	0.17366	-	0.009077	0.986758	-
14001-16000 CHF	0.248397	0.59209	-	0.294413	0.561712	-
More than CHF 16000	0.257279	0.47615	-	0.880288	0.080641	-
Smaller cities and suburbs	0.367956	0.01640*	1.44	-0.502944	0.233028	-
Rural areas	1.518642	0.0025**	4.56	-0.128435	0.840580	-
Household Size	-0.222739	0.07175	-	-0.142504	0.415183	-
Age	0.006625	0.48174	-	0.045083	0.000445 ***	1.04
Daily Distance	-0.005142	0.01601*	0.99	-0.003135	0.119610	-
N	280			160		

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, . $p < 0.1$

TABLE 2.3: Regression Results (Only odds ratios of significant variables are demonstrated)

Model 1 revealed that the second income class differs significantly between habit-driven and cost-sensitive PC users. The odds of being a habit-driven PC user in the second income category are around twelve times that of being a habit-driven PC user in the reference income category. Nonetheless, other income classes do not increase or decrease being habit-driven or cost-sensitive mobility user. The proportion of the cost-driven PC users in city centers is more than the habit-driven ones, while for rural areas, it is the opposite. The odds of being a habit-driven PC user in smaller cities/suburbs are about one and half times that of being in cities. In the same vein, the odds of being a habit-driven PC user in rural areas are about four and half times that of being in cities. That is, using a PC out of habit is more likely than out of cost in rural areas, while the situation is the opposite in cities. Thanks to the multicollinearity control that we carried

out, we can also say that we observed no correlation between income level and DEGURBA (see Table 2.5 in Appendix); thus, the relationship between the DEGURBA categories and being habit/cost driven is independent of income level. Model 1 revealed also that increasing daily distance decreases the likelihood of being a habit-driven PC user significantly. In other words, the model showed that the difference between the means of the daily distances of habit-driven and cost-sensitive PC users in Figure 2.4 is statistically significant. Holding all else constant, one unit increase in daily traveled distances will have a 0.99 unit decrease in the odds ratio of being a habit-driven PC user. Model 2 revealed that the only significant variable that increases the likelihood of being a habit-driven PT user is age. As illustrated in Figure 2.3, the most frequently observed age of the habit-driven PT users is around 60, which is approximately 20 for the cost-sensitive ones. The means of both groups are 45.5 and 36.1, respectively. Younger respondents have a higher likelihood of being a cost-sensitive PT user. Holding all else constant, one unit increase in age will have a 1.04 unit increase in the odds ratio of being a habit-driven PC user. Similarly to DEGURBA, we controlled the correlation between age and income level (see Table 2.6 in Appendix). The result shows us that the relationship between age and being a habit/cost-driven PT user is independent of income level.

Model	X-squared	p-value
1	6.2209	0.6225
2	8.0654	0.4271

TABLE 2.4: The Hosmer-Lemeshow goodness of fit test results

We applied the Hosmer-Lemeshow (Hosmer Jr, Lemeshow, and Sturdivant, 2013) goodness of fit test for our multiple logistic regression models (Model 1 and Model 2). It is a widely used standard test for logistic regression models

(Harrell Jr, 2015; Hosmer Jr, Lemeshow, and Sturdivant, 2013). We aimed to determine whether a set of observed values match those expected under the applicable models. The null hypothesis for the test is that the observed and the expected proportions are similar, which means that the model is correctly specified. The test produces a p-value—if it is low (e.g., below .05), we reject the null hypothesis. If it is higher, then our model passes the test. The test results are illustrated in Table 2.4. The p-values indicate that there is no evidence for a poor fit. That shows us that the models are indeed correctly specified.

2.6 Discussion

The chapter analyzes the origin of the behavioral heterogeneity observed between two groups of individuals by examining their characteristics. That is, there is an observed behavioral heterogeneity among the two groups; they have different reasons/motives for their mobility behavior. The chapter aims to examine the source of that behavioral heterogeneity by analyzing the socio-economic and demographic variables in the characteristics. In other words, it rationalizes the behavioral heterogeneity by analyzing the structural heterogeneity in the characteristics of both groups. To be more concrete, the chapter explores whether certain ex-ante chosen explanatory variables increase or decrease being habit-driven or cost-sensitive mobility user. The results show that these specifically analyzed two segments (i.e., habit-driven and cost-sensitive) have differences regarding the distribution of the explanatory variables. For instance, the distribution of age among habit-driven and cost-sensitive PT users is significantly different, as Model 2 shows. Or the distribution of traveled distances among habit-driven and cost-sensitive PC users is significantly different,

as Model 1 shows. The chapter analyzes two behavioral segments empirically. Nevertheless, the chapter does not quantify the behavioral heterogeneity. That is, it does not analyze the quantity of behavioral heterogeneity among Swiss households in terms of their mobility behavior (i.e., how many user/consumer segments we have based on mobility behavior is not explored). That will be analyzed in Chapter 3 in detail. From that aspect, Chapter 2 provides input for Chapter 3.

An important take away message of the chapter is that the representative agent models are not sufficient to model mobility behavior of Swiss households since we observed heterogeneous behavioral groups. Thus, the chapter also signals Chapter 4 regarding the heterogeneity to be modeled in the model. The model, which will be built and explained in Chapter 4 in detail, captures behavioral heterogeneity among Swiss households. Agents in the model have different preferences, emotions, habit stocks that impact their mobility behavior. So, agents differ from each other in terms of their mobility behavior. Besides, this chapter shows that a group of individuals declare their mobility consumption motive as a habit. Following the assumption that their mobility behavior is motivated by habit, the model involves a time-inseparable utility function that enables habit formation. That is, agents' past mobility consumption can partially affect the current utility in the model. More discussion on it will be provided in Chapter 4.

2.7 Limitations and Future Work

The most important limitation of the chapter is measuring habit. The MTMC data contains the mobility activities of the respondents for a particular day. It

is "one-shot" data. So, there is no continuity. Therefore, we could not measure habit as a latent variable for which we need multiple data points in the past belonging to the same individual. We used the stated reasons to identify the respondents whose mobility consumption is motivated by habit. An extension of the present chapter would be to measure the habit as a latent variable through a dataset with continuity and compare it with the stated reasons to strengthen the analysis. Another important limitation is the sample size of the selected households. We filtered the households who stated their mobility mode choice reasons as habit and cost. Nonetheless, they constitute about 13% of the overall responses. There are more popular answers to that question than habit and cost. However, since the goal of the study is to compare the characteristics of these two groups due to the implications from the literature, they have been picked up on purpose. Finally, a future work target would be quantifying behavioral heterogeneity in the mobility behavior of Swiss households. Chapter 3 carries out that extension by conducting a segmentation study. The final extension would be building an economic model that captures behavioral heterogeneity among agents. Chapter 4 presents an economic model that includes both behavioral and structural heterogeneity among agents. Agents can have different socio-economic and demographic attributes as well as emotions, preferences, and habit stocks in the model.

2.8 Conclusion

The overall chapter aims to explore the behavioral heterogeneity in mobility consumption of Swiss households. This chapter analyzes the characteristics of Swiss individuals who declared their mobility consumption motives as habit

and cost. That is, the chapter filters the habit-driven and cost-sensitive respondents in the MTMC data; thus, it creates two user segments with different consumption motives. Then, it analyzes the characteristics of these two behavioral segments to observe differences. It examines whether chosen variables of characteristics increase or decrease to be a habit-driven or cost-sensitive mobility user. The findings show that the likelihood of being a habit-driven private car user decreases with the increasing daily traveled distance and increases in rural areas and smaller cities. Age is the only significant variable that increases to be a habit-driven public transport user. The overall findings show that there are two behavioral segments (groups) which have heterogeneity in the characteristics. The next chapter will quantify the number of such segments among Swiss households regarding their mobility behavior and characteristics. Chapter 4 considers that heterogeneity in economic modeling.

2.9 Appendices

Appendix A Test Results

Variables	X-squared	df	p-value
Income level - DEGURBA	159.01	22	2.2e-16

TABLE 2.5: Pearson's Chi-squared Test Results

	Income Level	DEGURBA
Household Size	2e-16 (2828.9)	2e-16 (206.1)
Daily-distance	2e-16 (16.47)	0.0523 (2.95)
Age	2e-16 (4671)	3.12e-07 (14.98)

TABLE 2.6: One Way ANOVA Test Results (*F-Values are in the parenthesis)

	Household Size	Daily Distance	Age
Household Size	-	0.007178006	-0.5602912
Daily-distance	0.007178006	-	-0.01573075
Age	-0.5602912	-0.01573075	-

TABLE 2.7: Pearsons' Correlation Test Results

	VIF Factor
Model 1	1.023035
Model 2	1.130049

TABLE 2.8: VIF values for both models

Chapter 3

Market Segmentation: Obtaining Swiss mobility segments to analyze heterogeneity at the group level¹

Abstract

This chapter carries out market segmentation research to analyze heterogeneity at the group level among Swiss households in terms of their mobility consumption behavior and characteristics. That is, it identifies the group of individuals whose mobility behavior and characteristics are similar. It thus quantifies the observed behavioral heterogeneity in the previous chapter. That enables us to learn about the amount of behavioral heterogeneity among Swiss households and consider it for the model building in the following chapter. In methodological terms, the chapter develops and presents a new structured market segmentation procedure. Then it applies the developed procedure to the Mobility and Transport Microcensus (MTMC) data to obtain the Swiss mobility consumer segments. That reveals thirteen consumer segments with different

¹A scientific paper based on this chapter has been peer-reviewed and published by the IEEE Xplore Library (Bektas and Schumann, 2019).

mobility behavior and characteristics. In four of the obtained segments, the primary mode of mobility is the private car. These segments constitute about 60% of the overall population. The chapter discusses them in particular, because whose mobility consumption behavior should be targeted primarily to reduce mobility emissions. They are the relevant target audiences for decarbonization policies. The overall chapter provides both empirical and methodological contributions that improve the body of knowledge on mobility consumption behavior and market segmentation. Empirical contributions can be used for decarbonization policy formulation (e.g., targeting different segments with tailored policies). They can also be used in economic modeling studies like the following chapter to determine the degree of heterogeneity modeled in an economic model, e.g., using segment-level representative agents instead of a single representative agent for the whole population. Methodological contributions can be beneficial for future market segmentation studies.

Keywords: heterogeneity, segmentation, economic modeling, mobility behavior, cluster analysis, climate change, policy analysis

JEL Classification: C38, C61, C63, C80

3.1 Introduction

There is clear scientific evidence that Greenhouse Gas (GHG) emissions from economic activities, especially burning fossil fuels for energy, are the main

drivers of climate change that has destructive environmental and economic impacts (Stern, 2007). Like other disciplines, climate change is also a growing concern in economics. Many countries have ambitious emission reduction commitments, so they put extensive effort into transforming their economies towards lower emissions. Marginal emission reductions would not be enough to mitigate the impacts of climate change; we need massive GHG emission reductions (IPCC, 2018). Therefore, all possible emission reduction potentials in any economic sector should be analyzed and considered.

A part of the effort in climate change is about the mobility sector, where scientists and policy-makers have been motivated by the fact that the sector is one of the most significant contributors to GHG emissions. The sector accounts for around 16% of the GHG emissions and around 23% of the CO₂ emissions globally (Ritchie and Roser, 2020). The share of mobility emissions in the total emissions is even higher in Switzerland; the sector accounts for around 40% of the domestic GHG emissions and around 48% of the domestic CO₂ emissions (Boulouchos et al., 2021). While other major sectors of the Swiss economy, such as industrial production and construction, could manage to reduce their emissions in the period of 1990-2019, the mobility emissions grew in that period due to the growth in the demand. Considering Switzerland's climate change commitments, which is achieving net-zero emissions by 2050 (Bundesrat, 2021), a massive emission reduction in the mobility sector is indispensable.

The advancements in technology help countries transform different sectors of the economy towards lower GHG emissions by way of increasing efficiency in production and consumption. If we look at the mobility sector, in particular, these advancements have improved the efficiency of vehicles thanks to developments in different aspects such as more efficient engines, lightweight materials

in vehicle's body and chassis, more environmentally compatible power trains (e.g., hybrid, full electric) and so on. For instance, in Switzerland, the CO₂ emissions per car decreased on a yearly average basis by 1.5% per year until 2006 and 3.6% in the period of 2007-2015. The fleet is partially renewed every year and becomes lower-emitting relative to the previous year thanks to the newly registered vehicles that are more efficient. This transformation in the fleet has a positive impact on mobility emissions. However, achievements on the technology side alone will not lead to reaching net-zero emissions in the mobility sector (EEA, 2015). For instance, the emissions from the mobility sector in Switzerland continued to grow in the last decade due to increasing mobility demand, although the fleet has become more efficient and less emitting per vehicle kilometer (vkm). Emission reductions out of the efficiency improvements in the fleet fell far behind emission increases due to the growth in demand. Therefore, changes in individual mobility consumption behavior also play a pivotal role in reducing mobility emissions. That is, individuals should be incentivized to change their mobility consumption behavior and to move from unsustainable modes (e.g., vehicles with fossil fuel engines) to environmentally compatible ones (e.g., soft mobility). The aggregate of individual-level behavior changes would lead to vast demand shifts towards low emission modes (IPCC, 2018). Hence, it is needed to formulate economic policies targeting the demand side of the mobility sector to change individual behavior and demands, in addition to advancements in technology that largely transform the supply side of the sector (e.g., more efficient vehicles, more comfortable public transportation, etc.).

Heterogeneity among individuals should be kept in mind while formulating policy to change individual mobility consumption behavior. Individuals can

differ in many dimensions, such as preferences, resources, behavior, or overall characteristics. As a consequence of these differences, they can respond to the policy differently, or even some groups of individuals do not respond to it (Pennings, Garcia, and Irwin, 2005; Kirman, 2006). So, the impact of policy implementation may fall short of its ex-ante projection (Maréchal, 2010; Ferraro and Miranda, 2013). Therefore, it is beneficial to ex-ante analyze heterogeneity among individuals and consider the results for the policy formulation as well as for policy evaluation in economic models. In an optimal but extreme case, each individual would be targeted by a tailored policy separately according to his preferences, needs, characteristics, or behavior. That is, heterogeneity at the individual level can be considered in the policy formulation, and each individual is targeted according to his properties. However, it is almost impossible and even unnecessary to consider heterogeneity at the individual level due to the cardinality issue (i.e., the number of individuals). Therefore, analyzing and considering heterogeneity among individuals at the group level can be more appealing for policy formulation. Individuals with very similar preferences, needs, behavior, or overall characteristics, can be grouped together. Then, these groups can be targeted with group-specific policies to change their behavior, which is a very common way in marketing segmentation research (Wedel and Kamakura, 2012). Splitting similar individuals in different groups and targeting them separately would yield more behavior change and consequent aggregate demand shift towards low emitting modes (Anable, 2005).

This chapter aims to carry out a market segmentation study to analyze heterogeneity among Swiss households at the group level. So, the main research question to be answered in this chapter is what are the Swiss mobility consumer

segments (profiles)² which are relatively homogeneous internally and dissimilar from each other in terms of their mobility behavior and characteristics. The chapter employs the large dataset called the Mobility and Transport Microcensus (MTMC) (ARE/BfS, 2017). The dataset contains information in various dimensions, such as stated preferences, daily mobility routines (e.g., trips), mobility consumption behavior, demographics, resources (e.g., vehicle or seasonal ticket ownership), for approximately sixty thousand Swiss households. Since the dataset contains a vast number of variables, the chapter chooses a set of them ex-ante to be used in segmentation.

The chapter employs cluster analysis as for the methodology, which is a multivariate quantitative method often used for market segmentation (Saunders, 1980; Disdier and Tongeren, 2010). The method hinges on exploiting similarities among observation units (e.g., households) to split them into segments, in which internal heterogeneity (variance) is minimized. This chapter first creates pairwise dissimilarities among Swiss households based on each variable in the variable set. The households are placed in a latent space based on their similarities. For instance, if two households are similar in many dimensions (e.g., preferences, demographics, etc.) they stay close in the latent space. Then, the clustering algorithm splits households into segments according to their position in the latent space. The algorithm produces internally compact (homogeneous) and externally separated (heterogeneous) segments as output.

The chapter answers technical questions in addition to the research question to improve the body of knowledge on the application of cluster analysis in market segmentation studies. Previous economic studies often determine the number of segments ex-ante with informal statements. The chapter here

²The terms segment, profile, and group are used interchangeably throughout the thesis.

presents an explicit data-driven procedure to measure the quality of segments and determine the number of segment information. The quality is measured according to how good is intra-segment homogeneity and inter-segment heterogeneity. The procedure finds the optimal number of segments that provide the highest segmentation quality. So, the number of segment information is determined based on a data-driven approach. Secondly, the chapter answers the question of how the quality of segments can be further improved? Since the quality is measured based on how good intra-segment homogeneity and inter-segment dissimilarity are, the question boils down to how to further improve these two criteria. The presented procedure optimizes the clustering algorithm and maximizes intra-cluster homogeneity with a given variable set. That improves the existing practice and provides researchers a way to obtain segments whose numbers are determined by a data-driven approach and qualities are optimized by a state-of-the-art concept.

The chapter makes two main contributions. The first contribution is to the literature on the Swiss mobility sector and its decarbonization. The chapter is one of the first studies to explore and present the Swiss consumer mobility segments. The second contribution of the chapter is to the literature on the application of cluster analysis in economic studies. The chapter presents a new procedure to measure the segments' quality and find the optimal number of segments with a data-driven approach. The procedure optimizes the clustering algorithm to obtain segments with higher quality. That improves the existing practice and knowledge.

The following section discusses the application of cluster analysis in mobility research for market segmentation purposes. It also touches on the application of the methodology in other economic sectors and domains. Then, a new

overall concept for the market segmentation studies is introduced. Afterward, the concept is configured with the MTMC data. The findings section presents the obtained Swiss mobility consumer segments and their characteristics. The discussion section discusses the results and explains how this chapter links the previous chapter and the following chapter. Finally, the chapter ends with the limitations and conclusion section.

3.2 Literature Review

Cluster analysis methods are often used in mobility research to segment units being studied (e.g., households, firms, countries, etc.) according to their preferences, needs, behavior, or characteristics. It is a practical way to understand the characteristics of different market segments and develop strategies accordingly. The obtained segments are often suggested to be used for formulating tailored segment-level decarbonization policies that may lead to a higher behavior change and consequent demand shift towards lower-emitting modes. For instance, Soto, Cantillo, and Arellana (2021) use the K-means clustering algorithm for market segmentation for incentivizing mobility decarbonization policies in Colombia. The paper identifies three-car owner segments with different preferences, attitudes, and socio-demographic characteristics. The paper finally recommends strategies to target each segment by tailored policies considering heterogeneous characteristics and preferences. Similarly, Morton, Anable, and Nelson (2017) carry out a market segmentation of the emerging market for Electric Vehicles (EVs) by the combination of K-means and hierarchical clustering algorithms. The paper presents five household segments in the UK characterized by their preferences towards EVs, socio-economic characteristics (e.g., income,

age, gender, number of adults in the household, education level), current car ownership, and physiological traits. The paper asserts that there is a significant level of stratification in the preferences of households towards EVs. Therefore, a single policy targeting whole households is not efficient in accelerating EV diffusion. It finally suggests segment-level policies for each obtained segment considering their preferences and characteristics. Alonso-González et al. (2020) employ a latent class cluster analysis to segment Dutch households according to their attitudes and preferences towards the mobility as a service (MaaS) systems. The paper presents five segments concerning individuals' inclinations to adopt MaaS in the context of urban mobility. The paper projects the future diffusion direction of MaaS adoption based on the preferences and characteristics of the segments. Li et al. (2013) implement a market segmentation research for bicycle commuting to identify distinct bicycle market segments in China and develop specific policies or strategies for increasing bicycle usage in each segment. The paper employs the K-means clustering algorithm and presents six different segments with a different attitudes towards commuting travels. The paper finally suggests specific policies for each segment. Axsen, Bailey, and Castro (2015) explore the lifestyle and preference heterogeneity among potential EV buyers. The paper determines the number of segments as six and splits potential EV owners in Canada based on their preferences and lifestyles. Paper asserts that following a "one size fits all" approach to marketing, information campaigns, or policy design may backfire with some segments. Therefore, policies regarding EV incentives, e.g., subsidies, high-occupancy vehicle lane access, or reduced electricity rates, will be valued differently by various segments. In brief, the starting point of a vast majority of cluster analysis applications for market segmentation studies in the mobility sector is learning about heterogeneity among the individuals

and using this information to formulate tailored segment-level policies accordingly. This chapter follows the same line of research. It carries out a market segmentation analysis for the Swiss households based on their mobility consumption behavior, and characteristics. Then, it presents the market segments and signals the policy-making side to use them to formulate segment-level decarbonisation policies.

The studies with application of cluster analysis in other economic sectors or different research areas of economics constitute a large body literature. Therefore, a specific JEL code (C38) is allocated for these studies. The common characteristics of the large part of these studies is to understand the group level behavior of units being studied (e.g., firms, countries, households, banks, etc.). For instance, Boughaci et al. (2021) use the K-means clustering algorithm to segment credit applicants into segments according to their characteristics. The obtained segments are then used to create credit scores and bankruptcy predictions. The number of segments in this paper is chosen arbitrarily. Lutz and Newlands (2018) carry out consumer segmentation analysis within the scope of sharing economy. The paper segments home sharing customers based on their needs and characteristics. Sütterlin, Brunner, and Siegrist (2011) use cluster analysis to segment Swiss energy consumers based on their energy-related behavioral characteristics. The paper presents six different energy consumer segments, their characteristics, and specific marketing and policy strategies for each segment.

Previous studies rarely discuss measuring the quality of segments and how to improve it. They do not assess how homogeneous the instances in the same segment are and how different instances in different segments are. Considering the assumption that segments are internally homogeneous, it is relevant to measure the extent of intra-cluster homogeneity. This chapter, therefore, provides

a data-driven segmentation procedure, which is discussed in the following section in detail, to measure the quality of obtained segments and improve it as much as possible. Thus, the chapter improves the existing practice in market segmentation studies. Moreover, the chapter is one of the first market segmentation research to analyze heterogeneity among Swiss households regarding their mobility consumption behavior and characteristics. Thus, it improves the body of knowledge and provides a basis for policy-makers to formulate tailored decarbonization policies targeting different market segments. Thus, the chapter contributes to the efforts in the decarbonization of the Swiss mobility sector.

3.3 The overall concept of a market segmentation procedure

This section presents a new procedure for market segmentation research. The procedure hinges on a quantitative multivariate method: cluster analysis. The procedure aims for segmenting units being studied e.g., households, firms, countries based on their behavior and characteristics. As for the cluster analysis method, the procedure employs the K-medoids clustering algorithm (Kaufman and Rousseeuw, 2009). The algorithm is also called as PAM (partition around medoids) in some studies (Schubert and Rousseeuw, 2019). The K-medoids algorithm is a variant of the K-means algorithm but it has certain advantages over the K-means (Park and Jun, 2009). Firstly the K-medoids is more robust to the outliers and noises (Kaufman and Rousseeuw, 2009). Besides, it presents medoids (i.e., exemplar or centroid) of each segments as output. A medoid is the most characteristics member of the segment by having the lowest pairwise

dissimilarity with other units in the same segment. Obtaining the most characteristic member of the cluster is beneficial to understand the common behavior, preference or characteristics in the segment.

The overall concept of the procedure consists of two steps. In the first step, the optimal number of clusters (segments) is obtained. Since the K-medoids algorithm is unsupervised, the number of cluster information must be pre-specified. The procedure uses a statistical method to determine the optimal number of clusters in an explicit way. In the second step, the procedure carries out an optimization process to maximize cluster quality. The process aims to minimize heterogeneity (variance) in the clusters and maximize the dissimilarity (separation) among clusters. The procedure presents internally compact and coherent, and externally separated clusters as output. The outputs then will be called market segments. So, the terms clusters and segments refer to the same thing throughout this chapter. Both steps and their stages are illustrated in Figure 3.1. The following parts explain each step of the procedure one after another for the sake of clarity.

The 1st Step

The first step aims for obtaining the optimal number of clusters with an explicit way. It measures the quality of clusters with different number of cluster values (i.e., candidates to be the optimal number of clusters). Then, it selects the optimal number of clusters according to the cluster quality; the value with the highest cluster quality is appointed as the optimal number of clusters. In general, the cluster quality is measured according to two criteria:

- Compactness: it reflects intra-cluster homogeneity. A higher intra-cluster homogeneity is an indicator of good clustering.

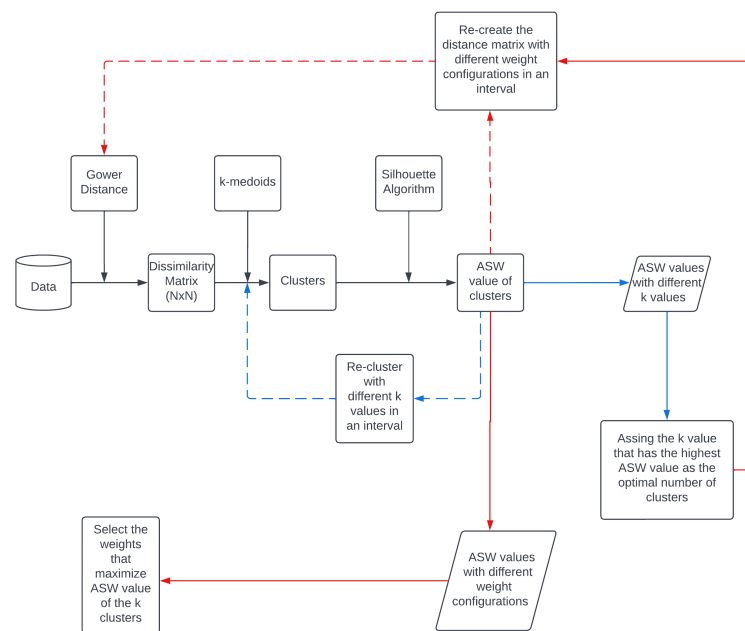


FIGURE 3.1: The Overall Concept of the Segmentation Procedure (contribution of the colors to the steps; blue: the first, red: the second, black: contributes to both)

- Separation: it reflects how-well a cluster is separated from other clusters. A higher inter-cluster dissimilarity is an indicator of good clustering.

The first step consists of stages which are illustrated with black and blue lines in Figure 3.1. It is operationalized by providing data as input. The data should have the individual level of granularity, e.g., households, firms, regions, etc. Ideally, the data should contain information about units in various dimensions such as preferences, behavior, and characteristics. According to the size of dimensions, a set of variables should be chosen ex-ante to be used in the procedure. A common way to choose a set of variables is using the factor analysis method that enables to reduce a large number of variables into fewer numbers of factors. However, the variables can be chosen heuristically according to the needs of particular research or based on implications from previous studies.

Once the procedure gets the input data, it creates pairwise dissimilarities among individuals in the data. It uses the Gower distance metric to create pairwise dissimilarities (Gower, 1971). The Gower distance metric is preferred to Euclidean in the procedure because it can create pairwise distances based on continuous, nominal and ordinal variables. The Euclidean distance metric can handle only continuous variables. For instance, it can not create pairwise distances among households based on their income levels since it is an ordinal variable. The Gower distance metric creates pairwise dissimilarities based on each variable in the variable set, combines and normalizes them between 0-1, and calculates the final pairwise dissimilarity. It calculates dissimilarities based on categorical (e.g., nominal and ordinal) variables based according to Equation 3.1 (Gower, 1971).

$$D_{i,j,k} = \begin{cases} 1, & X_{i,k} \neq X_{j,k} \\ 0, & X_{i,k} = X_{j,k} \end{cases} \quad (3.1)$$

Equation 3.1 calculates the pairwise dissimilarity between units (e.g, households) i and j based on the nominal variable k . The equation returns zero if both units have the same value on the variable k , one otherwise. For instance, if two units have the same gender, the equation produces zero dissimilarity for them in terms of gender variable. The dichotomous output can be adjusted for ordinal variables by dividing the scale. For instance, if an ordinal variable such as income level has 4 levels (e.g., 1, 2, 3, 4), the distance between 4 and 3 becomes 0.25 (Han, Pei, and Kamber, 2011).

$$D_{i,j,k} = \frac{|X_{i,c} - X_{j,c}|}{r_c} \quad (3.2)$$

The Gower distance metric calculates the pairwise dissimilarities for continuous variables via Equation 3.2. The equation calculates the difference between the values of two units on the variable c in absolute terms and divides it by the range. The range is simply the difference between the maximum and minimum values of the variable c among all units: $\max X_c - \min X_c$.

Calculated pairwise dissimilarities based on each variable are combined and normalized to obtain the final dissimilarity between two units. The Gower distance metric assigns weights for each variable in the combination phase. Weights represent importance of variables. If a variable has a higher weight, the dissimilarity based on this variable impacts the final dissimilarity more than other variables. As default the Gower distance metric assigns equal weight, which equals one, for all variables. The weights can be customized, which is the aim of the second step of the procedure, to improve the cluster quality. Equation 3.3 combines the dissimilarities based on all variables and computes the final value. In the equation, N represents the number of variables, w_k represents the weight of a particular variable k , and $D_{i,j,k}$ represents the dissimilarity between units i and j on the variable k .

$$D_{i,j} = \frac{\sum_{k=1}^N w_k D_{i,j,k}}{\sum_{k=1}^N w_k} \quad (3.3)$$

After all the final pairwise dissimilarities are created, the Gower distance metric puts them in a dissimilarity matrix ($N \times N$, N is the number of units) that functions as a latent space. The units with lower dissimilarity stay close in the space. The K-medoids algorithm gets this matrix as the input. It then splits units into k clusters based on their positions in the space (i.e., units staying cluster are

placed in the same cluster). Another input for the clustering algorithm is the number of clusters, which must be pre-specified. The procedure uses a particular statistics called the Average Silhouette Width (ASW) that measures the quality of clusters based on intra-cluster compactness and inter-cluster separation. It is a standard index to estimate the optimal number of clusters (Rousseeuw, 1987). Equation 3.4 calculates the Silhouette width value for a particular unit i . In the equation, a_i represents the average dissimilarity of the unit i with all other units in the same cluster, b_i represents the average dissimilarity of the unit i with all the units in the closest cluster to its cluster³. The equation returns a score between -1 and 1. A higher score indicates that the unit i is placed in the proper cluster. The average of the Silhouette Widths of all units provides statistics about the quality of clusters as well as the performance of the clustering algorithm.

$$Sil = \frac{b_i - a_i}{\max\{a_i, b_i\}} \quad (3.4)$$

The procedure uses the ASW statistic as a basis to determine the optimal number of clusters. It tries in a loop different options each with different number of clusters value (dashed blue line in Figure 3.1). It calculates an ASW score for each option. Then, the option that achieves the highest ASW value is determined. The value of that option is appointed as the optimal number of clusters. That is, the value that reaches the highest intra-cluster homogeneity and inter-cluster separation is appointed as the optimal number of clusters.

³Both a_i and b_i vary in the interval of [0-1].

The 2nd Step

The output of the first step of the procedure is the optimal number of clusters. The second step aims to improve further the ASW score of the optimal number of clusters. That is, it investigates how to minimize heterogeneity in the clusters. For that purpose, it optimizes the default weights of the Gower distance metric. As mentioned in the previous section, the Gower distance metric assigns equal weights to all variables by default. The second step aims to optimize these default weights to increase the ASW value of the optimal number of clusters. To find the optimal weights, the optimization function `optim` in R language is utilized, which is detailedly explained in the following section. According to each weight configuration, pairwise dissimilarities is re-calculated (red dashed line in Figure 3.1). That is, the positions of units change in the latent space. The K-medoids algorithm partitions units into k clusters (k here is the optimal number of clusters obtained in the first step). Then, the ASW value of the k clusters for each configuration is calculated and compared. The weight configuration that achieves the highest ASW value is appointed as the optimal Gower distance weights.

At the end of the two steps, the optimal number of clusters and the optimal Gower distance weights are obtained. Then, using this information, units are placed in clusters and presented as final market segments. In the following section, the procedure is applied with the MTMC data to segment Swiss households based on their mobility consumption behavior and characteristics. That is a case study of the procedure.

3.4 Detailed Configuration and Descriptive Statistics

This section presents an application case of the overall concept of the segmentation procedure with the Swiss data; MTMC (ARE/BfS, 2017). The goal is to segment Swiss households based on their mobility consumption behavior and characteristics. The following variables are chosen ex-ante to be used in the segmentation procedure in line with the goal. The variables describe mobility behavior and characteristics, which are also used in similar studies with mobility behavior segmentation purposes, such as (Haustein, 2012; Hunecke et al., 2010; Diana and Pronello, 2010; Morton, Anable, and Nelson, 2017).

1. Number of cars in the household
2. Number of daily trips
3. Mode-choice
4. Has Half-fare Travelcard
5. Multimodality
6. Average Daily Distance

The first variable in the list is the *number of cars in the household*. The variable indicates the availability and quantity of a resource (i.e., private car) that is relevant information for individuals while making economic decisions regarding their mobility activities. The variable can be a direct determinant of mobility consumption behavior because one can not consume a resource (or service, product) if it is not available for his use. Previous empirical studies revealed that

having more than a car in the household influence both private car and public transportation consumption (Li, Guan, and Wang, 2022; Mifsud, Attard, and Ison, 2017). Therefore, the variable is contained in the variable list to be used in the segmentation procedure. If we look at the distribution of the variable in the population, the majority of the Swiss households have one car (around 50%). Two and three-car available households come after. Around 8% of the Swiss households have no car available to be used for mobility activities.

The second variable is *the number of trips per day*. Previous empirical studies revealed that mobility behavior is associated with the number of trips per day (Loder and Axhausen, 2018). Therefore, the variable is considered for segmentation. The majority of the Swiss households perform two trips per day. They consist of one-fourth of the population. Around 4% of the population perform just one trip. It means that they do not turn back to their residential addresses. Some outliers perform up to twenty trips. However, overall, the bulk of the population performs between two and six trips per day.

The third variable is the *mode choice*, which is an essential variable because it implies mobility consumption behavior directly. It is a categorical variable including almost all the mobility modes in Switzerland (e.g., private car, train, tram, bike, walking, etc.). The MTMC respondents were first asked what their typical daily mobility activities are. After answering this question, they were asked how they carried out these activities. In other words, which mobility modes do they use for their daily mobility activities. A respondent can have more than an activity. It means that he can use more than a mode on an average day. For instance, he can prefer his private car to go to the workplace while walking to buy lunch. Therefore, the most frequently used mode is chosen as the mode choice for the respondents, using more than a mobility mode per day.

The mode choice variable indicates the primary mode of mobility for them.

The fourth variable is the binary variable of the *half-fare (HF) travelcard ownership*. The HF travelcard provides a 50% discount on traveling by public transportation modes (i.e., train, tram, bus, ships, etc.) in Switzerland. It is a resource that is relevant and known to economic agents (e.g., individuals) while making a particular economic decision for mobility activities. As previous studies revealed, the ownership of season tickets such as HF travelcard is central to mobility behavior (Loder and Axhausen, 2018). One could expect that HF travelcard owners tend to use more public transportation services more than non-owners from a cost point of view. In Switzerland, about 40% of the population has the HF travelcard. The ownership ratio is higher among the households who commute daily and lower in French and Italian speaking cantons of the country.

The fifth variable is *multimodality*, which is also a binary variable. The respondents in the MTMC data can have more than a trip on a particular day. They can use more than a single mode for these trips. The multimodality variable is zero for those using only a single mode for all their mobility activities and one for others. Approximately 70% of the Swiss population are mono-modals. A vast majority of mono-modals have their own properties, e.g., private car or bike. On the other hand, about 30% of the population uses at least two different modes to perform their mobility activities. Multi-modal respondents are more likely to be regular public transport users, live in city centers, and have an HF travel card.

The last variable is *the daily traveled kilometers*. It indicates the total traveled

distance on for routine mobility activities on an average day. It is a continuous variable. It impacts the mobility consumption behavior because individuals traveling long distances have different preferences than individuals traveling short distances such as comfort and speed (Carse et al., 2013; Chen et al., 2017). As a consequence of different preferences they can have different mobility behavior. As will be discussed in the following section, people having daily long commuting distance are more likely to prefer train over other modes. Conversely, people having short daily distance are more likely to prefer soft mobility modes such as bike, sharing systems, or walking. The Swiss population performs on average approximately 37 kilometers on a particular day. This average is higher among train users, as mentioned, people with longer daily distances prefer train over private car. The average daily distance is around 10 and 3 kilometers among bike users and walkers respectively.

After choosing mobility behavior relevant variables in the final variable set, the first step of the segmentation procedure is carried out to obtain the optimal number of clusters. The procedure creates pairwise dissimilarities among the Swiss households in the MTMC data by the Gower distance metric. It then generates the latent space (i.e., dissimilarity matrix) to be given to the clustering algorithm as input. The procedure tries different number of clusters candidates in the interval of [2-15]. In other words, based on the same latent space individuals are split into 2, 3, 4,...15 clusters respectively (see blue dashed line in Figure 3.1). The upper bound of the loop is set to 15 due to high computation time, which will be further discussed in the limitations section. For each iteration, the ASW value is calculated. The candidate that achieves the highest ASW score is appointed as the optimal number of clusters. The results of the process will be shown in the following section.

In the second step of the segmentation procedure, the following two functions are employed. Functions are described in the Appendix A in detail. Function A gets Gower distance weights as input, creates pairwise dissimilarities and the latent space accordingly. It then partitions units into k clusters, which is obtained as the optimal number of clusters in the first step. It finally measures the quality of the clusters and returns the ASW score as output. The function B gets default Gower distance weights and upper and lower bounds of the weights as input. It uses function A as the target function and aims to maximize its return. That is, the optim function tries all Gower weight combinations within the boundaries to maximize the output of Function A. Finally, function B returns the Gower distance weights combination that brings the highest return from Function A. These weights are appointed as the optimal Gower distance weights since they provide the highest ASW value, the highest intra-cluster homogeneity, and inter-cluster separation. The upper and lower bounds of the weights are set as 1 and 3, respectively ($w \in [1 - 3]$). The lower bound is set to the one due to the comparability reason with the default Gower distance weights. The upper bound is set to three to prevent the dominance of one variable in the characteristics of clusters.

3.5 Results

3.5.1 Optimal number of clusters and Gower distance weights

The overall concept of the segmentation procedure (see Figure 3.1) is implemented step by step with the MTMC data according to the configuration mentioned in the previous section. The results of each step are presented and discussed in this section. The first step of the procedure finds the optimal number of clusters with the given set of variables. Different values (i.e., candidates for the optimal number of clusters) in the interval of [2-15] are assessed, and the AWS score for each value is illustrated in Figure 3.2 below.

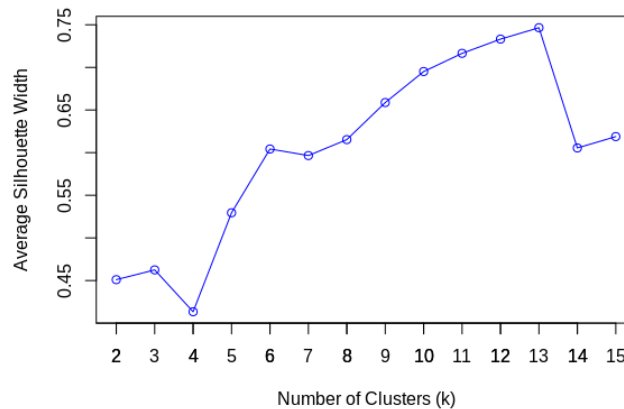


FIGURE 3.2: Number of clusters and the corresponding ASW values

The segmentation procedure yields the highest cluster quality when the households are divided into 13 clusters (ASW value=0.7465), as shown in Figure 3.2. The clustering quality is low when the number of clusters is few (e.g., 2, 3, 4, etc.). The reason is that there is high intra-cluster heterogeneity (variance) when households are divided into very few clusters. The clusters get relatively large

sizes; thus, some dissimilar households appear in the same cluster. The quality also becomes lower with a relatively higher number of clusters (e.g., 14 and 15). The reason is that some similar units might be placed in different clusters if the number of clusters becomes high. That influences the quality of clusters negatively in terms of weak inter-cluster dissimilarity. According to the results, both cluster quality criteria intra-cluster homogeneity and inter-cluster dissimilarity become optimal if the units are split into 13 clusters. Therefore, the optimal number of clusters is appointed as 13 at the end of the first step. The second step aims to improve the ASW score (0.7465) of the 13 clusters further. It finds the optimized Gower distance weights that maximize the ASW value obtained in the first step. The optimized Gower distance weights are illustrated in Table 3.1.

TABLE 3.1: The Optimized Gower Distance weights per variable

Features	Optimized Weights
Number of cars	1.000000
Has HF travelcard	2.469693
Number of Trips	1.000000
Daily Distance	1.000000
Mode-Choice	3.000000
Multimodality	2.640402

The optimization process finds relatively higher weights for some variables, such as mode choice and multimodality. It means that these variables are more distinctive and play more roles in splitting households. On the other hand, the optimization does not change the default weights of other variables such as daily distance, number of trips, and number of cars. Nevertheless, they still contribute to the structure of the clusters. Finally, pairwise distances are created based on the obtained Gower distance weights. Households are placed in a latent space based on these dissimilarities, and the K-medoids algorithm splits them into 13

clusters based on their positions in the space. The ASW score of the clustering with the optimized Gower distance weights is observed as 0.8458, which was 0.7465 with the default weights. The optimization process thus improves the quality of clusters significantly. In the end, the segmentation procedure presents the internally compact and coherent clusters that are well separated from each other. These clusters will be called the Swiss mobility consumer segments hereafter. The characteristics of segments will be discussed below in more detailedly.

3.5.2 The Swiss mobility consumer segments

The segmentation procedure has produced 13 Swiss mobility consumer segments with different behavior, sizes, and characteristics. Table 3.2 illustrates the dominant mobility behavior (i.e., primary mode of transport) information of each segment. Besides, it shows whether the individuals in a segment prefer multi/mono-mobility and have a HT travelcard. For the continuous variables used for segmentation (e.g., daily distance), Appendix B provides intra-segment density distributions.

TABLE 3.2: Some characteristics of the Swiss mobility consumer Segments

Segment Number	Primary mode of mobility	Multi/Mono-modality	HT Travelcard
1	Walking	Mono-modal	No
2	Private Car	Multi-modal	Yes
3	Tram	Multi-modal	Yes
4	Private Car	Mono-modal	No
5	Train	Multi-modal	Yes
6	Bus	Multi-modal	No
7	Bike	Mono-modal	Yes
8	Private Car	Multi-modal	No
9	Bus	Multi-modal	Yes
10	Walking	Mono-modal	Yes
11	Private Car	Mono-modal	Yes
12	Bike	Mono-modal	No
13	Train	Multi-modal	No

In four of the segments, the members' primary mode of mobility is the private car. They constitute approximately 60% of the Swiss population. They differ from each other due to other variables contributing to the definition of clusters, such as daily distance, multimodality, HF travel card ownership, and so on. The largest private car segment is segment four that alone constitutes about 27% of the population. The members of the segments are mono-modal; they only use the private car for all kinds of mobility activities. Almost no members of this segment have a HF travelcard, which is understandable because public transportation use is not typical in the segment. The members perform about 47 kilometers per day that is longer than the population average. Almost all the members have at least one private car in the household, and the majority of them have two cars. The members have, on average, two trips per day.

The second private car segment is segment eleven that constitutes around 12% of the population. Almost all the members of the segment as a HF travelcard. That is the distinguishing trait of the segment from segment four. The members travel around 48 kilometers daily. They are largely mono-modal. The majority have one car in the household that is lower than segment four.

The third private car segment is segment eight that constitutes around 10% of the population. The segment differs from other private car segments in terms of the multimodality variable. Almost all the members of the segment consume more than a mode on an average day to perform their mobility activities. Another distinguishing trait of the segments is the number of trips; the majority of the segments perform four trips per day that is more than the previously discussed private car segments.

The last private car segment is segment two, constituting around 6% of the population. The members of the segment are primarily multimodal, similarly

to segment eight and differently than segments four and eleven. However, almost all the members have an HF travel card differently than segment eight. The members of the segments have the highest daily traveled distance among all private car segments. They have, on average, four trips per day. The members are likely combining private car and public transportation, considering their multimodality characteristics (i.e., they use one more mode than private car to perform their mobility activities) and the HF travelcard ownership.

There are two segments whose members prefer walking as their primary mode of mobility. The first one is segment one that constitutes around 13% of the population. As one might expect, the members of the segment have relatively short daily traveled distances. The majority of the members have a car in the household and have no HF travel card. The members are primarily monomodal. The second walking segment is segment ten that constitutes about 9% of the population. The main difference between the characteristics of the two walking segments is that almost all the members of segment ten have the HF travelcard. Besides, most of the members of segment ten have four mobility activities per day, while it is two for segment one.

The segmentation procedure revealed two Swiss mobility consumer segments where the train is the primary mode of mobility. The first train segment is segment five that constitutes approximately 4% of the population. All the members of the segment are multimodal. That is, they use at least one other mode in addition to the train to perform their daily mobility activities. Similarly, all the members have HF travelcard. The members travel around 50 kilometers on a particular day, which is significantly more than the population average. They have, on average, two mobility activities on a particular day, although the share of members with more than four activities is significant. The second train

segment is segment thirteen that constitutes about 5% of the population. Different than the former train segment, the members of segment thirteen have no HF travelcard. Another difference is that the members travel on average around 106 kilometers on a particular day, which is the highest distance among all the obtained segments. Almost one-fifth of the members have no car at the household that is also the highest among the segments. Like the previous train segment, almost all the members are multimodal, but they do not have the HF travelcard. In brief, individuals traveling long distances are more likely to prefer train over other modes. That is in line with the findings of the previous studies, such as (Reichert and Holz-Rau, 2015; Cao et al., 2013; Garmendia, Ribalaygua, and Ureña, 2012) that demonstrate train is preferred for long distances due to speed and comfort.

There is only one segment whose members use the tram as their primary mode of mobility. That is segment three that constitutes around 3% of the population. One would assume that the members of the segment are more likely living in urban areas considering the availability of the tram in Swiss cities. Almost all the members use at least one another mode out of tram; so, they are multimodal. The members travel on average around twenty kilometers that are almost half of the population average. The majority of members have an HF travel card and four mobility activities per day.

There are two bus segments. Both are relatively small and constitute around 4% of the total population. The first one is segment six whose members are multimodal. They use at least one other mode than bus to perform their daily mobility activities. They have on average two trips per day. Almost a quarter of the members have no private car in the household. The members travel on average around twenty kilometers that is almost half of the population average.

They have no HF travelcard. The second bus segment is segment nine that differs from the former bus segments in terms of HF travel card ownership. All the members of segment nine have an HF travelcard. Besides, the members of segment nine travel more; the members travel on average 27 kilometers on a particular day.

Finally, the segmentation procedure produced two segments whose members prefer bike as their primary mode of mobility. The first bike segment is segment seven, whose members are largely mono-modal. They have HF travelcard and perform on average four mobility activities on a particular day. The majority of the members have a private car in the household. The second bike segment is segment twelve. It differs from the former because its members have no HF travelcard. The members of both bike segments travel around twenty three kilometers on an average day.

3.6 Discussion

The chapter carries out one of the first comprehensive market segmentation analyses for the Swiss mobility sector. It aims to observe heterogeneity in the mobility behavior and characteristics of Swiss households at the group level. It provides thirteen different household segments with different mobility behavior and characteristics. The segments are internally quasi-homogeneous; instances in the same segment have similar mobility behavior and characteristics. They are also well-separated from each other thanks to the presented segmentation procedure.

The previous chapter revealed that there are heterogeneous groups in the MTMC data in terms of their mobility behavior. However, it did not quantify

that. The present chapter therefore analyzed the heterogeneity among Swiss households in terms of their mobility behavior as well as characteristics, and revealed thirteen mobility segments. It thus examined the quantity of heterogeneity in mobility behavior of Swiss households. The present chapter signals the following chapter that there are heterogeneous groups in the population, so models with a single representative agent are not sufficient to model their behavior. The following chapter considers the information to choose the modeling approach as well as the degree of heterogeneity to be modeled in the model.

3.7 Limitations

The chapter has certain limitations. The first one is the boundaries of the interval for the number of clusters values. In the first step of the procedure, all the integer values between [2-15] as the candidates of the number of clusters. It appointed the value with the highest ASW score, which is 13, as the optimal number of clusters. However, there might be potential values out of the boundaries of the interval that might improve the ASW value even more. Due to the computational time, the first peak was chosen as the optimal. That is a limitation and could be improved in future research. The second limitation is the upper bound of the Gower distance weights. It was set to three to avoid the dominance of a variable in the structure of segments. The lower bound is set to the default Gower distance weights to be consistent. Besides, the optimization function searches for the optimal weights with seven digits. It means it tries all possible weight combinations with seven digits. So, keeping the upper bound of the weights higher would lead to high computational time. That is another

reason for setting the upper bound of the weights at three. There might be a better weight combination out of the boundaries that could improve the ASW score of the optimal number of clusters more. So, setting the upper bound at three is a limitation that could be improved in future work. Besides, in future work, the lower bound of the weights can be set to zero, enabling us to eliminate distinctive variables in the segmentation process. That is, the optimization process can reveal a score of zero for a variable if it is not distinctive. So, it is not considered in the segmentation procedure. Finally, the chapter has chosen six variables to be used in the segmentation procedure. The variables have been chosen according to the implications in previous studies. They could be chosen with a data-driven approach in future research, such as factor analysis. Furthermore, the number of variables is limited to six due to computational complexity. Since the units to be segmented are large (approximately 29.000 households), creating an $N \times N$ dissimilarity matrix is a memory-demanding process. Therefore, we kept the number of variables in the final dataset six, which could be improved by adding new variables in future research.

3.8 Conclusion

The overall chapter aims to explore heterogeneity among Swiss households in terms of their mobility behavior and characteristics. It introduces a structured segmentation procedure and implements it with the Swiss data. That reveals the Swiss mobility consumer segments that quantify the heterogeneity in mobility behavior of the Swiss households. In methodological terms, the introduced

procedure measures the quality of segments and improves the quality by an optimization process. Finally, it presents internally compact (homogeneous) segments that are well-separated from each other. The segments can be used to formulate tailored segment-level mobility behavior change policies. They can also be used in economic models that aim to capture heterogeneity at the group level. Instead of a single representative agent for the whole population, a representative agent can be created for each segment (i.e., segment-level representative agents). That has the potential to improve the bias of assuming the whole population identical in economic analysis.

3.9 Appendices

Appendix A

Algorithm 1 Function A to find the ASW value of the clustering given Gower distance weights

Input: weights

Output: ASW value

- 1: Calculate a Gower distance matrix with the weights
 $Gower_dist \leftarrow daisy(data, weights = weights)$
 - 2: Partition the cases into k clusters
 $clusters \leftarrow pam(Gower_dist, k)$
 - 3: Calculate the ASW value of the cases in the k clusters
 $ASW\ value \leftarrow clusters\$silinfo\$avg.width$
 - 4: **return** ASW value
-

Algorithm 2 Function B to optimize the default weights

Input: default weights, upper bound(u), lower bound(l)

Output: optimized weights

1: Calculate the optimized weights through the optim function

$optimized_weights \leftarrow optim(par = weights, fn = function A, lower = 1, upper = u, method = 'L - BFGS - B')$

2: **return** $optimized_weights$

Appendix B

In this section, distributions of the variables for each cluster are illustrated in the following charts. For the multimodality variable, 1 and 0 indicate yes and no, respectively. In the has-HT variable, 1 and 2 indicate owners and non-owners, respectively. In the Modal-choice variable; 9=car, 2=train, 6=bus, 15=walking, 5=tram, 14=bike.

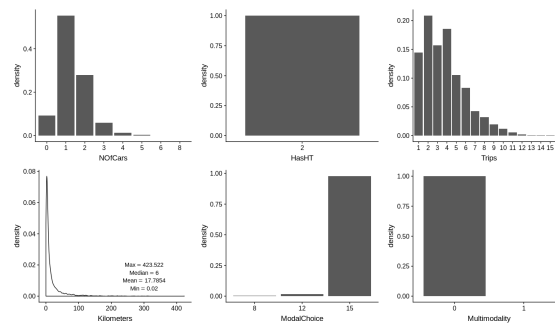


FIGURE 3.3: Segment 1

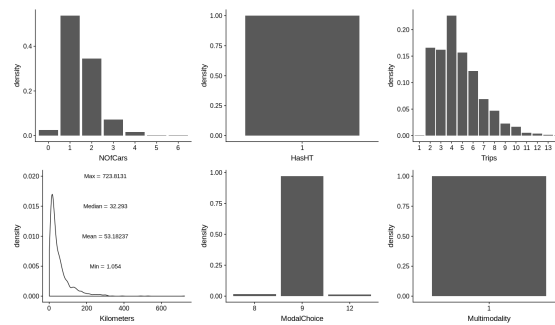


FIGURE 3.4: Segment 2

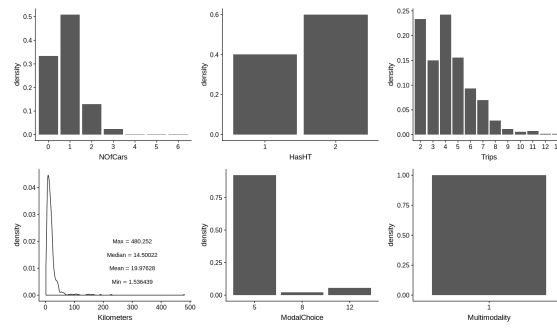


FIGURE 3.5: Segment 3

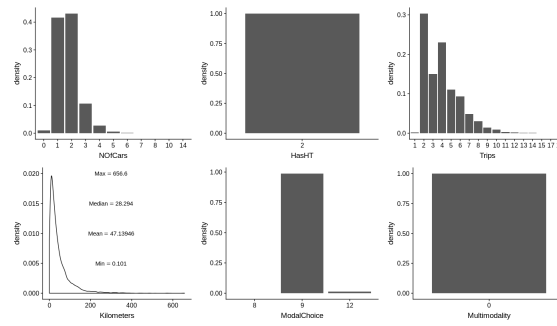


FIGURE 3.6: Segment 4

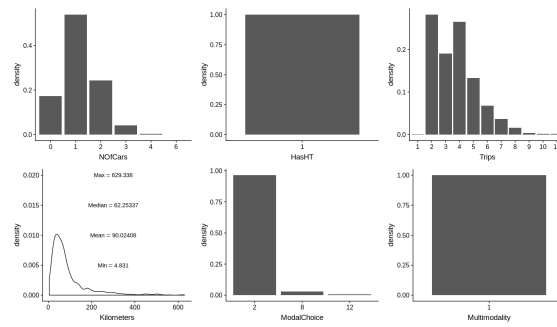


FIGURE 3.7: Segment 5

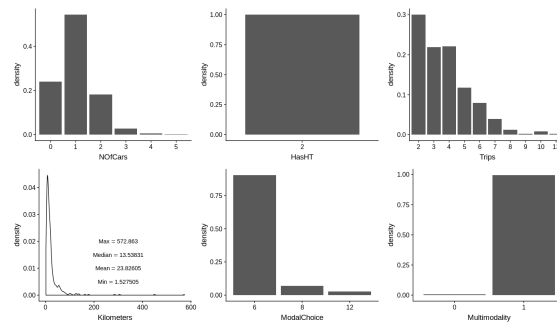


FIGURE 3.8: Segment 6

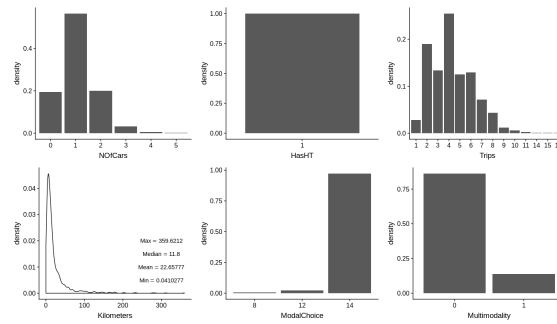


FIGURE 3.9: Segment 7

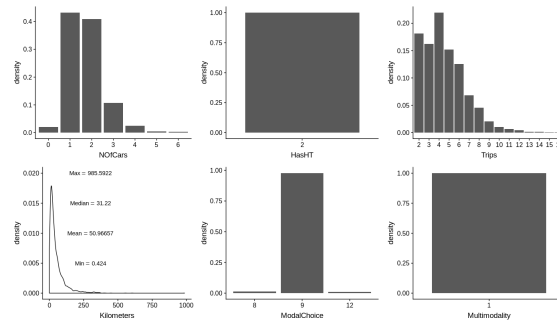


FIGURE 3.10: Segment 8

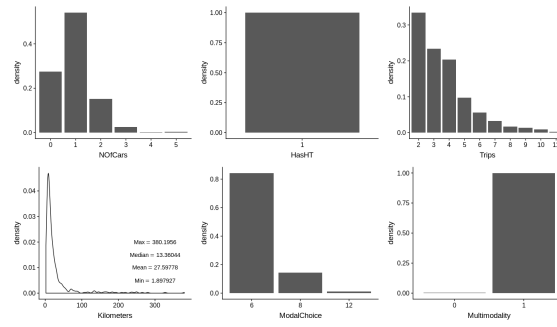


FIGURE 3.11: Segment 9

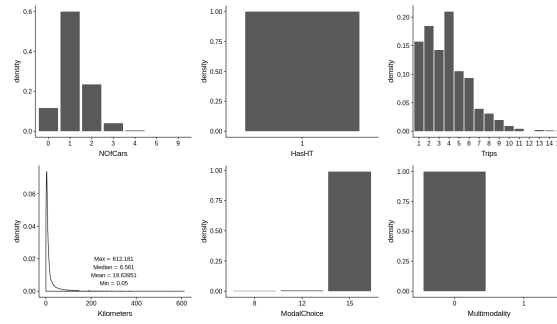


FIGURE 3.12: Segment 10

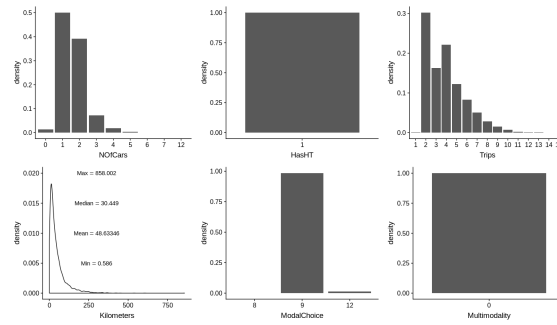


FIGURE 3.13: Segment 11

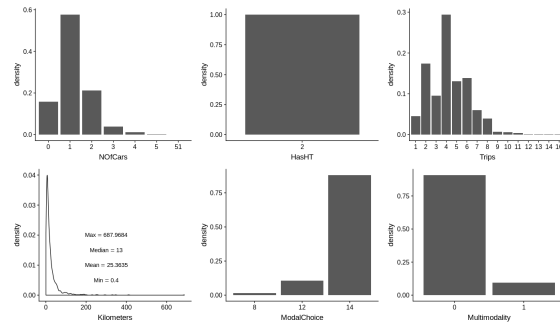


FIGURE 3.14: Segment 12

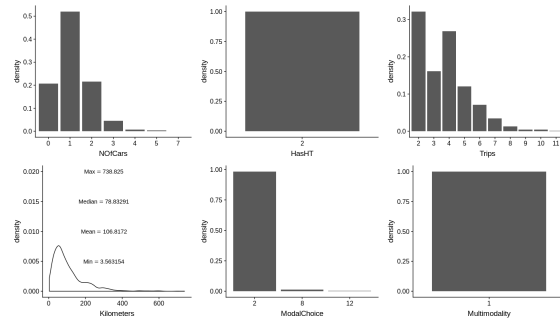


FIGURE 3.15: Segment 13

Chapter 4

Building an ACE demand model and validating it with a new meso-level empirical validation method¹

Abstract

The previous two chapters empirically analyzed and examined the heterogeneity among Swiss households regarding their mobility consumption behavior and characteristics. This chapter employs the findings of the first two chapters as a foundation. It builds and validates an agent-based computational economics (ACE) demand model for modeling mobility demand in Switzerland. The model consists of agents that are behaviorally and structurally heterogeneous. Agents are behaviorally heterogeneous because they have heterogeneous utility functions. Agents' utility functions are configured according to their behavioral segment information. So, agents belonging to different behavioral segments have different utility functions in the model. That enables the model to capture the behavioral heterogeneity among agents that we observed in the past

¹A scientific paper based on this chapter has been peer-reviewed and published by the SN Business & Economics Journal in the special issue on New Frontiers in Economics: the Agent-Based Approach (Bektas, Piana, and Schumann, [2021](#))

chapters. Besides, agents' utility functions are time-inseparable; thus, their past consumption partly influences the current utility. Agents in the model are structurally heterogeneous because they can have different endowments, constraints, socio-economic characteristics, and demographics. Thus, the model can capture both structural and behavioral heterogeneity among agents. As a consequence of that heterogeneity, agents in the model can react to the changes in their economic environment (e.g., economic policies) differently as empirical studies observed in the real world. That will be discussed and examined in the following chapter in more detail. The model produces the individual level mobility demands that can be aggregated to find macro patterns such as total mobility demand and CO_2 emissions. The model can be used as an artificial world to test decarbonization policies in silico, as will be demonstrated in the following chapter. More discussion about the model is given in the case-study and model building sections.

An important line of research in ACE is empirical validation (Tsfatsion, 2021). As mentioned in many influential contributions in the ACE literature, ACE models lack empirical validation methods that undermine the credibility of modeling results (Fagiolo et al., 2019; Barde, 2020). Therefore, this chapter first discusses the empirical validation problem of ACE models in depth and proposes a replicable method to empirically validate the output of ACE models, a specific indicator of "goodness-of-validation" and its statistical distribution, leading to a statistical test in some way comparable to the p-value. The method involves an unsupervised machine learning algorithm hinging on cluster analysis. It clusters the ex-post behavior of real and artificial individuals to create

meso-level² behavioral patterns. By comparing the balanced composition of real and artificial agents among clusters, it produces a validation score in $[0,1]$ which can be judged thanks to its statistical distribution. In synthesis, it is argued that an ACE model can be initialized at the micro-level, calibrated at the macro-level, and validated at the meso-level with the same data set.

After introducing the proposed empirical validation method, the chapter discusses the model building in the case-study section. It introduces the ACE model for modeling mobility demand in Switzerland, as mentioned above. It discusses the properties and novelties of the model, and its contribution to the ACE literature in detail. Finally, the proposed empirical validation method is applied to the built ACE model to assess its output to examine whether the model produces some macro or micro stylized facts observed in real data. We aimed to validate the built model because if the model is validated, the answers derived from the model can be utilized to answer questions directed to the original system (Klügl, 2008). In the end, a validated model is presented and suggested for decarbonization policy assessment. It will be used to evaluate future mobility scenarios and decarbonization pathways in Switzerland in the following chapter.

Keywords: agent-based computational economics, demand modeling, empirical validation, unsupervised machine learning, cluster analysis

JEL Classification: C52, H31, C63, C38

²The meso-level means clusters (patterns) throughout the thesis. It is situated between the micro-level, which is the individual level, and macro-level, which is the aggregated output of the micro-level

4.1 Introduction

Modeling economies as complex systems has been attracting many scholars (Hamill and Gilbert, 2016). Agent-based computational economics (ACE) models are one of the modeling tools for complex systems, which can provide a practical way to model economies or specific economic sectors; thus, their usage has been growing in the field of economics (as well as in other disciplines) during the last three decades (Fagiolo et al., 2019; Hamill and Gilbert, 2016). ACE models consist of autonomous and decentralized entities (agents); each can have dynamic behavior and heterogeneous characteristics (Geanakoplos et al., 2012). The dynamic behavior of heterogeneous agents is governed by decision-making mechanisms (rules) derived from established empirical and theoretical foundations (Dawid et al., 2014). Thus, agents do not necessarily make decisions based on the assumption of a representative agent who is intertemporally optimizing an objective function under rational expectations (Colander et al., 2008).

ACE models have certain features that distinguish them from top-down models, such as DSGE models. (Arthur, 2006). First of all, ACE models have a bottom-up perspective. The macro-dynamics in these models are the emergent properties of micro-level interactions and agents' behavior, which is not necessarily constrained with equilibrium and hyper-rationality (Heckbert, Baynes, and Reeson, 2010). These emergent properties at the macro-level can be used to analyze complex and decentralized systems quantitatively (Duffy, 2006). As Arthur (2006) states, emerging properties often feedback micro-level decisions, which leads to a perpetual novelty in the behavior. Thanks to the bottom-up perspective, ACE models are capable of modeling each individual's micro behavior separately, which allows us to have a potentially high level of heterogeneity

(Dawid et al., 2012). That is, ACE models offer a way to handle higher degree of heterogeneity when models with a single representative agent are not suitable. Secondly, agents in ACE models are not necessarily to have perfect information and complete computing (thinking) capacity. That allows ACE models to capture the bounded-rationality concept that arises due to limited information and computing capacity (Caiani et al., 2016; Fagiolo and Roventini, 2012). Thirdly, ACE models can contain non-trivial interactions, which were governed by ex-ante defined rules of behavior. The interactions can lead to learning and adaptation, which are difficult to capture with top-down models (Farmer and Foley, 2009).

It is an asset for ACE models (like other economic models) demonstrating how well the model Data Generating Process (mDGP) represents the real-world Data Generation Process (rwDGP) (Fagiolo, Moneta, and Windrum, 2007; Klügl, 2008; Bianchi et al., 2007; Murray-Smith, 2015; Beisbart and Saam, 2019). One way to do that is to compare the data generated by the mDGP and the rwDGP statistically; we call this procedure empirical validation (Windrum, Fagiolo, and Moneta, 2007). ACE models are more complex than representative agent models as well as top-down models with heterogeneous agents such as DSGE models due to the potential existence (by no means necessary) of non-linearities, macro-micro feedback, heterogeneous interactions (Fagiolo, Moneta, and Windrum, 2007). That makes the relationship and the comparison of ACE model-generated data and real data problematic, which leads to complexity and consequently, methodical problems regarding the empirical validation of ACE models (Heckbert, Baynes, and Reeson, 2010). Although there are contributions in the last decade, such as (Barde, 2020; Lamperti, 2018b; Guerini and Moneta, 2017), we still do not have standardized empirical validation methods for ACE models

that inevitably lead to a lack of robustness in terms of validation (Fagiolo et al., 2019).

Previous research recommends the involvement of machine learning techniques detecting and comparing patterns in data as for empirical validation methods (Fagiolo et al., 2019; Barde and Van Der Hoog, 2017), which allows us to perform more thorough comparisons of mDGP generated data and rwDGP generated data. A class of these algorithms, which will be discussed in the following section, can help us to reveal stylized facts in the data that the rwDGP generates and assess the mDGP's ability to produce them. The recommendation of the previous research has motivated the present chapter; it proposes an unsupervised machine-learning algorithm³, specifically *cluster analysis* (Russell and Norvig, 2002), as an empirical validation method. The goal of cluster analysis methods and their principles were discussed in the previous chapter in detail. The proposed empirical validation method focuses on the ACE models that use micro-data as input for configuration and produce results accordingly to address questions from the real world. That is, micro-level is used in these models for configuration (e.g., the specification of agents' and environment's characteristics). Aggregation of micro-outputs yields macro-patterns, which are often used for calibration. Therefore, the proposed validation method aims to compare model-generated data and real data at the meso-level. The rationale to target on the meso-level is the following. First, validation at the micro-level can lead to overfitting since it is the level of model configuration. Second, validation at the macro level is not sufficient because ACE models can produce outputs more granular than macro-level (Fagiolo et al., 2019). Validation at the meso

³Unsupervised machine learning detects previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision. Cluster analysis is considered as one of the most common unsupervised machine learning algorithms (Kassambara, 2017).

level can avoid both the overfitting issue of the micro-level and insufficiency of the macro-level. Thus, the method can improve the existing practice in the validation of ACE models. The proposed method clusters the ex-post behavior of real individuals and artificial agents, who have the same ex-ante given characteristics. Then, it quantitatively assesses how well the clusters overlap in a multidimensional latent space. It then produces a qualitative validation score and its statistical distribution for the assessment.

Once the new empirical validation method is introduced and proposed, in the case-study section an ACE model is built to model heterogeneous mobility demand in Switzerland. The model incorporates the empirical findings of the previous chapters, such as the existence of heterogeneity. That is, the model captures both behavioral and structural heterogeneity among agents by its design. The agents in the model are heterogeneous in many dimensions such as utility functions, endowments, constraints, and socio-economic characteristics. These differences are relevant and known to agents while making a particular economic decision. Therefore, agents' responses (i.e., decisions) towards changes in their economic environment, including those caused by policies, can be heterogeneous. The model produces individual mobility demands as output. These can be aggregated to find macro-patterns such as total mobility demand in Switzerland or total domestic CO₂ emissions produced by the mobility sector. The model and its features and novelties are discussed in detail in the model building section. The proposed empirical validation method is applied to the built ACE model to assess its output in the case-study section. It is aimed to analyze whether similar empirical regularities occur in model-generated and real data. Finally, the chapter presents an empirically validated ACE model to test

climate policies to decarbonize the Swiss mobility sector. The model can demonstrate demand shifts towards more sustainable mobility modes and consequent CO₂ reductions due to policy implementation. The following chapter presents a study with the usage of the model to analyze different decarbonization policies targeting mobility behavior change in Switzerland.

The rest of the chapter is organized as follows. In Section 4.3, we first discuss the theoretical background of the validation of ACE models in light of existing literature. Then, we touch on the recently introduced validation approaches. After that, we explain the proposed method and discuss how it could expand the existing literature. Section 4.4 builds the ACE model and discusses its properties and novelties in detail. Then the proposed validation method is applied to the output of the built model. The validation results are discussed in Section 4.5. Then, the value of the method and its applicability to other ACE models in different sub-domains of the study of economics such as macroeconomics, decentralized markets, industrial organization is discussed in the Section 4.6. It also discusses what kind of ACE models could be assessed by the method and provides some example models for the sake of clarity. Finally, the chapter ends with the future works and conclusions section.

4.2 Validation Methods

4.2.1 Theoretical Background of Validation of ACE models

In this section, we follow a general-to-specific way to discuss the validation of ACE models in the light of existing literature. Firstly, we introduce the types of validation techniques (stages) for ACE models in general terms. We utilize a

procedure to validate ACE models, which was introduced by Klügl, [2008], and discuss the validation stages that are ordered in that procedure. Then, we discuss one of these stages (the last one) called empirical validation in detail, since we introduce a novel method for that stage in this paper.

One of the major valuable aspects of using ACE models is to explain and understand a real-world phenomenon that is costly and sometimes difficult to analyse in real-world (e.g., field experiments, real laboratory experiments, etc.) (Xiang et al., [2005]). As Farmer and Foley ([2009]) state,

"ACE models allow for the creation of a kind of artificial (virtual) universe in which many players act in complex and realistic ways"
(p. 686)

Thus, such models enable to analyze *-in silico-* the future status of the original system under novel conditions. Assessing how well the artificial universe (i.e., ACE models) represents a proportion of the original system (i.e., a part of the real world that is aimed to be modeled) is an asset for models that potentially makes the modeling results more credible (Klügl, [2008]). This assessment is called validation in the literature (Windrum, Fagiolo, and Moneta, [2007]; Bianchi et al., [2007]). If the model is validated, the answers derived from the model can be utilized to answer questions directed to the original system (Klügl, [2008]).

Klügl ([2008]) introduced a procedure (see Figure [4.1]) that involves different stages to explain how to get a model from implementation to usage. These stages are also discussed in (Balci, [1994]). According to the procedure, once a model is implemented, the face validation stage is at first carried out. In that stage, the modelers are supposed to contact domain experts to assess whether the model behaves reasonably. The experts provide subjective judgments on the accuracy

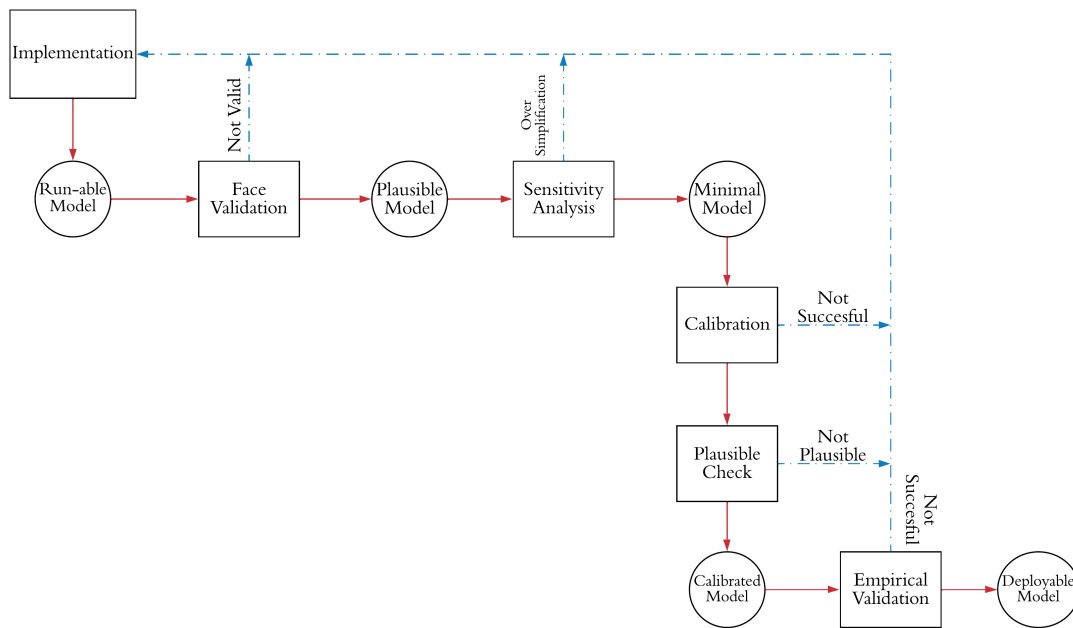


FIGURE 4.1: A general procedure to validate ACE models (KlÜgl, 2008)

of the model. Sensitivity analysis comes next, where the impacts of different parameters on the model output are assessed. It is assumed that the relationship between a parameter⁴ and the output occurring in the model should occur similarly in the original system as well. Once such impacts are analyzed, then the appropriate values are assigned in calibration⁵ for the parameters. Calibration aims for finding the -optimal- parameter set, which resembles the output of the model to the output from the original system. In general, ACE model parameters are calibrated in regard to aggregated (macro) patterns that the mDGP produces (Guerini and Moneta, 2017). That is, the macro patterns that the model produces are compared with the real data, and the values that make them the most similar are assigned to the parameters. The plausibility check comes after

⁴A model parameter is a configuration variable whose value can be estimated or learned from data (Heckman, 2000).

⁵Calibration is a process for finding values for the parameters of artificial economic worlds (i.e., economic models). (Gregory and Smith, 1990)

calibration, where human experts assess the plausibility of the model outcome (e.g., dynamics and trends of the different output values of model runs). It is technically the same as the previously discussed face validity, as Klügl (2008) states. Finally, statistical tests are applied to compare model-generated data and real data as named empirical validation. That is, the model becomes deployable and be ready to use after its output is validated.

Empirical validation is the last stage of the procedure in Figure 4.1 and aims to compare the data coming from the rwDGP and the mDGP statistically. For instance, assume that we have real data generated by the rwDGP, which contains different data points. The data points can be at the micro-level as Equation 4.1 denotes (Pyka and Fagiolo, 2007; Windrum, Fagiolo, and Moneta, 2007), where I represents the population of individuals whose heterogeneous behaviors are observed and contained in the vector of z in a finite time steps of n . For instance, for a mobility demand model, z would be individual level mobility consumption behavior. In the equations, t stands for time.

$$(z)_i = \{z_{i,t}, t = t_0, \dots, t_n\} \quad i \in I, n \in \mathbb{N} \quad (4.1)$$

$$(Z) = \{Z_t, t = t_0, \dots, t_n\} \quad n \in \mathbb{N} \quad (4.2)$$

The data points that the rwDGP generates at the micro-level can be aggregated to obtain macro data points, as denoted in Equation 4.2 (Pyka and Fagiolo, 2007; Windrum, Fagiolo, and Moneta, 2007), where the vector of Z contains aggregated micro data points of a population (i.e., I). For instance, a household's consumption behavior is represented by a micro-level data point, while

the aggregation of all households in a population I is represented by a macro-level data point. Modelers aim to approximate values for the vector of z or Z for which finding the optimal micro (θ , e.g., agent preferences) and macro (Θ , e.g., the environment) parameters is needed for calibration. Once the optimal parameters are set, which is the one step before the empirical validation in Figure 4.1, then the output of the model can be compared empirically to real data from the original system (Fagiolo, Moneta, and Windrum, 2007; Guerini and Moneta, 2017). As Klügl (2008) states,

"calibration and validation must use different data sets for ensuring that the model is not merely tuned to reproduce given data, but may also be valid for inputs that it was not given to before" (p. 6).

However, having two data sets from the same original system is often not possible. In such cases, the available data can be used on all available levels, as Klügl (2008) asserts. For instance, a model can use micro-data as input, be calibrated at the macro-level, and be validated at the meso-level. So, the same data set can be exploited at different levels without over-fitting.

4.2.2 Related Works Regarding Empirical Validation

In this section, we first discuss recently introduced validation methods. Then, we explain why our method is related to the discussed methods and how it could expand them for the sake of readers.

Barde (2016) and Barde (2017) has introduced a new information theoretic criterion as a validation method for ACE models. The method follows the minimum description length (MDL) principle, which hinges on the efficiency of data compression to measure the accuracy of models' output (Grünwald and

Grunwald, [2007](#)). It first uses model-generated data to create a Markov transition matrix for the model, then uses the real data to produce a log score for the model on the data. As the author clearly states in his recent paper (Barde, [2020](#)), the method can be applied only to ACE models that are reducible to a Markov process. So, the method can assess only memoryless ACE models' validation. Our method differs at this point that it can be applied to non-Markov processes, which are to be expected in many ACE models since "history matters" (e.g., everything that agents learn or store at a local state) in many of them (Axelrod, [1997](#)). For instance, as will be discussed in the next sections, the model that we use for the case study is non-Markovian since agents have individual memories to record and consider their past decisions; thus, it can contain habits transitioning across many periods. As the author states, the method doesn't include estimation; instead, it is applied to already calibrated models to assess their output. It is similar to our method from that aspect.

Lamperti ([2018a](#)) has offered another information theoretic criterion called General Subtracted L-divergence (*GSL-div*) as a validation method for ACE models. The method measures the similarity between model generated and real-world time series. It assesses the extend of models' capability to mimic patterns (e.g., distribution of time series such as changes of values from one point in time to another) occurring in real-world time series. It is related to our method because our method aims to compare the similarity among patterns occurring in real data and model-generated data as well. However, *GSL-div* focuses only on aggregated time-series data as Fagiolo et al. ([2019](#)) indicate, while our method focuses rather on meso-level behavioral patterns that are constructed by micro attributes. We discuss the advantages of the meso-level approach later. The authors state that the *GSL-div* can overcome certain shortcomings of the method of

simulated moments (MSM) (McFadden, 1989), e.g., it does not need to resort to any likelihood function and provides a better representation about the behavior of complex time-series. Their method could be applied technically to any ACE model that produces times series data. Detailed explanation of the method, illustrative examples, and case studies can be found in (Lamperti, 2018b; Lamperti, 2018a).

Grazzini and Richiardi (2015) discuss estimation methods for Dynamic stochastic general equilibrium modeling (DSGE) models and analyze whether such models can also be applied to ACE models. The authors mention the simulated minimum distance (SMD) methods (Hall and Rust, 2003), such as the MSM, as a natural approaches to the estimation of ACE models. Such methods aim for estimating model parameters by minimizing the distance between the aggregates between model output and real data. Our approach differs from these methods because it focuses on the last step of the procedure of Klügl (2008) (see Figure 4.1). In other words, it is applied to already calibrated models, similarly to the method of Barde (2016). Thus, the estimation methods in the class of SMD can be only complementary to our method. As will be discussed in the future works section, in a future paper, we plan to couple a SMD method with our method to apply together on an ACE model.

Differently from the previously discussed methods, Guerini and Moneta (2017) offers a method that aims to compare causal relationships in model-generated data and real-world data to validate ACE models. The method hinges on estimating Structural Vector Autoregressive (SVAR) models through real and artificial time series and comparing them to get a validation score. The method is related to ours since our method also aims to compare causal relationships between ex-ante characteristics and individuals' ex-post behavior. However, we

compare such relationships at meso-level while the method of Guerini and Moneta (2017) focuses only on aggregate time-series.

To conclude, as Fagiolo et al. (2019) state in their critical review,

“all these recently developed validation methods focus only on aggregate time series, while most of ACE models have been able to replicate both micro and macro stylized facts”(p. 14).

Some of the discussed methods could be applied in principle at the micro-level, but there is no “proof-of-concept” yet. Besides, applications of such methods at the micro-level could lead to over-fitting if a model gets micro-data as input and its parameters are estimated to fit individual behavior one-to-one (e.g., fitting behavior of artificial agent to its real counterpart). Considering the increasing availability of micro-data, the number of ACE models using micro-data as input is growing (Macal and North, 2014; Hamill and Gilbert, 2016). Therefore, in this paper, we offer a meso-level validation method for the models drawing on micro-data. The method involves an unsupervised machine-learning algorithm. along the lines suggested by (Fagiolo et al., 2019) and (Barde, 2016). They represent contributions regarding machine-learning involvement on the side of model estimation (Hoog, 2019). However, such involvement is still lacking on the side of validation. Our method could expand the existing validation methods towards the direction of machine-learning and encourage future contributions. The further text is structured as follows: we discuss the overall concept of our method in the next section in detail. We also discuss for what kind of ACE models the method could be applied and provide some example models from recent research in Section 4.6.2.

4.2.3 The Overall Concept of the Meso-level Validation Method

This section introduces a meso-level empirical validation method for ACE models drawing on micro-data first as a broad methodological choice, and then we describe it in detail. In broad terms, we sharply distinguish the different phases and goals of the relationship between real (empirical) data and the model in the following way: the meso-level is exclusively used for validation, whereas the micro-level is used for input micro-data into the agents in terms of parameters (not of outcomes of their decision-making process, because this could lead to over-fitting) and the macro-level for calibration. By this distinction, we radically eliminate any source of overlap between what is given to the model as input, what is used for calibrating its overall results and macro-micro parameters, and what is used for validation. More specifically, our method consists of sequential steps for which we created an overall concept as in Figure 4.2. We explain each step one after another, according to their sequence in the concept. The main goal of the concept is to compare model-generated data and real data at the meso-level to understand how well the mDGP can produce the behavioral patterns that occur in real data. It produces a quantitative score in an interval of [0-1] according to which we can assess the validity.

The overall concept gets two data sets as input. The first data set contains information regarding the ex-ante characteristics of artificial individuals (i.e., agents) and their ex-post behavior, generated by the mDGP. The second data set involves information regarding the ex-ante characteristics⁶ of real individuals and their ex-post behavior, generated by the rwDGP. Both data sets contain information at the individual level since the mDGP of ACE models produce

⁶The variables constituting characteristics of individuals are their socio-economic and demographic attributes that are relevant for decision-making. We will discuss it later in detail.

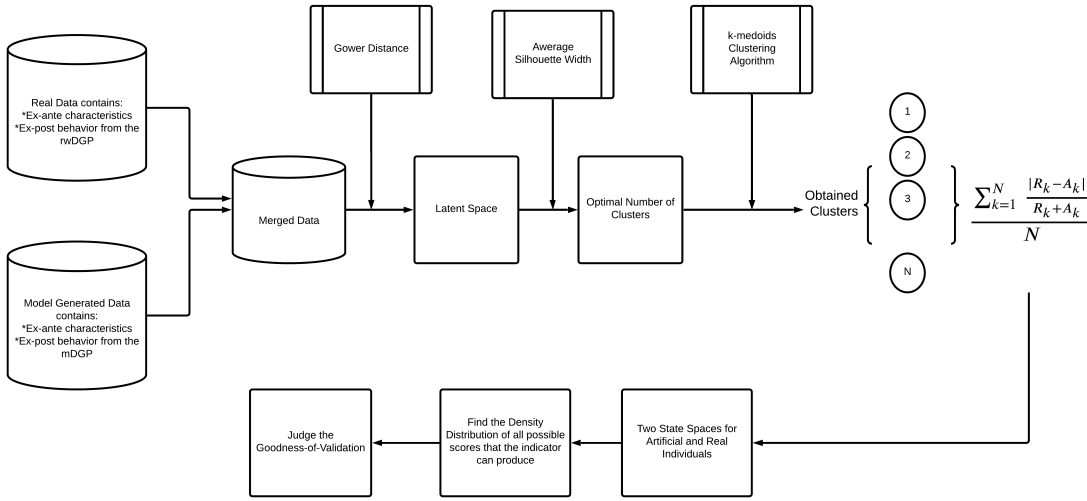


FIGURE 4.2: The Overall Concept of the proposed empirical validation method

data at the individual level (i.e., micro-level). Individuals are clustered according to their characteristics and behavior in the data sets, and these clusters are compared at meso-level quantitatively. An essential point for the comparison according to the method is that the real data should be the one that is used to initialize the model. In this case, individuals in real data are mapped to artificial agents one-to-one; thus, the number of real and artificial individuals become equal. That is a prerequisite to apply the validation method. The data sets can differ in what model-wise is an ex-post behavior because an artificial agent might behave differently than a real individual with the same characteristics. The variables constituting the ex-ante characteristics should ideally be the ones influencing the ex-post behavior. By having this, the clusters involve a combination of the variables in individuals' characteristics and consequent behavior. Hence, by comparing clusters, we can study the behavioral patterns (e.g., the relationship between the characteristics and the behavior) in model-generated data and real data.

Instead of clustering artificial and real data sets separately, we merge them as indicated in Figure 4.2, and cluster them together to analyze the balance in the clusters (i.e., how many real and how many artificial individuals are in each cluster). Individuals in the merged data are placed in a multidimensional latent space based on their attributes (i.e., ex-ante characteristics and ex-post behavior). The latent space is represented by a symmetric distance matrix⁷. Several metrics exist to create that matrix, such as Euclidean, Manhattan, Gower, etc. In the overall concept, we utilize the Gower distance metric since it can handle different column types⁸ (e.g., categorical, continuous, ordinal, etc.) to place instances in the latent space (Gower, 1971). For instance, the merged data might contain some attributes of households that can be categorical such as income level or continuous such as age. Gower distance can determine the positions of the individuals in the latent space based on these columns without any transformation, while other metrics such as Euclidean accepts only continuous ones, as discussed in the previous chapter.

To determine the positions in the latent space, all pairwise distances among individuals are created variable by variable. Distances on categorical variables are calculated according to Equation in (4.3), while it is according to Equation in (4.4) for continuous ones. The equations are explained in the previous chapter in detail. Therefore, we do not describe them again to avoid redundancy. Nevertheless, the equations are illustrated again below for the sake of readers.

⁷We provide to readers an example latent space, which we generated for the experiment in Appendix C.

⁸Following the dominant language conventions in the distance metrics and cluster domains, columns mean attributes throughout the article. Rows are cases (agents) that are clustered.

$$D_{i,j,k} = \begin{cases} 1, & X_{i,k} \neq X_{j,k} \\ 0, & X_{i,k} = X_{j,k} \end{cases} \quad (4.3)$$

$$D_{i,j,k} = \frac{|X_{i,k} - X_{j,k}|}{r_k} \quad (4.4)$$

Based on each variable, pairwise distances are generated and scaled in the interval between zero and one. The calculated n distances (i.e., n is the number of variables) are aggregated and normalized according to Equation 4.5 to obtain a final distance between two instances (Gower, 1971). Variables can have different weights (importance) in the aggregation. Weights are represented by w in the equation, which are identical by default and equal to one. Once all pairwise distances are calculated, they are placed in a symmetric distance matrix. The matrix acts like a latent space whereby a clustering algorithm partitions the instances into a pre-specified number of clusters according to the positions of the instances in the space.

$$D_{i,j} = \frac{\sum_{k=1}^N w_k D_{i,j,k}}{\sum_{k=1}^N w_k} \quad (4.5)$$

$$Sil = \frac{b_i - a_i}{\max\{a_i, b_i\}} \quad (4.6)$$

As for the clustering algorithm, we utilized the k-medoids clustering algorithm since it is compatible with the latent space created by the Gower distance metric, as discussed in the previous chapter. However, k-medoids algorithm requires to specify ex-ante the optimal number of clusters. There are the goodness of fit metrics in the literature (e.g., Average Silhouette Width (ASW), Calinski

and Harabasz Index (CH) and Pearson version of Hubert's Γ (PH) (Campello and Hruschka, 2006)), which can provide quantitative measurement scores regarding the quality of clustering with the different number of clusters. The ASW is one of the most widely used approaches that measures how well an instance is matched with its own cluster (Maulik and Bandyopadhyay, 2002). As a goodness of fit measure, it reflects how well intra-cluster homogeneity and inter-cluster dissimilarity are maximized (Rousseeuw, 1987). As discussed in the previous chapter, the idea for pre-specifying the optimal number of clusters is to try different k-values in an interval and appoint one of them, which has the highest ASW value, as the optimal number of clusters. For each k number, the ASW value of the clusters is calculated according to Equation 4.6, which depicts the Silhouette value of instance i . The feature a_i represents average dissimilarity of i to all other objects in the cluster a (the smaller the value, the better the assignment). Another feature b_i ⁹ reflects the minimum dissimilarity of the instance i to all objects in any other cluster (the closest cluster to i except its own cluster). The Equation 4.6 returns values between -1 and 1. Values close to 1 indicates that instance i is assigned to the proper cluster. Average *Sil* values of all instances (ASW) gives an idea about the quality of the clustering (Rousseeuw, 1987).

After the instances are placed in a latent space, and the optimal number of clusters are found, the k-medoids algorithm (See Algorithm 3 in Appendix D) partitions the instances into k (the optimal number) clusters. To understand how well clusters from real and artificial data overlap, we compare the quantity of artificial and real individuals in each cluster according to the indicator (4.7). In the formulation of the indicator (4.7) R represents the number of real

⁹Both a_i and b_i vary in the interval of [0.1].

instances, A represents the number of artificial instances, and N is the optimal number of clusters. The indicator finds the dissimilarity in the balance of artificial and real instances for each cluster. Finally, it returns a normalized score in a spectrum between zero and one.

$$\frac{\sum_{k=1}^N \frac{|R_k - A_k|}{R_k + A_k}}{N} \quad (4.7)$$

If an artificial agent behaves observationally equivalent to the real individual with whom it has the same characteristic, they are placed in the same position in the latent space; thus, they are supposed to be in the same cluster. If all artificial agents behave observationally equivalent with the real individuals with whom they have the same characteristics, it is expected that the clusters would have fifty percent artificial fifty percent real instances (as in the simple experiment in Appendix B). In this case, the indicator's outcome (4.7) becomes zero, which indicates a perfect match. In other words, a zero score demonstrates that the behavior patterns in real data are perfectly overlapping with the ones from the artificial data. Conversely, if an artificial agent produces different ex-post behavior than his real counterpart, they are placed in different positions in the latent space. Thus, they are supposed to be in different clusters. That leads to unbalanced clusters and, consequently, a weak validation score according to the indicator in (4.7).

The overall concept is completed with the determination of the place of the score in the distribution of all possible scores it could theoretically take, which allows us to interpret it. The idea is similar to the permutation test that is used to assess a specific value through its distribution (Collingridge, 2013). We obtain

all possible scores that the indicator in 4.7 can generate and their frequency in the exhaustive list of all possible cases, which is the state space. The state space contains all possible alternative ways in which a total can be distributed¹⁰. In the case at hand, we study the scores that the indicator (4.7) generates in all possible subdivisions of the total number of artificial agents and of the total number of real individuals in the clusters. Accordingly, we obtain the distribution of possible scores, which allows us to judge the specific score - that a model achieves in the previous steps - concerning all other possible scores.

Overall, this procedure builds on the idea that a validated model should produce "indistinguishable" results from real data. We deliver a method having a quantitative indicator of "goodness-of-validation," taking values from zero to one. The method can be used for the ACE models having micro-data as input and produce results accordingly. We discuss such models and provide examples in Discussion Section. The method provides to these models two advantages: avoiding over-fitting of the micro-level validation and having more detailed validation than macro-level, as recommended in (Fagiolo et al., 2019). In the next section, we apply the method to a specific model in the personal mobility domain, implementing a certain simulation framework.

4.3 Case Study

This section consists of three subsections. In the first, we describe an agent-based simulation framework called Behavior Driven Demand Model (BedDeM) that we configured to build a specific ACE model to model individual mobility demands in Switzerland. The framework building process is discussed in (Nguyen

¹⁰We provided a numerical example in Appendix C regarding how the total can be distributed among clusters. We also provide example state spaces for different clustering configurations.

and Schumann, 2019)¹¹, and a use case is addressed in (Bektas et al., 2018). In the second subsection, we discuss the model building process by configuring the generic framework with empirical (real) data. The proposed validation method is applied to the built model, and the results are discussed in the next section. In the third subsection, we discuss the specific variables that constitute individuals' ex-ante characteristics and ex-post behavior in the built model. These variables are used to place real and artificial individuals in a multidimensional latent space to compare meso-level patterns.

4.3.1 The Simulation Framework

General Features

The BedDeM framework has been developed as a generic tool that can be configured to address specific issues from different research domains regarding household behavior (e.g., household consumption, mobility demand, tourism demand, etc.). It comprises of the key theoretical tenets of the multi-agent cognitive system, in which heterogeneous agents are capable of making choices (decisions). It enables modeling the micro-behavior of each individual (agent) separately.

The core element of the BedDeM framework is an agent-based simulator, written in Java based on the RePast library (Nguyen and Schumann, 2019), complemented by key concepts from Triandis' Theory of Interpersonal Behavior (TIB) (Triandis, 1979), described in Section 4.3.1. TIB explains the origin of

¹¹The source code of the framework is available online at (BedDeM, 2020)

individual behavior, which is utilized as a decision-making mechanism (component) in the framework. Hence, agents make their choices (decisions) according to the contained determinants in the TIB.

Overview of Agent's Design

BedDeM consists of agents that have (not necessarily) heterogeneous characteristics and preferences. Agents are assigned tasks and are supposed to choose an option to perform their tasks according to the ex-ante defined behavioral rules (e.g., decision-making mechanism). Tasks and options are specified according to the application domain. For instance, according to a configuration, agents might choose a tourist place to visit or choose a mobility mode to perform their trips. Configuration here stands for applying the framework for a specific domain such as mobility, tourism, household consumption, etc.

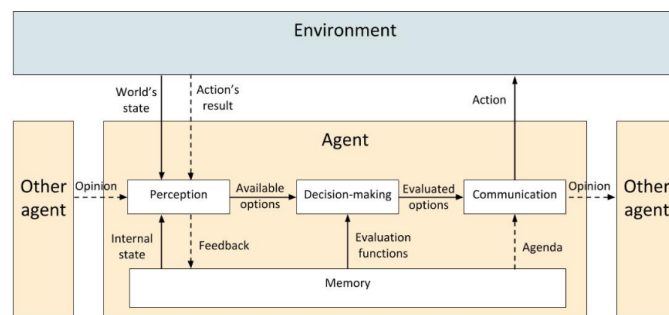


FIGURE 4.3: Overview of Agent's Design (Nguyen and Schumann, 2019)

When an agent performs a task, he first collects information. The perception module (see Figure 4.3) gets information about the present state of the environment and combine it with other agents' opinions. It then brings the information to the decision-making module for reasoning. The obtained information is combined with heterogeneous preferences and also with the past decisions in

the memory. As agents maintain their local state (the individual level memory, see Figure 4.3), the decision-making becomes time-inseparable. In the end, the agent lists all available options to perform the task and choose the most preferable one according to his individual reasoning. After the choice, he informs the other agents about his choice that can be used as information in others' decision-making modules.

Agent's Decision Making Mechanism

Triandis' Theory of Interpersonal Behavior (TIB) Since micro-behavior is the main output of the mDGP, it should be well-specified to obtain precise emergent properties with the original system. While the BedDeM framework was being constructed, firstly, the origin of individual behavior was addressed. The idea was obtaining a standard theory that depicts the origin of individual behavior and using this theory as agents' decision-making mechanism. In cognitive science, there exist such theories e.g., Ajzen's Theory of Planned Behavior (TBP) (Ajzen et al., 1991) and Ajzen and Fischbein's Theory of Reasoned Action (TRA) (Chang, 1998). These theories state that individual's intention to act is the key determinant of behavior (Bamberg, Ajzen, and Schmidt, 2003). There are several ACE models and platforms they attempted to incorporate these theories (Nguyen and Schumann, 2019).

Triandis extended these theories in his TIB model (see Figure 4.4). He added two new components over them; habits and facilitating conditions. According to TIB, the frequency of past behavior forms a habit that partly impacts current behavior. Hence, the current behavior is determined by the current status of the environment (e.g., economic parameters) and the previous decisions in the individual memory. The theory as well as other empirical research state that

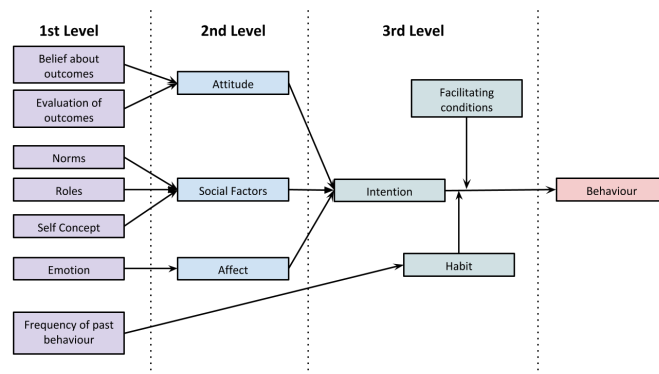


FIGURE 4.4: Triandis TIB model (Triandis, 1979)

intention is moderated by habit that leads to non-deliberate decision-making (Verplanken et al., 1994; Bamberg, Ajzen, and Schmidt, 2003).

As Nguyen and Schumann, 2019 state, TIB includes all aspects of TRA and TPB, as well as additional components such as habits that potentially improve its predictive power and descriptive accuracy. Although there is no proof which theory is more suited to build an agent-based framework, TIB was chosen for the BedDeM framework since it provides a more comprehensive understanding of the origins of individual behavior.

Implementation of TIB as agent decision-making mechanism The full implementation of the TIB model as an agent decision-making mechanism is illustrated in Figure 4.5. The function in Eq. 4.8, which will be called the agents' utility function throughout the chapter, is used by agents to execute the decision-making (see Fig. 4.5) and calculate the expected utility of each option according to the determinants in the TIB model. That is, each determinant in the TIB model is a part of the agents' utility function, so agents consider them while calculating expected utilities. When a task is assigned to an agent, he first gets information

from the environment to have available options to perform the task. Then, according to each determinant (d) (i.e., box in Figure 4.5) in the first layer, the agents sorts available options (opt) in a list according to their expected utility (see Eq. 4.8). The expected utility is calculated by comparing the property of an option with other's ($EU_d(opt)$). To calculate the expected utility in the first level, either a real numerical system (for quantitative determinants such as price), or a ranking function (for the determinants such as emotions) are utilized. Both numerical values and rankings can come from empirical data or be calibrated through experts' assessment (Nguyen and Schumann, 2019).

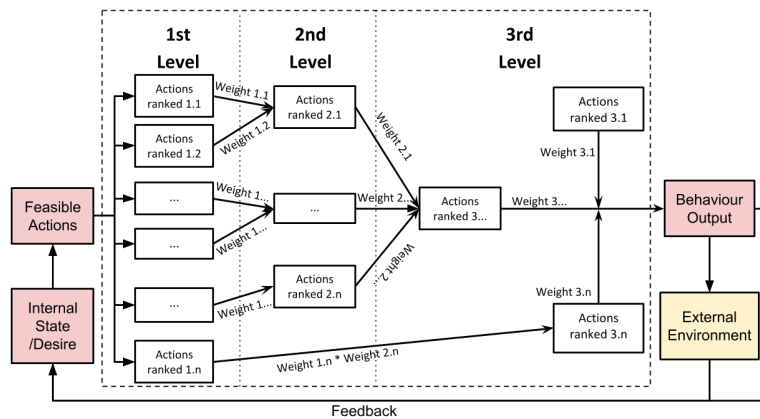


FIGURE 4.5: Agent's decision-making mechanism with mapping to TIB (Bektas et al., 2018)

$$EU_d(opt) = \sum_{c=1}^C (EU_c(opt) * w_c / (\sum_{o=1}^O R_c(o)))$$

- $EU_d(opt)$ is the score of an option (opt) at determinant d.
- C is the set of the children of d (i.e., determinants connects with d in the previous level). (4.8)
- O is the set of all available options.
- w_c is the weight of child determinant c.

Once all options are ranked in lists according to each determinant in the first layer, the lists are merged and normalized with associated weights (w_d) to pass in the next layer (see Eq. (4.8)). The expected utility of each option according to each determinant is multiplied with the associated weight, which becomes the new expected utility of the option. The weights in decision-making represent the importance of determinant. For instance, if it is desired to have time-separable decision-making, the weight of habit can set to zero, which means that the memory (i.e., past decisions) does not impact the current behavior. The weights are determined based on the segments obtained in the previous chapter. That is, agents belonging the same segment have similar weights; thus similar preferences and behavior. By this means, the model captures the behavioral heterogeneity among agents observed in the previous chapter. Detailed information about the use of behavioral segments to calibrate the model (i.e., assign the weights) can be found in (Nguyen and Schumann, 2019). Once all decision-making steps are merged, the agent ends up with a sorted list of options according to their expected utilities. According to the configuration, he can choose the first best option deterministically in the list, or with certain probabilities can

be created over the scores; thus, he can choose an option stochastically. More detailed information regarding the framework and its decision-making mechanism can be found in (Nguyen and Schumann, 2019).

4.3.2 Model Building

We are currently applying the framework in the mobility domain. The BedDeM framework becomes an ACE mobility demand model through the configuration, which aims to generate heterogeneous mobility demands at the household level. The model allows for mode choices for mobility trips based on price and non-price signals through its decision-making mechanism. The model incorporates the insights from the first two chapters. It captures both structural and behavioral heterogeneity among agents. Structural heterogeneity stands for dissimilarities in socio-economic and demographic attributes. That is, agents in the model can have different income levels, resources (e.g., private cars, travel cards, etc.), place of living, household size, and so on. Behavioral heterogeneity stands for the differences in agents' behavior. That is, agents in the model can have different mobility preferences, pro-environmental values, emotions towards certain modes, and habit stocks. Agents can have heterogeneous utility functions thanks to the heterogeneous weights in the decision making discussed in the previous section. This chapter uses the information obtained in the previous chapters. As a consequence of behavioral and structural heterogeneity that the model can capture, agents in the model can respond to economic stimuli heterogeneously, as we observe in the real world. The model allows for ex-ante evaluation of decarbonization policies targeting individual behavior change. The next chapter presents a study regarding the usage of the model to evaluate different

decarbonization policies to change the mobility behavior of Swiss households.

Input of the model We utilize the "Mobility and Transport Micro-census (MTMC)" (ARE/BfS, 2017) data of the Swiss statistical office to build the model. The data is at the micro-level and can be easily mapped to the agents. The data contains information regarding Swiss households' socio-economic characteristics (e.g., location, income level, car/travelcard ownership, etc.) and daily mobility activities. We map the real respondents one-to-one our agents; thus, each agent represents a real Swiss household by having all his characteristics, including mobility activities. Besides, all the exogenous variables that are used to shape the environment, such as fuel prices, reflect the Swiss system. In addition, the Swiss mobility segments that were obtained in the previous chapter are used in the model to calibrate agents' behavior. They are used to assign values for the weights in the decision-making mechanism as discussed in the previous section. That is, agents behavior is calibrated according to the behavioral heterogeneity that we observed in the real data. So, the model involves behavioral heterogeneity in addition to structural heterogeneity among agents.

Output of the model Each agent in the model is assigned a task list (i.e., trips of the real households) to perform. Agents evaluate existing options (e.g., car, public transportation, soft mobility, etc.) according to the decision-making mechanism introduced in the previous section and choose a mode for each of their trips sequentially. The model simulates each agent's micro-behavior separately and generates micro-level heterogeneous mobility demands as the core output. The output can be aggregated to obtain macro-patterns such as total mobility demand or aggregated CO₂ emissions per mobility mode. The model

is initialized through the MTMC data, as mentioned above. The MTMC data contains upscale factors for each respondent to find the aggregates for the whole population. These factors are used ex-post to find macro-patterns over the micro outputs (e.g., individual mobility demands) that the model produces as the core output. That is, individual consumption (i.e., outputs) are multiplied by these upscale factors to get macro-outputs.

4.3.3 Variable Selection for the Case Study

To apply the validation method, the variables identifying the ex-post behavior and the ex-ante characteristics of individuals should be chosen and given to the validation method as input (see Figure 4.2). As ex-post behavior, mode-choice should be in the chosen variables (the last variable in the list below). As for the variables constituting ex-ante characteristics, we identified the ones influencing mode-choice in the previous chapter and used them in this case study (the first four variables in the list below). The complete list of the used variables is demonstrated below:

- Number of cars in the household
- Number of daily trips
- Having a half-fare travelcard
- Daily distance
- Mode-choice

4.4 Empirical Validation Results

We apply the overall concept step-by-step and discuss the results of each step sequentially. We commenced the overall concept with a merged data set containing 3000 artificial and 3000 real (MTMC) individuals with the chosen variables. Before clustering the individuals, we obtained the optimal number of clusters. We utilized the Average Silhouette Width (ASW) score that gives statistics to determine the optimal number of clusters. As illustrated in Figure 4.6, we clustered individuals in the final data set into a different number of clusters within an interval ([2:15]) for each, we calculated the ASW score. The results show that we get the highest cluster quality when we cluster the individuals into six clusters. In other words, we obtain the optimal intra-cluster homogeneity and inter-cluster heterogeneity by dividing individuals into six clusters. We utilized the score as a statistical ground and used the obtained optimal number of clusters to proceed the method.

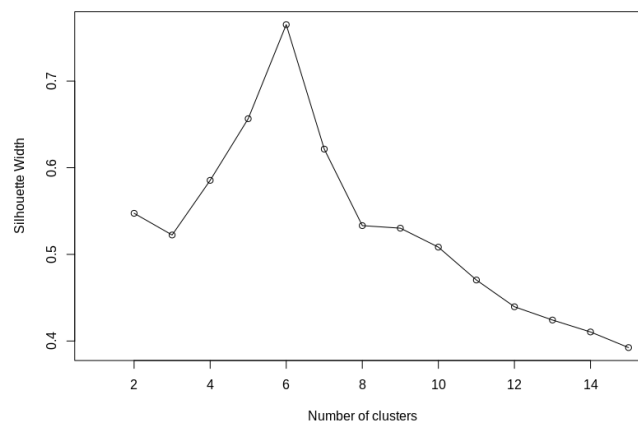


FIGURE 4.6: ASW values with different cluster numbers

We divided the individuals placed on the multidimensional latent space into six clusters and analyzed such clusters' composition. For each cluster, we got

the quantities of artificial and real individuals to compute the indicator. The obtained quantities are demonstrated in Table 4.1.

Clusters	Number of Artificial Agents	Number of Real Individuals	Score
Cluster 1	944	586	0.2339
Cluster 2	1329	990	0.1461
Cluster 3	203	439	0.3676
Cluster 4	175	382	0.3716
Cluster 5	164	322	0.3251
Cluster 6	185	281	0.2060
Overall	3000	3000	0.2750 (Average)

TABLE 4.1: Quantity of Artificial Agents and Real Instances in the clusters and corresponding scores

After we obtained the quantity of artificial and real instances in the clusters, we applied the indicator (4.7) as the overall concept indicated and got the value 0.2750. By construction, it is between zero and one; the lower, the stronger the validation. But how to judge this specific value in general (e.g., independently on the number of clusters)? As anticipated in the previous section, we iterate the computation of this indicator (4.7) for all possible cases (i.e., balance combinations), which is the matrix product of two identical state spaces. An example of such a case is the situation where the 3000 real agents are all in one cluster. That can be matched by the situation in which all 3000 artificial agents are in the same cluster (good) or in another cluster (bad). Alternatively, 2750 artificial agents are in that cluster or in another. Examples like this are many thousands, but Page, 2012 provides a computational method to elicit all of them. It computes not only how many but also enlist which ones they are. Mathematically speaking, it generates the weak composition of 3000 in 6. Since the full number is way too high to be computed in a reasonable time, we first quantize and then fit the results with a continuous function. We quantize the 3000 in 20 groups of 150 units each (in a procedure that is similar to bootstrap). We perform Page's

algorithm that would call `shapes(20, 6)`: a state space enlisting the ways in which 20 units (in our case groups of agents) can be separated into six classes (in our case: clusters).

For each of its rows, the outcome of the indicator can be computed. By having 20 groups (i.e., 150 quantum size of 3000 individuals) and 6 clusters, we obtained two states (for artificial and real individuals), each with around fifty thousand combinations. For all possible number of artificial and real individual distributions in the clusters, we applied the indicator. Then we obtained the density distribution of all the possible scores (i.e., outcomes of the indicator), which is illustrated in Figure 4.7. Thus, we defined the space to see where our model's validation score is, which enables us to judge the score.

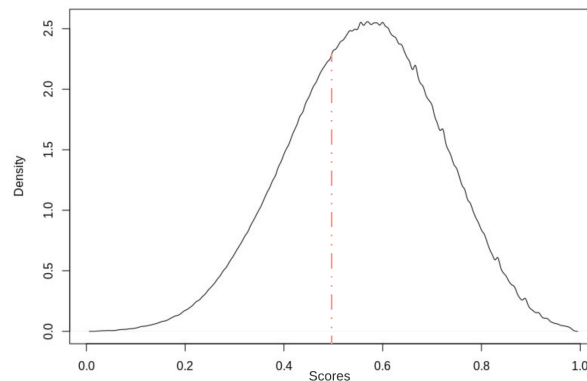


FIGURE 4.7: Density distribution of all possible validation scores

Thanks to the computed all possible scores, we could easily judge the specific score that the model achieves. We summed up the number of cases that have a better score than 0.2750. We divided it into the total number of cases to obtain the percentage. By this means, we calculate the area under the curve (integral) in Figure 4.7. The results show us that approximately 4.2% of cases would produce a score equal to or lower than the score of the model, 0.2750. We interpret these

results in this way that the model is validated at the conventional threshold of 5%.

4.5 Discussion

The findings show that the built model for the case study satisfactorily represents the original system at the meso-level for the given variables. The artificial agents in the model behave observationally similar to the real individuals, who have the same ex-ante given characteristics. It can be interpreted in this way that the mDGP mimics the rwDGP by producing observationally similar data (behavior) with the given input data sufficiently.

To judge the validation score that we obtained, we created the density distribution of all possible scores, as explained in Section 4.3.3. We utilized a simplification by setting the quantum size 150 to reduce high computational time. We tested whether the quantum size is relevant for the density distribution by applying another quantum size (300) (see Figure 4.8 in Appendix A). We compared the functional forms of the density distributions to see whether different quantum sizes lead to different curves. It was observed that both 150 and 300 quantum sizes produce almost identical curves. We report it to demonstrate that the result is robust to the changes in simplified assumptions.

4.5.1 Discussion of the potential application of the meso-level approach to validation upon further models

In this section, we discuss certain insights that we gained during the implementation in the case study that may turn out to be useful to other researchers. The

section provides information about the potential cases (models) for which the method can be applied and guidelines for the sake of usability.

Indeed, this is often the case: ACE models are frequently built and initialized with micro-data since they aim to model heterogeneous behavior of each individual separately and in a highly realistic way (Dawid et al., 2012). Such micro-data contains information at the individual level and can be mapped to artificial agents. Thus, artificial agents get their ex-ante characteristics from real individuals, and they are supposed to generate ex-post behavior according to the behavioral rules of the mDGP. As long as the ex-post behavior of real individuals (or reference group) is known, the proposed empirical validation method can be applied to an ACE model drawing on microdata for which a meso-level can be computed, at which agents can be clustered and compared. For instance, in the model that we built for the case study, agents represent households and it produces mobility mode-choice behavior. Since we also know real households' behavior with the same characteristics, we could apply the method in a rather straightforward way. In another ACE model, agents might represent real firms, and the micro-data containing information regarding real firms can come from accounting systems and declarations to the statistical offices. Real and artificial firms' behavior is clustered with their characteristics as the proposed method suggests and can be compared at the meso-level. Conversely, if in a macroeconomic ACE model there is a wide range of types of agents (firms, households, financial institutions, public institutions, etc.), our procedure might become too cumbersome if applied to all such types. In other words, the method is not dependent on the domain (scope) but its applicability is restricted to ACE models for which a meso-level can be computed from available micro-data, possibly of only one type.

In the procedure of clustering, one needs to select the variables upon which clustering occurs and determine the optimal number of clusters. After that, the application of the indicator in (4.7) can be carried out. The variables should be available for both the artificial and the real agents; they should be relevant for the main behavior that the model is called to describe. In our use case, we used the variables that a previous analysis demonstrated having a large impact on the behavior. But if one cannot proceed with such an analysis, one might take the neutral stance of taking all common variables across real and artificial agents. The optimal number of clusters can be obtained as we did (by taking the number of clusters for which ASW is maximal), but any method that would single out a non-arbitrary number of clusters might be used, if appropriate. Finally, one needs to compute the probability of the goodness-of-validation to be higher than a certain threshold, much alike the p-value. This probability is to be computed using the procedure indicated before¹².

4.6 Conclusion

4.6.1 Summary on model building

The present chapter builds an ACE demand model by configuring the generic agent-based simulation framework BedDeM with the MTMC data. The model produces the individual-level mobility demands that can be aggregated to obtain macro patterns such as total mobility demand and total CO₂ emissions in Switzerland. The model incorporates the findings of the previous chapters. Firstly, it captures structural heterogeneity among agents. Agents in the model

¹²One can download the state space from App. 1 of Piana, Bektas, and Khoa, 2020 or compute it by executing the Java code, downloadable from Appendix 2 of the same paper.

are mapped one to one to the real Swiss households in the MTMC data. Thus, each agent represents a real Swiss household by having all its characteristics. So, agents can be heterogeneous in terms of their income levels, car ownerships, travelcard ownerships, and budget constraints. Besides, agents' behavior is calibrated according to the obtained behavioral Swiss mobility segments in the previous chapter. Agents in the model can have different mobility preferences, emotions towards certain modes, habit stocks, and consequently mobility behavior. Due to the capture of behavioral and structural heterogeneity, agents in the model might react to economic stimuli differently according to the heterogeneity in the decision-making-related dimensions. Secondly, the model incorporates insights from behavioral economics, as suggested by (Stiglitz, 2018). The BedDeM framework involves a standard decision-making theory for mobility consumption behavior. That is, agents' behavioral rules are not defined based on ad-hoc assumptions without a theoretical foundation but are defined based on a standard theory that describes real agents' behavior. In this way, the model avoids the traditional *wilderness* critique of ACE models. Besides, the model is one of the first models to involve a standard theory to describe agents' consumption behavior. So, it extends the body of knowledge on ACE modeling in the sense of incorporating insights from behavioral economics. Finally, the model contains a time-inseparable utility, as suggested by the used decision-making theory. Agents have individual memories, and their past consumption (i.e., mode-choices) relatively impacts the current consumption. Thus, the model allows for habit formation.

To sum up, the chapter builds an ACE demand model that involves both behavioral and structural heterogeneity among agents. The model produces individual mobility demands as the core output that can be aggregated to obtain

the macro results such as total mobility demand and CO₂ emissions. The model has been validated by a new empirical validation method. That is, the model can successfully produce empirical regularities observed in the real data. So, the answers derived by the model can be used to answer questions directed to the Swiss mobility sector. Following the procedure of Klügl (2008) discussed in Section 4.3, the model is ready to use to evaluate decarbonization policies targeting individual mobility behavior change. That is, the validated model can produce potential mobility behavior changes and consequently aggregated demand shifts and CO₂ emission reductions as a consequence of policy implementations. The following chapter provides a study with the usage of the model to analyze future CO₂ emission trajectories of the Swiss mobility sector under the hypothetical implementation of decarbonization policies to change individual mobility behavior, as will be demonstrated in the next chapter.

4.6.2 Summary on model validation

The present chapter presents a new empirical validation method for ACE models drawing on micro-data. The method aims for validating the output of ACE model at the meso level by using the cluster analysis approach. The method intends to cluster the ex-post behavior of real and artificial individuals with the same ex-ante given characteristics. It produces a validation score in [0,1] by comparing the similarity among clusters. The clusters do not only contain the ex-post behavior of real and artificial individuals but also their ex-ante given characteristics that influence the behavior. Hence, comparing clusters enables us to compare behavioral patterns in model-generated data and real data. To

provide an instance of application of the method, we used the built ACE mobility demand model. The model obtained a satisfactory validation score that shows that, in this case, the mDGP can mimic the rwDGP successfully for the given variables.

In general, the proposed empirical validation method has certain advantages. First, it fully leverages the specificity of ACE models covering highly heterogeneous agents and their potential multi-level aggregation. An ACE model can be initialized at the micro-level, be calibrated at the macro-level, and be validated at the meso-level with the same data set and for the same time frame. A procedure that is often used in time series to calibrate the model for a first segment of time periods and then validate in out-of-sample successive time suffers from the necessity of assuming that there are no structural breaks over time. This assumption may not be particularly suitable for models looking for emerging properties, high non-linearities and, indeed, structural breaks. The second advantage is that with the meso-level validation, we can compare the behavioral patterns that the mDGP and the rwDGP generate, respectively. It is not easy with macro-level validation because it compares only the aggregates. Therefore, the relationship between the ex-ante given characteristics and ex-post generated behavior can't be easily compared. In short, we offer to the researchers using ACE models a method to empirically validate them, which is a crucial intermediate step in the overall useful application of this highly promising approach.

4.6.3 Future Work

In future work, we aim to couple the built model with other energy and mobility-related models in Switzerland. Our model requires some exogenous

information (i.e., input) such as fuel prices and the estimated fleet in the future years (e.g., the share of electric vehicles (EVs) in the fleet). That information could be the output of another model. Similarly, the output of the built model e.g., total mobility demand per mode in Switzerland, can be the input of another mode. So, coupling related models enable to exchange information dynamically.

4.7 Appendices

Appendix A

This appendix provides a numerical example concerning what happens if the method assesses a model having missing variables. We created a simplified experiment for the sake of clarity. In the experiment, we assume that thousand real agents take a decision in a deterministic way based on a linear combination of five variables with fixed coefficients. We prepared two very simple models for the validation procedure. In the first, the artificial agents use the same variables and the correct coefficients. In the second, the artificial agents use a truncated number of variables (the first three in the provided file)¹³ missing the last two. It uses the right coefficients for the three variables. Then, we obtained the models' outputs. For both outputs, we applied the overall structure of the validation method. We placed real and artificial instances in a latent space¹⁴ and found the optimal number of clusters thanks to the ASW algorithm, which was twelve. Then, we divided the individuals into twelve clusters and obtained the quantity of real and artificial ones as in Table 4.2.

¹³We provide the variables along with the decisions of real and artificial individuals in an excel file on this link: <https://drive.switch.ch/index.php/s/4qiWLhnz8YNaxla>)

¹⁴We provide a created latent space in an excel file for the sake of clarity: <https://drive.switch.ch/index.php/s/QcE10MUC2j8ThuL>

Clusters	Without Missing Variables (Model 1)			With Missing Variables (Model 2)		
	Artificial Agents	Real Individuals	Score	Artificial Agents	Real Individuals	Score
Cluster 1	109	109	0	103	122	0.08
Cluster 2	84	84	0	100	81	0.10
Cluster 3	78	78	0	83	113	0.10
Cluster 4	105	105	0	99	126	0.12
Cluster 5	68	68	0	80	52	0.21
Cluster 6	92	82	0	86	91	0.02
Cluster 7	75	75	0	94	97	0.01
Cluster 8	92	92	0	57	41	0.16
Cluster 9	85	85	0	83	73	0.06
Cluster 10	53	53	0	70	62	0.06
Cluster 11	74	74	0	65	77	0.08
Cluster 12	85	85	0	80	65	0.10
Overall	1000	1000	0	1000	1000	0.09

TABLE 4.2: Quantity of Artificial Agents and Real Instances

As shown in Table 4.2, the first model produces a perfect balance; the quantity of real and artificial individuals are equal among clusters since the model isolates and considers all required variables. That is the extreme case, but we used it as an example for lucidity. The second model misses two variables, so there are deviations between ex-post behaviors of real and artificial individuals. Therefore, the balance of real and artificial individuals is not as good as the first model. Based on the results in Table 4.2, the indicator produced two scores for both models. We judged these results thanks to having their statistical distributions and got goodness-of-validation percentages as in Table 4.3.

Without Missing Variables (Model 1)		With Missing Variables (Model 2)	
Indicator's Score	Goodness-of-validation	Indicator's Score	Goodness-of-validation
0	0.0%	0.09	0.032%

TABLE 4.3: The scores and their judgments by statistical distributions

The first model (the extreme case) got zero from the indicator since all the artificial individuals behave similarly to real ones. Thus, it is validated at zero percent since there is no better model for producing a better result. The second model could obtain the score of 0.09961, which is worse than 0.03216% of all

possible scores that indicator could produce. The numerical example shows us that missing variables could turn out to have a lower validation score. In other terms, the indicator is sensitive to the misalignment between the model and the empirical world arising from missing variables and produces lower validation for models for which this misalignment is larger. A systematic enquiry into the effects of biases in the coefficients, the number and features of missing variables is, needless to say, beyond the scope of this paper.

Appendix B

In this appendix, we develop a numerical example of how the indicator (see Eq. 4.7) works and how we can judge the obtained score thanks to the statistical distribution of all possible scores, for the sake of readers. Let's assume that we have 50 artificial and 50 real agents. As explained in section 4.2, the number of real and artificial individuals must be equal, which is the initial condition to get a validation score through the method. Let's assume that we have two configurations as examples for the sake of clarity; the optimal number of clusters is obtained as three and five, respectively. For each configuration, we assume three situations; perfect balance, fair balance, weak balance. For each situation, we created the fictitious real-artificial number of individuals among clusters as illustrated in Table 4.4 to employ the indicator and calculate the score. According to ex-post behavior, which the mDGP generates, one could technically end up with all these situations¹⁵.

¹⁵For both configurations, we created state spaces without quantization (i.e., with the quantum size of one). The exhaustive list of state spaces can be downloaded via: <https://drive.switch.ch/index.php/s/demnJYBRJ8xgyje> and <https://drive.switch.ch/index.php/s/uQYgktIiz8XBQ09>

Configurations and Situations				
Configurations	Clusters	Perfect Balance	Fair Balance	Weak Balance
3-Clusters	Cluster 1	25 R - 25 A	30 R - 15 A	40 R - 10 A
	Cluster 2	15 R - 15 A	15 R - 25 A	15 R - 25 A
	Cluster 3	10 R - 10 A	5 R - 10 A	5 R - 10 A
	Total	50 R - 50 A	50 R - 50	50 R - 50 A
5-Clusters	Cluster 1	15 R - 15 A	20 R - 12 A	30 R - 5 A
	Cluster 2	10 R - 10 A	8 R - 18 A	5 R - 15 A
	Cluster 3	12 R - 12 A	5 R - 12 A	3 R - 18 A
	Cluster 4	8 R - 8 A	10 R - 3 A	10 R - 2 A
	Cluster 5	5 R - 5 A	7 R - 5 A	2 R - 10 A
	Total	50 R - 50 A	50 R - 50	50 R - 50 A

TABLE 4.4: The number of real (R) and artificial (A) individuals in the clusters according to the assumed configurations and situations

Based on the balances among clusters as depicted in Table 4.4, the indicator in Eq. 4.7 would create the following scores as illustrated in Table 4.5 below.

Configuration	Situations		
	Perfect Balance	Fair Balance	Weak Balance
3-Clusters	0	0.3055	0.5888
5-Clusters	0	0.3530	0.6523

TABLE 4.5: The scores generated by the indicator

Thanks to the computed all possible scores, we could easily judge the scores in Table 4.5. We sum up the number of cases that lead to better scores than the specific score and divide it by the total number of cases. So, we could get p-value alike percentages, which are represented in Table 4.6.

The reasons for computing the statistical distribution of all possible scores to judge a specific score is the following: First, the scores that the indicator could generate are not evenly distributed (see Figure 4.7). Second, they are sensitive to the number of clusters. Accordingly, the percentage value enables comparability across models, which might have different numbers of clusters or agents. Similarly, it allows to compare how well a model is validated in two time-space contexts (e.g., where the model is better validated in two countries or when the

model is better validated across different years). Third, as a percentage and not absolute value is similar to a p-value; thus, the goodness-of-validation can be assessed using very well-established conventional thresholds (e.g., 1% and 5%).

Configuration	Situations		
	Perfect Balance	Fair Balance	Weak Balance
3-Clusters	0	0.0797	0.5163
5-Clusters	0	0.1222	0.6437

TABLE 4.6: Percentage values

An example for the interpretation of a particular score and its judgment would be as following: A model, who could achieve the score of 0.3055 (fair balance in Table 4.5) with the given inputs and the found optimal number of clusters (3 clusters for this example), performs worse than 0.0797% of the all possible scores. We interpret the results in this way that this model is validated by the threshold at 0.0797%.

Appendix C

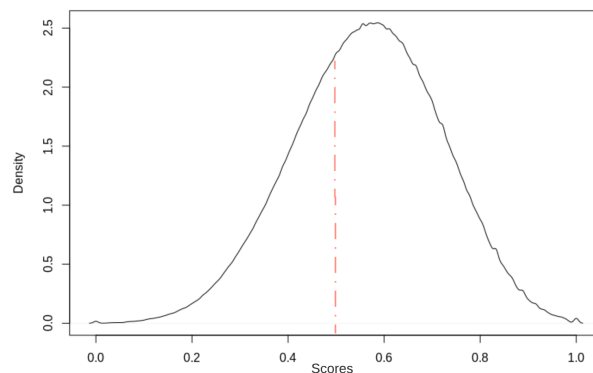


FIGURE 4.8: Density distribution of all possible validation scores with 300 quantum size (instead of 150 in Figure 4.7)

Appendix D

Algorithm 3 Pseudo-code of the k-Medoids Algorithm

Input: m (a dissimilarity matrix), k (the number of clusters)

Output: k clusters

Initialization: select arbitrary k of the n cases as the medoids

```
1: while The Mean Square Error (MSE) decreases do  
2:   Distribute  $n$  cases to  $k$  medoids  
3:   Update medoids from the cases added to the cluster  
4:   Calculate MSE for each update  
5:   if MSE does not decrease anymore then  
6:     break  
7:   end if  
8: end while  
9: return  $k$  clusters
```

Chapter 5

Analyzing Future Scenarios & Decarbonization Pathways for the Swiss Mobility Sector with an ACE model¹

Abstract

The chapter discusses a use case of the model to analyze different future scenarios and decarbonization pathways for the Swiss mobility sector. A scenario is the combination of the trajectories of several variables, including socio-economic dynamics, policy measures, technical parameters, new lifestyle patterns, Swiss mobility characteristics, and electricity and fuel system characteristics. The model computes for each scenario individual mobility demands, which are aggregated to get the macro patterns such as total mobility demand per mode and total CO₂ emissions. The chapter compares the macro outputs from each scenario and discusses whether they deliver the net-zero targets from the mobility

¹The chapter is a part of an umbrella project called JA Mobility. The analyses and results discussed throughout the chapter have been largely carried out through the BedDeM model. Deliverables of the umbrella project can be found in (Piana et al., [2020](#); Schumann, Weigt, and Kober, [2020](#))

sector in 2050, the target year of the Swiss Federal Council, as discussed in previous chapters.

Apart from the discussions at the macro level, the chapter dives into the micro-level analysis. Agents in the built model are heterogeneous in their utility functions, endowments, constraints, and socio-economic characteristics. We expect that, due to the heterogeneity, their responses to the changes in their environments would be different. The chapter conducts a micro-level analysis to test that. The analysis first investigates agents' reactivity to changes in their economic environment. It compares the micro-behavior of agents in two scenarios that differ in fuel taxes and mobility prices. It examines whether some agents change their behavior while others don't. In other words, it looks at whether agents' reactivity to the price stimuli is heterogeneous, which can't be captured in the model with representative agents. Then, the micro-level analysis introduces two policies targeting specific segments (i.e., sub-groups). The policies are also a part of the scenarios discussed in the previous paragraph. The chapter picks up the scenarios with the same scenario variables except for the formulated policy variable (i.e., the scenarios with and without a specific policy). It analyzes the micro-behavior of agents with and without policy. It examines which agents have responded to the policy and changed their behavior. It also builds a bridge between Chapter 3 and Chapter 5 by analyzing policy responses at the segment level. That is, it discusses which segments are more responsive to the implemented policies and whether they align with the target groups of the policy formulation phase.

5.1 Introduction

The second and third chapters of the thesis have analyzed, revealed, and quantified heterogeneity among Swiss households regarding mobility behavior and characteristics. Then, Chapter 4 has built and validated an ACE mobility demand model that incorporates the insights from Chapter 2 and 3. The model involves both structural and behavioral heterogeneity among agents. That is, agents in the model can have different socio-economic characteristics or endowments (i.e., resources) as well as utility functions. Chapter 4 has also presented an empirical validation method for ACE models and applied it to the built model. The results have revealed that the model can sufficiently produce stylized facts that we observe in real data. Once the model is validated, it becomes deployable and ready to use, as described in the procedure of Klügl (2008). The answers derived from the model can be utilized to answer questions directed to the original system. This chapter introduces a socio-economic study, a part of which was carried out through the built model.

This chapter aims to analyze future scenarios and decarbonization pathways for the Swiss mobility sector. In other words, it examines the evolution of mobility in Switzerland from 2015 through 2050, which is the time frame in which the Swiss Federal Council has committed to be climate neutral (Bundesrat, 2021). The evolution of mobility term is used in the meaning of a systematic change in the mobility sector for which a bunch of dimensions evolves over time, deflecting from a *ceteris paribus* assumption. The chapter first defines variables representing these dimensions whose changes are relevant for the evaluation of mobility. The variables can be passive, which have only one trajectory that does

not vary across scenarios (e.g., estimated population growth in the future, or energy prices), as well as active (e.g., fuel tax trajectory). The product of different assumptions of the trajectories of the active variables constitutes a scenario. To be more concrete, assume that we have three scenario dimensions, each represented by an active variable, and we assume four different trajectories for each of these active variables. The total number of scenarios that would exhaust the space of all possible combinations becomes 64 (4^3).

The scenarios consist of the combination of the trajectories of the ex-ante determined twelve active variables. The active variables and the assumed trajectories for them will be discussed in the following section. The combination of different trajectories of the active variables yields 640 scenarios to be analyzed. The chapter analyzes each scenario and computes the estimated CO_2 emissions and mobility demand (i.e., traveled kilometers) for each between 2015-2050. That analysis is called *macro-level analysis* throughout the chapter. The macro-level analysis helps us to foresee the mobility emissions and demand in 2050, by which the Swiss Federal Government has the net-zero emission target.

Two active variables are policy implementations, specifically targeting car user segments. The first policy is local regulations (LOCREG) that aim to close certain territorial areas of city centers to private cars. The second policy is public transportation evaluation, which seeks to improve the comfort of public transportation modes until 2050 gradually. The chapter analyzes the responses of agents to the policies. That is, it picks up the scenarios that differ only in the trajectories of policy trajectories (with and without policy) and compares their modeling results. It analyzes the micro-behavior of agents with and without policy implementation. It examines behavior changes of agents and their segments to understand whether the policies affected the ex-ante targeted sub-groups.

That analysis is called *micro-level analysis* throughout the chapter. The micro-level analysis also compares agents' reactivity to economic stimuli. Previous chapters discussed that individuals in the real world are observably heterogeneous, so they react differently to the changes in their economic environment. So models with a single representative agent cannot capture different reactivity to economic stimuli, which was the main argument for using the ACE approach in Chapter 4. The present chapter compares two scenarios; they differ only in the prices (i.e., with or without a high carbon tax and mobility pricing). The chapter demonstrates that some agents change their behavior following the price increase while others do not, which is difficult to be captured with representative agent models.

The chapter is a part of umbrella research called JA Mobility (Piana, Del Duce, and Schumann, 2019). The selection of active and passive variables has been carried out with collective effort in that research. Besides, different energy and mobility models, which will be discussed in the next section for the sake of clarity, have been coupled for scenario analysis. Nonetheless, other models and the coupling structure are out of scope for this chapter. This chapter discusses only the analyses and results, which have been carried out and produced by the BedDeM model.

5.2 The selection of active and passive scenario variables

Passive variables refer to dimensions on which no alternative futures need to be analyzed. That is, these variables have only one trajectory that does not vary

across scenarios. Active variables refer to dimensions on which more than one alternative is explored. That is, for these variables, different evolution trajectories are expected (assumed) because of divergent stress by various stakeholders or alternative socio-economic spontaneous dynamics. These variables directly define the characteristics of a scenario.

The list of passive scenario variables is provided in Table 5.1 below. These variables are exogenous, and there are no variants for them to be studied².

List of Passive Scenario Variables		
Socio-economic Dynamics	New Lifestyle Patterns	Swiss Mobility Characteristics
<ul style="list-style-type: none"> • Population Development • Economic Growth • Household Size • Household Income • Energy Prices 	<ul style="list-style-type: none"> • Flexible work forms / part time • Aging Population 	<ul style="list-style-type: none"> • Subscription to Mobility as a Service (MaaS) • Supply and Demand for Pooling Services

TABLE 5.1: Passive Scenario Variables and their categories (Schumann, Weigt, and Kober, 2020)

Table 5.2 shows active scenario variables and their acronyms that are used throughout the chapter. For these variables, more than a trajectory is assumed. Space of the combination of different trajectories for the active variables indicates the total number of scenarios to be evaluated. Nevertheless, the set of twelve active variables was judged still too broad, as it would have been required at least 2^{12} scenarios if all variables had only two alternative trajectories. Therefore, some active variables have been assumed to co-evolve together to obtain a manageable number of scenarios to evaluate in a certain time frame. That will be discussed in the experimental design section.

²Online Appendix A provides the assumed trajectories for all the passive variables in an excel file on this link: <https://drive.switch.ch/index.php/s/VgKZISIoQ9R69JI>

List of Active Scenario Variables	
Name of the active variable	Acronym
• Impact of on-line shopping and home delivery (e.g., in zero-carbon modes)	• ONLINE
• Home-office / Co-working spaces	• HOMEOFF
• Climate change awareness	• CCAWARE
• Cars occupancy rate	• OCCUP
• EV charging behavior	• CHARGE
• Taxation on fuel	• TAXFUEL
• Mobility pricing (a price per km traveled, with different rates per mode)	• MOBPRICE
• Local regulations for the transitions (e.g., access ban)	• LOCREG
• Evolution of public transportation (quality)	• PTEVO
• Subsidies	• SUBSID
• Penetration of autonomous taxis	• AUTOTAXI
• Behavioral policies(e.g.,nudges,awareness tools,etc)	• BEHAVPOL

TABLE 5.2: Active Scenario Variables and their acronyms (Schumann, Weigt, and Kober, 2020)

As mentioned in the introduction section, two active variables are policy variables (i.e., PTEVO and LOCREG). They are formulated to change the behavior of specific sub-groups in the population. Besides, TAXFUEL and MOBPRICE variables are used to analyze agents' responses to the changes in their economic environment. The scenario results of these policy variables will be analyzed at the micro-level.

5.3 Macro-level Analysis

The macro-level analysis hinges on creating a scenario space with the combination of different trajectories of the active variables, and analyzing aggregated agent behavior to find macro outputs for each scenario between 2015-2050. Macro outputs refer to aggregated CO₂ emissions and mobility demand throughout the chapter.

5.3.1 Model Coupling Structure

To evaluate different scenarios at the macro-level, four models have been coupled. Although the coupling is not in the scope of the thesis, it is mentioned shortly for the sake of clarity. One of the models is BedDeM, which has been built and validated in Chapter 4. The other three models are;

- STEM - a whole energy model in Switzerland (Ramachandran and Turton, 2014)
- MATSim - a microsimulation for mobility routing (W Axhausen, Horni, and Nagel, 2016)
- Swissmod - an electricity sector model (Schlecht and Weigt, 2014)

The STEM model estimates the optimal fleet between 2015 and 2050 based on mobility driving patterns that the MATSim model produces. So, one of the outputs of the STEM model is the fleet information, including the share of power trains (e.g., EV and internal combustion engine (ICE)) and fuel types (e.g., gasoline, diesel, gas, hydrogen). The fleet is disaggregated for each year and given to the agents in BedDeM. As discussed in the previous chapter, agents in BedDeM make economic decisions to carry out their mobility activities based on their decision-making mechanism. The model produces individual-level mobility demands over which macro-patterns are obtained through aggregation, such as total mobility demand per mobility mode and total CO₂ emissions. MATSim takes that individual-level trips and mobility demands, which BedDeM produces, and assigns them the origins and destinations. Finally, Swissmod filters the trips with EVs, generates a wealth of statistics referring to the electric system,

and checks for the robustness of the grid of the requested electricity. It evaluates possible effects on prices (e.g., at the national level).

The evolution of the fleet (e.g., technological developments) as well as energy prices (e.g., gasoline, diesel, hydrogen, gas, and electricity) are exogenous to BedDeM. It gets this information from the STEM model to be used in agents' decision-making. Two online appendices with the used fleet information and price estimations in [2015-2050] are provided for the sake of readers³

The results discussed in this chapter are produced by the BedDeM model. The structure of other models, as well as the coupling structure, is not relevant to the chapter. Nonetheless, we mention them briefly to clarify how exogenous variables are provided to the BedDeM model, such as the fleet.

5.3.2 Trajectories for the active variables

The first active variable is ONLINE for which two different trajectories are assumed. The variable is in the scope of the new lifestyle patterns dimension. It is assumed that in the future the usage of the online shopping will be increased that will reduce the number of mobility activities (i.e., trips) with shopping purposes. According to the first trajectory, total number of mobility activities is assumed to decrease by 3% per decade. That is, at the end of a decade (e.g., 2030) there will be 3% less trips than the beginning of that decade (e.g., 2020). The second trajectory assumes 10% decrease per decade instead.

The second active variable is HOMEOFF, for which two different trajectories are assumed. The variable is in the scope of the new lifestyle patterns dimension. It is assumed that more people will work remotely in the future, which will affect

³Online appendix B on the estimated fleet: <https://drive.switch.ch/index.php/s/13IBDtIAaoux1sV>. Online appendix C on the estimated energy prices: <https://drive.switch.ch/index.php/s/p2YLVw3EKY2xVRo>

total mobility demand in Switzerland. According to the first trajectory of the variable, the total number of mobility activities is assumed to decrease by 3% per decade. The second trajectory assumes a 10% decrease per decade instead.

The third variable is CCAWARE, for which two different trajectories are assumed. The variable refers to the climate change awareness of the Swiss population. The first trajectory assumes that there will be no change in that variable. It will be exactly the same as we observe in the MTMC data. The second trajectory assumes that 20% of the Swiss population will be more climate change aware, affecting their mobility consumption behavior. BedDeM model can capture behavioral heterogeneity among agents, as discussed in the previous chapter. That is, agents might have different CCAWARE values that affect their behavior. The model can capture mobility behavior changes (if any) due to becoming more climate change aware.

The fourth variable is OCCUP, for which five different trajectories are assumed. The variable is in the scope of the Swiss mobility characteristics dimension. According to the MTMC data, the current occupancy rate for private car trips is about 1.56 in Switzerland (ARE/BfS, 2017). Each of the trajectories assumes different occupancy rates for private car trips. The first trajectory assumes that it will stay steady at 1.56. That is, in 2050, the occupancy rate for private car trips will also be 1.56. The second trajectory assumes that it will decrease and equal to one in 2050. The rates in between [2015-2050] are linearized. The third trajectory assumes that it will slightly decrease and equal to 1.4 in 2050. The fourth trajectory assumes that it will slightly increase and equal to two 2050. And the final trajectory assumes that it will dramatically increase and equal three in 2050.

The fifth variable is CHARGE, for which two trajectories are assumed. The

variable is in the scope of the charging behavior dimension. It reflects the coordinated (controlled) charging percentage for electric vehicles (EVs). The first trajectory assumes a low coordinated charging: all the EV owners charge their cars as soon as they have access to the grid (e.g., when they are back home). The second trajectory assumes a high coordinated charging: all the EV owners let a centralized planner assign charging time slots. As a result, the centralized planner shifts a part of charging to off-peak hours (with lower electricity prices). Note that the centralized planner ensures that the batteries have enough energy stored for the EV owners to carry out their daily mobility activities.

The sixth variable is MOBPRICE, for which two different trajectories are assumed. The variable is in the scope of the mobility pricing dimension. The first trajectory assumes to charge sixteen Swiss Rappen (centimes) per traveled private car kilometers in 2050. The MOBPRICE values between 2015 and 2050 are linearized accordingly. The second trajectory assumes to charge thirty Swiss Rappen per traveled private car kilometer in 2050.

The seventh variable is TAXFUEL, for which two different trajectories are assumed. The variable is in the scope of the taxation on fuel dimension. According to the first trajectory, the taxation on fuel (e.g., gasoline, diesel, gas) will stay steady until 2050. The second trajectory assumes a 60% tax increase in 2050. We note that the increase affects only the tax component of the retail price. It does not change other components such as crude oil and distribution & marketing. The tax rates between 2015 and 2050 are linearized similar to other variables' trajectories.

The eighth variable is LOCREG, for which two different trajectories are assumed. The variable indicates the percentage of areas that are covered by personal car bans. That is, individuals having trips from or to these areas are not

allowed to use their cars. They can use other modes instead. The first and soft trajectory assumes that about 5% of the total mobility zones will be prohibited from private vehicles by 2050. The second and strong trajectory assumes that about 20% of the mobility zones will be prohibited for personal car use by 2050.

The ninth variable is PTEVO, which indicates the evolution of public transportation in Switzerland. Two different trajectories are assumed for the variable. The first trajectory assumes that the comfort attribute of public transportation modes will increase by 3% per decade until 2050. The second trajectory assumes the increase will be 20% per decade. Agents in the built model evaluate the comfort of the mobility modes while making decisions.

The tenth variable is AUTOTAXI, which indicates the penetration of autonomous taxis. The first trajectory assumes that autonomous taxis will be available first by 2030. Their pro kilometer cost will be sixty times the current per kilometer cost of a standard EV. That will reduce to forty times by 2050. According to the second trajectory, the pro km cost of autonomous taxis will be forty times of standard EVs. That will reduce to twenty times by 2050.

The eleventh variable is SUBSID, which indicates subsidies for the cost of durable goods and their installation (e.g., new charging stations, new vehicle technologies). The variable has two trajectories; weak and large subsidies. The variable is exogenous to our model. It has been used by the STEM model to forecast fleet composition. The fleet is distributed to the agents in the BedDeM model to forecast mobility demand.

The last variable is BEHAVPOL that indicates the implementation of behavioral policies such as nudges and awareness tools. Two different trajectories are assumed for the variable. The first trajectory assumes no behavioral policy implementation. The second trajectory assumes that the whole population

is targeted with nudges. The trajectories of the active variables as well as their current values can be found in Appendix A⁴.

5.3.3 The space of scenarios

The scenario space refers to all the combinations of the trajectories of the active variables. Passive variables such as energy prices have only one trajectory that does not vary among scenarios. We have twelve active variables, each with at least two trajectories. That means that the scenario space is not feasible in terms of the cardinality of scenarios. Therefore we assumed that some of the active variables will co-move. That is, their trajectories change across scenarios together. We assumed that ONLINE and HOMEOFF, CCAWARE and BEHAVPOL, and MOBPRICE and TAXFUEL variables co-move across scenarios. That shrinks the space of scenarios and reduces the total number of scenarios to be evaluated.

Scenario Number	CHARGE	SUBSIDIES	MOBPRICE	TAXFUEL	LOCREG	PTEVO	OCCUP	AUTOTAXI	CCAWARE	BEHAVPOL	ONLINE	HOMEOFF
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	2	2
3	1	1	1	1	1	1	1	1	2	2	1	1
4	1	1	1	1	1	1	1	1	2	2	2	2
5	1	1	1	1	1	1	1	2	1	1	1	1
6	1	1	1	1	1	1	1	2	1	1	2	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
640	2	2	2	2	2	2	5	2	2	2	2	2

TABLE 5.3: The space of scenarios as a matrix (the numbers represent the trajectories) (Piana, Del Duce, and Schumann, 2019)

The scenario space is represented by a matrix, the structure of which is shown in Table 5.3. The complete matrix, as well as detailed information about the trajectories of the active variables, are provided as an online Appendix A. Besides, detailed information about the sources used for the trajectory assumptions can be found in the appendix as well as in (Piana, Del Duce, and Schumann, 2019).

⁴Appendix A is online available on <https://drive.switch.ch/index.php/s/VgKZISIoQ9R69JI>

5.4 Micro-level Analysis

The micro-analysis firstly aims to show that agents in the model are heterogeneous; so that they react to the changes in their environment differently. To test that, two specific scenarios' outputs are compared. To be more concrete, the output of scenarios that differ only in the trajectories of price-related active variables (TAXFUEL and MOBPRICE) are compared. That is, the behavior of agents with and without pricing policies (e.g., with high and low fuel prices) are compared. The goal is to observe whether some agents change their behavior while others don't. So, we can demonstrate that using the ACE approach over representative agent models brings an added value in the thesis since representative agent models can't capture heterogeneous responses of agents to policies. They assume everybody in the population changes their behavior when the representative agent changes it.

The micro-analysis secondly aims to formulate policies targeting specific sub-groups in the population. The policies are active variables whose trajectories are discussed in the previous section. The first of these policies is the local regulations for transitions (e.g., access ban), LOCREG. That aims to block certain parts of city centers to private cars. The model uses the trips of the respondents in the MTMC data. For each trip, there is departure and destination information. There is also information regarding the location type of the departure and destination (e.g., city center, rural municipality, sub-urban area, etc.). When the policy is implemented, the agents with a trip to (or from) banned sites can't use their private cars for that trip. The policy thus primarily targets car user segments. The second policy is the evaluation of public transportation. As previously discussed, it has two trajectories with different improvements in public

transportation quality. The policy targets car user segments primarily too.

The third and last aim of the micro-analysis is to observe whether the formulated policies affect only target segments (i.e., private car users) as envisaged. To test this, the results of the scenarios, which differ only in the trajectories of the policy variables (i.e., LOCREG and PTEVO), are compared. The scenarios with the same active variable trajectories but differ in PTEVO trajectories are compared. The same is carried out for LOCREG. The results will be discussed in the next section.

5.5 Results

5.5.1 Results from the Macro-Analysis

Estimated mobility demands across scenarios

The modeling results show that the traveled personal kilometers (Pkm) will rise across all 640 scenarios in 2050. It is mainly due to the increase in demand as a consequence of the assumed population growth, which is a passive variable with one trajectory and does not change across scenarios. Figure 5.1 demonstrates the estimated Pkm across scenarios in 2050. The scenarios in the x-axis are ordered according to their Pkm estimations. As shown in the figure, two broad groups of scenarios are characterized by low and high Pkm growth.

The scenarios that assume a strong home office (HOMEOFF) and online shopping (ONLINE) behavior increase, and strong access bans for private cars (LOCREG) lead to a lower demand increase according to the modeling results (the first 160 scenarios in the x-axis in Figure 5.1). The model estimates slightly

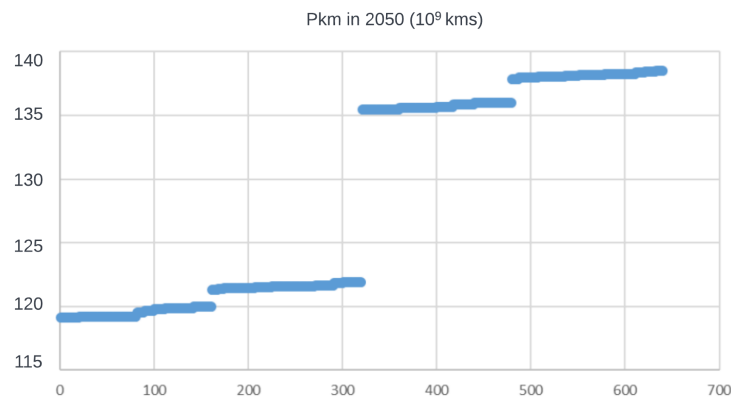


FIGURE 5.1: Each coordinate on the x-axis is a scenario, and y-axis is corresponding Pkm in 2050

more demand increase for the ordered scenarios between 160 and 320, which assume a weak LOCREG but a strong HOMEOFF and ONLINE behavior increase. Similarly, the model estimates a high demand increase for the ordered scenarios between 480 and 640, which assume a weak LOCREG and weak HOMEOFF and ONLINE behavior increase. The model estimates slightly lower demand increase for the ordered scenarios between 320 and 480 than those between 480 and 640, which assume a strong LOCREG but weak increase in HOMEOFF and ONLINE behavior. According to the modeling results, other active scenario variables exert limited influence on the Pkm. LOCREG, ONLINE and HOMEOFF are the scenario dimensions that impact Pkm significantly.

According to the modeling results, Scenario 156 leads to the lowest private car usage. The scenario assumes a strong access ban for private cars and a high increase in online shopping and home office behavior. The scenario also assumes a high rise in the taxation on fuel and mobility prices, a high climate change awareness, and a high increase in the improvement of the comfort of public transportation.

The model estimated the highest private car usage for Scenario 1, which assumes no private car access ban and a soft increase in online shopping and home office behavior. Besides, the scenario assumes low taxation on fuel and mobility prices, low climate change awareness, low increase in the improvement of the comfort of the public transportation, and no behavioral policy implementation.

According to the modeling results, Scenario 311 leads to the highest public transportation use. The scenario assumes a high tax on fuel and mobility prices, a high increase in the improvement in the comfort of public transportation, and high climate change awareness. Besides, the scenario assumes the implementation of behavioral policies and a soft rise in online shopping and home office behavior.

Estimated Carbon-dioxide emissions per scenario

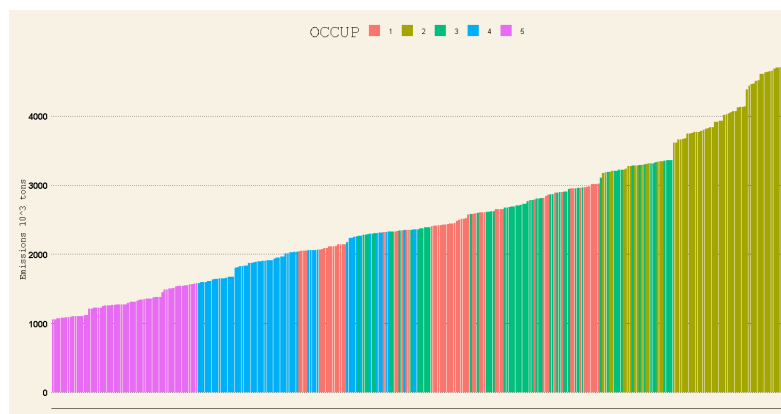


FIGURE 5.2: Ordered scenarios and their OCCUP trajectories

This section demonstrates the effect of the active variables' trajectories, which constitute the scenarios, on the estimated CO_2 emissions. In Figure 5.2 (and also in the following figures), the 640 scenarios are ordered according to their CO_2

emissions estimations. The scenarios in the figures are colored according to trajectories of active variables. Figure 5.2 shows the assumed trajectories of the occupancy rate variable in the scenarios and corresponding CO_2 estimations. As Figure depicts, occupancy rate plays a overarching role in determining mobility emissions. The scenarios with the occupancy rate trajectory five, which assumes that the occupancy rate will increase dramatically and equal to three in 2050, lead to the lowest CO_2 emissions in 2050. These scenarios lead to, on average 48% less CO_2 emissions in 2050 than the scenarios with the occupancy rate trajectory one, which assumes today's occupancy rate (i.e., 1.56) value constant until 2050. Similarly, the scenarios with the occupancy rate trajectory two, which assumes that the occupancy rate will decrease dramatically and equal to one in 2050, lead to the highest CO_2 emissions in 2050. They lead to, on average, 56% more CO_2 emissions in 2050 than the scenarios with today's occupancy rate (i.e., the trajectory one). The scenarios with the occupancy rate trajectories four and three, which assume slightly increase and decrease respectively, are in the center, as Figure 5.2 shows.

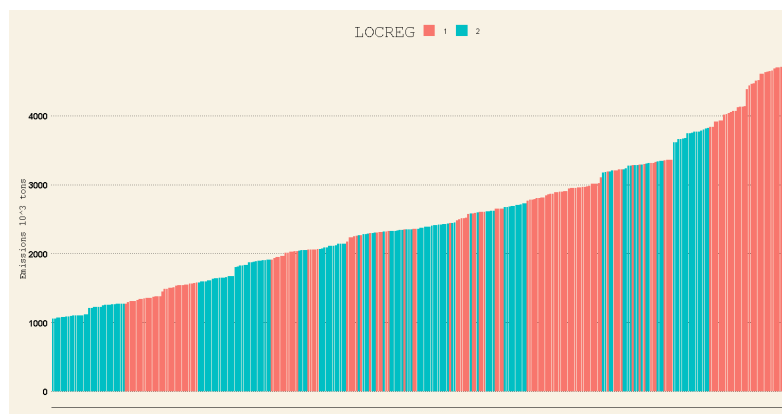


FIGURE 5.3: Ordered scenarios and their LOCREG trajectories

Figure 5.3 shows the ordered scenarios colored according to their LOCREG trajectories. The scenarios with the second LOCREG trajectory, which assumes a

strong access bans for private cars, lead to lower CO_2 emissions in 2050 keeping all other variables constant. The scenarios for which the model estimates high CO_2 emissions (i.e., the ones on the right end in Figure 5.3) assumes the first LOCREG trajectory that assumes a soft access bans for private cars.

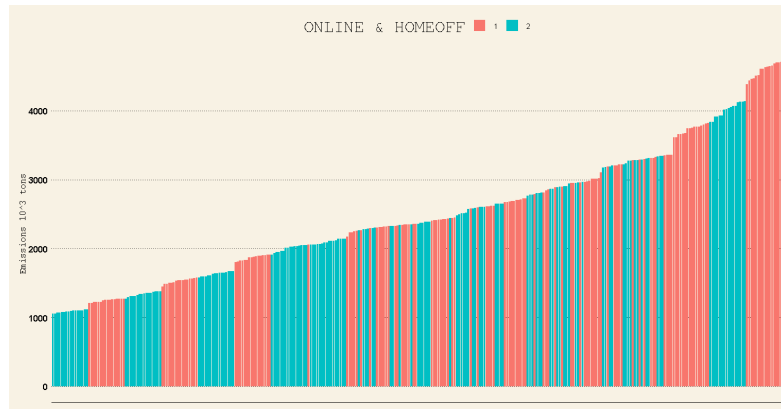


FIGURE 5.4: Ordered scenarios and their ONLINE and HOME OFF trajectories

The trajectories of online shopping and home office behavior variables are also pivotal to determining CO_2 emission estimations. Figure 5.4 shows that, the scenarios with the lowest CO_2 emissions estimations in 2050 (i.e., the ones on the left end in Figure 5.4) assume a strong increase in online shopping and home office behavior (i.e., the trajectory two). Similarly, the scenarios for which the model estimates the highest CO_2 emissions (i.e., the ones on the right end in Figure 5.4) assume the first trajectory for these two variables.

Figure 5.5 shows ordered scenarios with colors according to their public transportation evolution trajectories. The scenarios with the second trajectory, which assumes a significant improvement in the comfort of public transportation services, lead to lower CO_2 emissions in 2050. That is, in these scenarios,

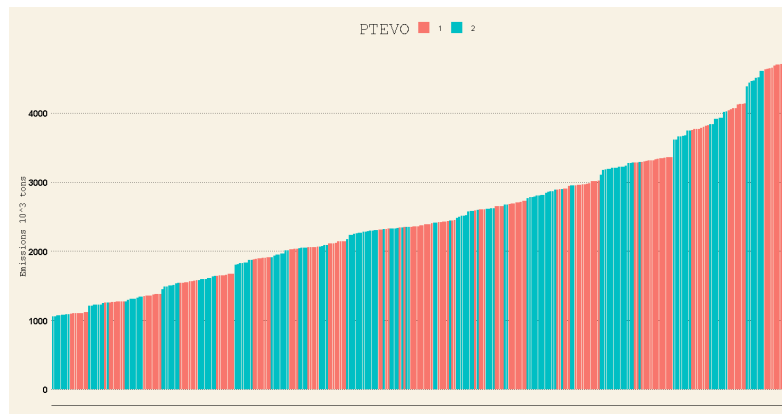


FIGURE 5.5: Ordered scenarios and their PTEVO trajectories

agents consumed more public transportation due to high comfort and consequently fewer private cars, which caused to reduction mobility emissions partially.

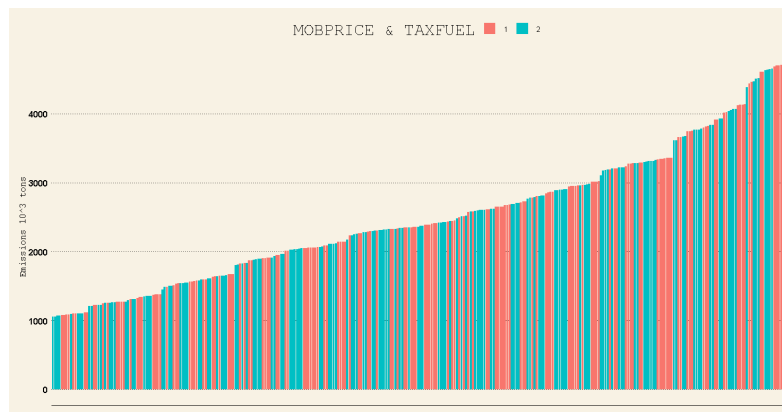


FIGURE 5.6: Ordered scenarios and their mobility price MOBPRICE and TAXFUEL trajectories

The trajectories of the price-driven decarbonization policies, which are mobility price and taxation of fuel, across the ordered scenarios are illustrated in Figure 5.6. These policies are not as significant as the previous variables such as occupation rate, online shopping, or public transportation evolution. Nevertheless, as Figure 5.6 shows, the scenarios with the second trajectory, which assumes a higher tax on fuel and mobility price, lead to lower CO_2 emissions in

2050. That is, some agents in the model reacted to price changes and chose another mode than the private car, which caused reductions in mobility emissions. Nonetheless, the reactions to the price changes are heterogeneous; some agents react, some do not, which we will discuss later. Similar illustrations for other active variables are provided in an online Appendix⁵.

5.5.2 Results from the Micro-Analysis

As discussed in the previous section, the micro-analysis first examines agents' reactions to the changes in their economic environment to observe whether agents react to the changes differently due to heterogeneity. That is, the goal is to analyze whether some agents respond to changes while others do not. To test that, we looked for the modeling results of two scenarios that only differ in the assumptions on the trajectories of fuel tax and mobility pricing. The scenarios are Scenario 1 and 161, respectively. The former assumes that the fuel tax in Switzerland will stay steady from 2015 until 2050, and there will be no mobility pricing implementation. The latter assumes that fuel tax will be 60% higher in 2050 than the 2015 level, and mobility pricing will be implemented (Online Appendix A provides the assumptions for the trajectories of both scenarios). The rest of the scenario variables (both active and passive) as well as the fleet are identical in both scenarios. We compared the modeling results for both scenarios for the final year 2050. We particularly analyzed the behavior of the same agent in Scenario 1 and 161, respectively. The results demonstrate that agents' responses to tax changes and mobility pricing implementation are heterogeneous.

⁵Online Appendix E available on: <https://drive.switch.ch/index.php/s/zAPbh94dIZfh8eu>

About 4% of the agents changed their behavior in Scenario 161 relative to Scenario 1 and used public transportation (especially train) and soft mobility modes instead of the private car. However, others did not change their behavior. That is, agents' demand is relatively inelastic to price (both taxation on fuel and mobility prices per kilometer). The rationale for that would be; firstly, there is a large existing price gap between private cars and public transportation (in favor of the former). That is, even the prices are raised in some scenarios, private cars are cheaper than public transportation, especially for the agents without any seasonal tickets. Secondly, the fleet is assumed to be more efficient and EV-dominated in the future (see the online Appendix C for the estimated fleet trajectory). That counterbalances the increased fuel prices that can explain why agents do not change their mobility behavior. Thirdly, budget exhaustion is very unlikely because incomes are rising, along with the GDP, which is a passive scenario variable (see online Appendix A for the estimated income trajectory). Finally, some agents using a private car can choose it out of their judgments on comfort as well as their habits. Thus, they can continue to use their cars for particular trips even if it's more expensive than public transportation. Online Appendix D⁶ provides micro-level behavior of agents in Scenario 1 and 161. The appendix also provides the micro-level behavior that is different in two scenarios (i.e., 4% of the agents change their behavior) for clarity. So, that part of the analysis showed that agents in the model are heterogeneous. They have different responses to policies that can't be captured by the models with a single representative agent since they assume that the whole population changes (or does not change) its behavior when the representative agent changes (or does

⁶Appendix D is online available on: <https://drive.switch.ch/index.php/s/DWJt3M9mZGNKBVi>. Please note that we hid the agents' household IDs, which are given to the agents from the MTMC data, due to data privacy concerns.

not change) its behavior.

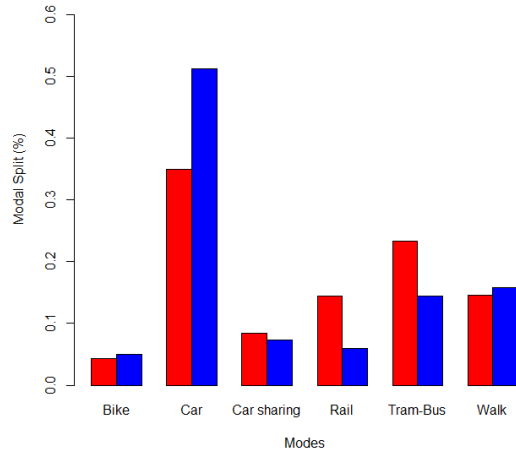


FIGURE 5.7: Modal Split with (red) and without (blue) LOCREG

Another goal of the micro-analysis was to target specific segments with policies to change their behavior and observe in the model whether the policies affect the targeted segments as envisaged. As discussed in the previous chapter, two policies have been formulated for private car user segments; LOCREG and PTEVO. The modal split with and without LOCREG policy in the final year of the analysis (2050) has been illustrated in Figure 5.7. As shown in the figure, there is a behavior shift from the car to train (rail) and tram-bus. Then, we looked at the proportion of the agents who change their behavior among the behavioral segments obtained in Chapter 3. The results revealed that the policy affects primarily private car user segments, as ex-ante envisaged (see Table 5.4). The proportion of agents who changed their behavior following the LOCREG policy is remarkably higher in segments 2 and 11, whose members primarily have a half-fare travel card. We also observe minor changes in other segments than private car user segments. These are mainly due to the stochasticity that

the model has. That is, two different modes can have similar expected utilities; so that the agent's choose both in different simulation runs.

Changes per Segment			
Segment Number	Change	Primary Mode of Transport	Segment size (% of the population)
1	1.2%	Walking	13%
2	9.7%	Private Car	6%
3	0.8%	Tram	3%
4	7.2%	Private Car	27%
5	1.3%	Train	4%
6	0.5%	Bus	2%
7	0.3%	Bike	4%
8	6.3%	Private Car	10%
9	0.1%	Bus	2%
10	0.7%	Walking	9%
11	10.4%	Private Car	12%
12	0.2%	Bike	3%
13	0.6%	Train	5%

TABLE 5.4: The effect of LOCREG at the segment-level

Figure 5.8 demonstrates the modal split with and without the PTEVO policy in 2050. As shown in the figure, there is a demand shift from private car to public transportation modes. The proportion of other modes are almost the same across both scenarios. The improvement in the comfort of the public transportation modes affect mainly private car users, as they shift their behavior. We see that in Table 5.5 that the proportion of agents who changed their behavior after the implementation of the PTEVO policy is significantly higher in private car user segments. It is even higher among segments whose members have largely half-fare travel pass and are multimodal.

Changes per Segment			
Segment Number	Change	Primary Mode of Transport	Segment size (% of the population)
1	0.1%	Walking	13%
2	9.2%	Private Car	6%
3	0.1%	Tram	3%
4	4.5%	Private Car	27%
5	0.4%	Train	4%
6	0.3%	Bus	2%
7	0.1%	Bike	4%
8	5.7%	Private Car	10%
9	0.1%	Bus	2%
10	0.7%	Walking	9%
11	8.1%	Private Car	12%
12	0.2%	Bike	3%
13	0.3%	Train	5%

TABLE 5.5: The effect of PTEVO at the segment-level

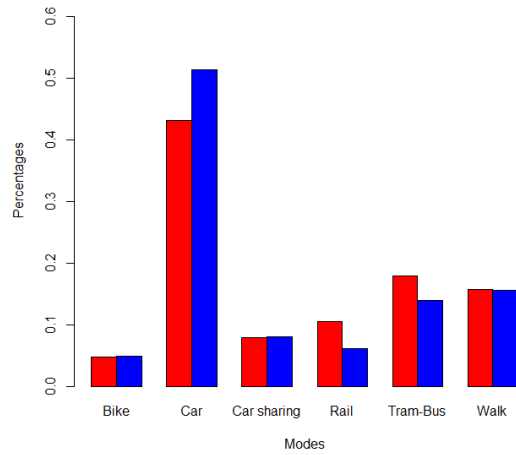


FIGURE 5.8: Modal Split with (red) and without (blue) PTEVO

The micro-analysis show that the model involves heterogeneity, and agents can react to policies differently. The results revealed that with the implementation of LOCREG and PTEVO policies, private car use could be reduced significantly. That would reflect positively on mobility emissions. The policies target mainly private car user segments; the proportion of other modes stays relatively steady.

5.6 Conclusion

This chapter presented a use case of the BedDeM model for socio-economic analysis. It conducts and discusses macro and micro-level analyses using the use case results. The results from the macro-level analysis show that, in all scenarios, mobility demand in terms of personal kilometers traveled will rise in 2050 in Switzerland. The main driver of that rise is the estimated population growth in 2015-2050, which is used as a passive variable across scenarios. Nevertheless, the

estimated increase in demand is heterogeneous; there are low and high demand scenarios. For the former, the model forecasts about a 13-18% demand increase in 2050 relative to today's demand. For the latter, that ratio is about 28-32%. Low demand scenarios assume a strong access ban for private cars and a substantial increase in online shopping and home office behavior. These two assumptions differentiate low-demand scenarios from high-demand ones. Besides, none of the scenarios lead to net zero emissions in 2050. The model estimates about ten million tons of CO₂ emissions for the best scenario (156) and about fifty million tons for the worst (169). The former is about 80% less than the emissions in 2015 according to the (ARE/BfS, 2017) data, and the latter is about 55%. Three active variables are pivotal to reducing mobility emissions according to the modeling results; occupancy rate, home office, online shopping behavior, and local regulations. Other active variables such as taxation on fuel, climate change awareness, or public transportation evolution are also supporting to reduce emissions. For instance, the scenario with the lowest CO₂ emissions estimation assumes high taxation on fuel, a high climate change awareness, and a substantial improvement in the comfort of public transportation. Nevertheless, the effect of these variables on total emissions is limited compared with other active variables such as local regulations or occupancy rate.

The results from the micro-level analysis revealed that improving the comfort of public transportation and closing territorial areas in city centers can effectively shift individuals' behavior from private cars to public transportation. These policies have been designed for private car user segments, and the results show that these segments are more reactive to them, as ex-ante envisaged. It has been discussed in the previous chapters that users can be grouped according to their characteristics and behavior, "tailored" policies can be formulated at

the segment level, and models incorporating heterogeneity can capture the responses of segments towards that policies. The micro-level analysis has realized that claim by formulating policies targeting specific segments (e.g., private car user segments) and testing the responses of agents in different segments to these policies with an economic model consisting of heterogeneous agents.

Chapter 6

Conclusion

6.1 General Conclusion

This chapter summarizes the overall contributions of the thesis. It discusses how the stated research questions are answered throughout the thesis and how these answers can potentially improve the existing practice and extend the body of knowledge in different lines of research in economics.

Chapter 2 carries out an empirical analysis to explore the origin of heterogeneity in mobility consumption behavior. The chapter answers the particular research question, "What characteristics increase (or decrease) the likelihood of being habit-driven and cost-sensitive mobility user?". The chapter filters the Swiss individuals in the MTMC data who stated their mobility consumption reasons as habit and cost, respectively. Thus, two distinct consumer segments with different mobility consumption reasons (motives) are obtained. The chapter analyzes the variables (e.g., socio-economic and demographic) whether they increase or decrease the likelihood of being a habit-driven and cost-sensitive mobility user. The analysis enables us to explore the structural heterogeneity behind the heterogeneous mobility behavior motives. As for the methodology, multiple logistic regression was carried out. The results show that the likelihood

of being a habit-driven private car user decreases with the increasing traveled distance (kilometers) and increases in rural areas. Besides, being in the second income level increases the likelihood of choosing a private car out of habit. With the increasing age, the likelihood of being a habit-driven public transportation user increases. The overall findings show that there are two behavioral segments with different mobility consumption motives, and there is structural heterogeneity between them. Nevertheless, the chapter does not quantize behavioral heterogeneity, e.g., how many such behavioral segments do we have? It only observes and shows it. Therefore, the following chapter will analyze the quantity of the observed behavioral heterogeneity, as will be discussed. The overall chapter contributes to the economic literature on mobility consumption behavior by presenting empirical findings.

Chapter 3 conducts one of the first extensive segmentation research on the demand side of the Swiss mobility sector. It seeks an answer to a particular research question, "What are the segments of Swiss households based on their mobility consumption behavior and characteristics?". The chapter develops and presents a new structured market segmentation procedure, which hinges on multivariate quantitative method cluster analysis, to answer the research question. It applies the segmentation procedure with the MTMC data to find Swiss household segments regarding mobility consumption behavior and characteristics. It presents thirteen different Swiss mobility consumer segments, which are internally relatively homogeneous in terms of mobility behavior and characteristics. The segments quantify the observed behavioral heterogeneity in the previous chapter. The overall chapter contributes to the economic literature on market segmentation and the decarbonization of the Swiss mobility sector by presenting empirical findings and providing methodological improvements.

Chapter 4 answers the research question, "How can an economic model with heterogeneous agents that model mobility demand in Switzerland be built and validated?". The chapter uses the findings of the two previous chapters as input. It builds and validates an ACE demand model that captures behavioral and structural heterogeneity among agents. Agents can have structural heterogeneity such as income level, budget constraints, endowments, place of living, as well as behavioral heterogeneity that we observed in the previous chapter. Regarding the latter, agents' utility functions (i.e., decision-making mechanisms) are configured and calibrated according to their segments. To be more concrete, the weights used for agents' utility functions, as detailed explained in Chapter 4, are set according to agents' segment information. So, each segment has its own utility function, which allows the model to incorporate behavioral heterogeneity. The model simulates individual mobility behavior as core output. That can be aggregated to find macro patterns such as total mobility demand and CO₂ emissions.

Finally, Chapter 5 provides a study with the use case of the built model for socio-economic analysis. The study involves macro and micro-level analyses. As for the macro-level analysis, the chapter answers the question, "How to use an economic model to estimate future decarbonization and mobility demand trajectories in Switzerland?". The chapter first determines different dimensions whose changes are relevant to the evolution of the Swiss mobility sector. Different trajectories for the dimensions are determined, whose combinations yield a scenario space. The model evaluated 640 different scenarios for the future of the Swiss mobility sector between the years 2015-2050. It simulated for each scenario individual mobility behavior that is aggregated to find macro outputs such as mobility demand and CO₂ emissions. The results from the macro-level

analysis show that mobility demand in terms of personal distance traveled will rise in Switzerland in all the scenarios in 2050 relative to today's demand. However, the scenarios that assume a strong access ban for private cars, a higher increase in online shopping and home office behavior lead to a lower demand increase. The model estimates significant CO₂ emission reductions for the scenarios in the range of 50-80% compared to today's emissions. Even the scenarios that assume today's values for the trajectories of active variables (i.e., no evolution in the active variables) lead to about 50% of emission reductions due to the estimated efficiency improvements and penetration of EVs on the supply side. These reductions increase and become about 80% for the scenarios that assume a high increase in the home office and online shopping behavior, a significant improvement in public transportation, a high tax on fuel, a high climate change awareness, strong access bans for private cars, and implementation of behavioral policies. As for the micro-level analysis, the chapter answers the question, "How to formulate segment-level decarbonization policies and test them with an economic model?". The chapter formulates two policies targeting private car segments specifically. The policies are public transportation evolution that aims to gradually improve the comfort in public transportation modes until 2050 and local regulations that aim to ban territorial areas in city centers to private cars. The model computed the micro-behavior of agents with and without policies and compared whether the policies affected the targeted segments as ex-ante envisaged. The results show that as a result of the implementation of both policies, the agents from private car segments changed their behavior primarily. The micro-level also examines the responses of the agents to the changes in their economic environment. The reason is that the previous chapters discussed that the

Swiss households are heterogeneous in many dimensions, such as their mobility behavior, characteristics, endowments, and constraints. They can respond to economic policies differently due to the heterogeneity. The chapter compared the modeling results of two scenarios that differ only in terms of the implementation of the price related policies (e.g., tax on fuel and mobility price per traveled kilometer). The chapter analyzed whether the same agents change their behavior for the same trips due to the price increase. The results demonstrate that the responses of agents are heterogeneous. Few agents (about 4%) changed their consumption behavior; they used public transportation and soft mobility modes instead of the private car. The rest kept their existing behavior. The analysis showed us that agents in the model are heterogeneous in many dimensions, so they can respond to the changes in their environment differently.

6.2 Final Remarks - Policy Discussion

The thesis provides readers with empirical and methodological insights. It follows a research narrative, first observing and analyzing heterogeneity, then incorporating the heterogeneity in economic modeling, and finally using the model for socio-economic analysis. The narrative has been built by chapters and their connections. Each chapter provides readers with methodological novelties such as the market segmentation procedure, the ACE model, and the empirical validation method. I believe that the readers can profit from these novelties, use them in their research, and even improve them further.

The thesis also signals the policy-making side by providing the results from the scenario analysis. The macro-level analysis shows that mobility emissions

will not be zero in 2050 with given assumptions for the trajectories of the scenario variables. In order to achieve direct zero emissions from the mobility sector, further policies are needed, such as phasing out the sales of cars with fossil fuels. If further such policies are not possible to implement for some reason, other parts (sectors) of the Swiss energy system could have negative emissions, which can counterbalance mobility emissions to achieve the net zero emission target. The micro-level analysis shows that the policies of public transportation evolution and local regulations can largely affect private car user segments. That is, these policies lead to behavior change among private car users. They lead to a demand shift from private cars to primarily public transportation. The micro-level analysis also revealed that agents' demand is relatively inelastic to price, whose reasons have been discussed in Chapter 4. Increasing tax on fuel could lead to demand shifts from private cars to other modes. However, it is somewhat limited compared to other variables or policies used in the analysis.

6.3 Future Work

The thesis answered all the research questions formulated in Chapter 1. So, the research in the scope of this thesis is completed. Nonetheless, the presented results and methodologies can be expanded in future work depending on the financing of it. For instance, the built model can be further developed by adding new modules. In the current version of the model, agents choose mobility modes for their mobility activities. That is, the decisions are short-term. A long-term decision-making mechanism can be added to the model in future work. Thus, agents can make long-term investments that can also influence their short-term decisions. For instance, agents can buy a car or seasonal ticket that can affect

their short-term decisions like choosing mobility mode for today's trips. Thus, the fleet evolution, which is exogenous to the model at the current stage, can be captured and produced by the model. So, a part of the supply side (e.g., personal cars) that is currently exogenous to the model can also be modeled. Besides, the model can be used in further studies concerning the decarbonization of the Swiss mobility sector depending on its financing again. It has been built and used in the scope of the thesis. The number of use cases would be increased in future works so that we can get feedback from other researchers and improve the model accordingly.

Bibliography

- Abel, Andrew B. (1990). "Asset Prices under Habit Formation and Catching up with the Joneses". In: *The American Economic Review* 80.2, pp. 38–42. ISSN: 00028282.
- Ajzen, Icek et al. (1991). "The theory of planned behavior". In: *Organizational behavior and human decision processes* 50.2, pp. 179–211.
- Alexandre, Michel and Gilberto Tadeu Lima (2020). "Combining monetary policy and prudential regulation: an agent-based modeling approach". In: *Journal of Economic Interaction and Coordination* 15.2, pp. 385–411.
- Alonso-González, María J, Sascha Hoogendoorn-Lanser, Niels van Oort, Oded Cats, and Serge Hoogendoorn (2020). "Drivers and barriers in adopting Mobility as a Service (MaaS)—A latent class cluster analysis of attitudes". In: *Transportation Research Part A: Policy and Practice* 132, pp. 378–401.
- Anable, Jillian (2005). "'Complacent car addicts' or 'aspiring environmentalists'? Identifying travel behaviour segments using attitude theory". In: *Transport policy* 12.1, pp. 65–78.
- Antonelli, Cristiano (2011). *Handbook on the economic complexity of technological change*. Edward Elgar Publishing.
- ARE/BfS (2017). "Verkehrsverhalten der Bevölkerung Ergebnisse des Mikrozensus Mobilität und Verkehr 2015". In: *Federal Office for Spatial Development and Swiss Federal Statistical Office*.

- Arthur, W Brian (2006). "Out-of-equilibrium economics and agent-based modeling". In: *Handbook of computational economics* 2, pp. 1551–1564.
- Axelrod, Robert (1997). "Advancing the art of simulation in the social sciences". In: *Simulating social phenomena*. Springer, pp. 21–40.
- Axsen, Jonn, Joseph Bailey, and Marisol Andrea Castro (2015). "Preference and lifestyle heterogeneity among potential plug-in electric vehicle buyers". In: *Energy Economics* 50, pp. 190–201.
- BAFU (2019). *Treibhausgasinventar der Schweiz*. <https://www.bafu.admin.ch/bafu/de/home/themen/klima/zustand/daten/treibhausgasinventar.html>. (Accessed on 11/18/2021).
- Balci, Osman (1994). "Validation, verification, and testing techniques throughout the life cycle of a simulation study". In: *Annals of operations research* 53.1, pp. 121–173.
- Bamberg, Sebastian, Icek Ajzen, and Peter Schmidt (2003). "Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action". In: *Basic and applied social psychology* 25.3, pp. 175–187.
- Barde, Sylvain (2016). "Direct comparison of agent-based models of herding in financial markets". In: *Journal of Economic Dynamics and Control* 73, pp. 329–353.
- (2017). "A practical, accurate, information criterion for nth order markov processes". In: *Computational Economics* 50.2, pp. 281–324.
- (2020). "Macroeconomic simulation comparison with a multivariate extension of the Markov Information Criterion". In: *Journal of Economic Dynamics and Control* 111, p. 103795.

- Barde, Sylvain and Sander Van Der Hoog (2017). "An empirical validation protocol for large-scale agent-based models". In: *Bielefeld Working Papers in Economics and Management*.
- Barnes, William, Myles Gartland, and Martin Stack (2004). "Old habits die hard: path dependency and behavioral lock-in". In: *Journal of Economic Issues* 38.2, pp. 371–377.
- Baumgartner, Anne, Uros Tomic, Iljana Schubert, Annika Sohre, Corinne Moser, and Paul Burger (2019). "What determines mode choice in leisure travel? An analysis of financial and non-financial measures aiming at a less car-based form of leisure mobility". In: *Swiss Mobility Conference*.
- Becker, Felix and Kay W Axhausen (2018). "Predicting the use of automated vehicles for Zurich, Switzerland". In: *15th International Conference on Travel Behavior Research (IATBR 2018)*.
- BedDeM (2020). *GitHub - SiLab-group/beddem_simulator*. https://github.com/SiLab-group/beddem_simulator. (Accessed on 07/16/2020).
- Beisbart, Claus and Nicole J Saam (2019). *Computer simulation validation*. Springer.
- Bektas, Alperen, Khoa Nguyen, Valentino Piana, and René Schumann (2018). "People-centric policies for decarbonization: Testing psycho-socio-economic approaches by an agent-based model of heterogeneous mobility demand". In: *Computing in Economics and Finance (CEF) Conference*.
- Bektas, Alperen, Valentino Piana, and René Schumann (2021). "A meso-level empirical validation approach for agent-based computational economic models drawing on micro-data: a use case with a mobility mode-choice model". In: *Springer Nature Business & Economics* 1.6.

- Bektas, Alperen and René Schumann (2019). “How to Optimize Gower Distance Weights for the k-Medoids Clustering Algorithm to Obtain Mobility Profiles of the Swiss Population”. In: *2019 6th Swiss Conference on Data Science (SDS)*. ISBN: 978-1-7281-3105-4. IEEE, pp. 51–56.
- Berger, Thomas (2001). “Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis”. In: *Agricultural economics* 25.2-3, pp. 245–260.
- Bernhardt, K (2007). “Agent-based modeling in transportation”. In: *Artificial Intelligence in Transportation* 72.
- BFE (2021). *CO₂-Emissionsvorschriften für neue Personen- und Lieferwagen*. <https://www.bfe.admin.ch/bfe/de/home/effizienz/mobilitaet/co2-emissionsvorschriften-fuer-neue-personen-und-lieferwagen.html>. (Accessed on 11/18/2021).
- Bianchi, Carlo, Pasquale Cirillo, Mauro Gallegati, and Pietro A Vagliasindi (2007). “Validating and calibrating agent-based models: a case study”. In: *Computational Economics* 30.3, pp. 245–264.
- Boldrin, Michele and Fabio Canova (2001). “Inequality and convergence in Europe’s regions: reconsidering European regional policies”. In: *Economic policy* 16.32, pp. 206–253.
- Boughaci, Dalila, Abdullah AK Alkhaldeh, Jamil J Jaber, and Nawaf Hamadneh (2021). “Classification with segmentation for credit scoring and bankruptcy prediction”. In: *Empirical Economics* 61.3, pp. 1281–1309.
- Boulouchos, Konstantinos, Christian Bach, Christian Bauer, Dominik Bucher, Davide Cerruti, Amin Dehdarian, Massimo Filippini, Maximilian Held, Stefan Hirschberg, and Ramachandran Kannan (2021). *Pathways to a net zero CO₂ Swiss mobility system: SCCER Mobility Whitepaper*. Tech. rep. ETH Zurich.

- Brette, Olivier, Nathalie Lazaric, and Victor Vieira da Silva (2017). "Habit, decision-making, and rationality: comparing Thorstein Veblen and early Herbert Simon". In: *Journal of economic issues* 51.3, pp. 567–587.
- Browning, Martin, Jesus Carro, et al. (2007). "Heterogeneity and microeconomics modelling". In: *Econometric Society Monographs* 43, p. 47.
- Bundesrat, Der (2021). *Climate protection: Federal Council adopts Switzerland's long-term climate strategy*. <https://www.bafu.admin.ch/bafu/en/home/topics/climate/news-releases.msg-id-82140.html>. (Accessed on 11/18/2021).
- Caiani, Alessandro, Antoine Godin, Eugenio Caverzasi, Mauro Gallegati, Stephen Kinsella, and Joseph E Stiglitz (2016). "Agent based-stock flow consistent macroeconomics: Towards a benchmark model". In: *Journal of Economic Dynamics and Control* 69, pp. 375–408.
- Campbell, John Y and John H Cochrane (1999). "By force of habit: A consumption-based explanation of aggregate stock market behavior". In: *Journal of political Economy* 107.2, pp. 205–251.
- Campello, Ricardo JGB and Eduardo R Hruschka (2006). "A fuzzy extension of the silhouette width criterion for cluster analysis". In: *Fuzzy Sets and Systems* 157.21, pp. 2858–2875.
- Cao, Jing, Xiaoyue Cathy Liu, Yinhai Wang, and Qingquan Li (2013). "Accessibility impacts of China's high-speed rail network". In: *Journal of Transport Geography* 28, pp. 12–21.
- Carattini, Stefano, Garth Heutel, and Givi Melkadze (2021). *Climate policy, financial frictions, and transition risk*. Tech. rep. National Bureau of Economic Research.

- Carroll, Christopher D, Jody Overland, and David N Weil (2000). "Saving and growth with habit formation". In: *American Economic Review* 90.3, pp. 341–355.
- Carse, Andrew, Anna Goodman, Roger L Mackett, Jenna Panter, and David Ogilvie (2013). "The factors influencing car use in a cycle-friendly city: the case of Cambridge". In: *Journal of transport geography* 28, pp. 67–74.
- Cellina, Francesca, Anna Förster, Davide Rivola, Luca Pampuri, Roman Rudel, and Andrea Emilio Rizzoli (2013). "Using smartphones to profile mobility patterns in a living lab for the transition to e-mobility". In: *International Symposium on Environmental Software Systems*. Springer, pp. 154–163.
- Chan-Lau, Mr Jorge A (2017). *ABBA: An agent-based model of the banking system*. International Monetary Fund.
- Chang, Man Kit (1998). "Predicting unethical behavior: a comparison of the theory of reasoned action and the theory of planned behavior". In: *Journal of business ethics* 17.16, pp. 1825–1834.
- Chang, Myong-Hun (2015). *A computational model of industry dynamics*. Routledge.
- Chen, Ching-Fu and Wei-Hsiang Chao (2011). "Habitual or reasoned? Using the theory of planned behavior, technology acceptance model, and habit to examine switching intentions toward public transit". In: *Transportation research part F: traffic psychology and behaviour* 14.2, pp. 128–137.
- Chen, Faan, Jiaorong Wu, Xiaohong Chen, and Jianjun Wang (2017). "Vehicle kilometers traveled reduction impacts of transit-oriented development: evidence from Shanghai City". In: *Transportation Research Part D: Transport and Environment* 55, pp. 227–245.

- Cherchi, Elisabetta and Cinzia Cirillo (2014). "Understanding variability, habit and the effect of long period activity plan in modal choices: a day to day, week to week analysis on panel data". In: *Transportation* 41.6, pp. 1245–1262.
- Christiano, Lawrence J, Martin S Eichenbaum, and Mathias Trabandt (2018). "On DSGE models". In: *Journal of Economic Perspectives* 32.3, pp. 113–40.
- Cogliano, Jonathan F and Xiao Jiang (2016). "Agent-based computational economics: simulation tools for heterodox research". In: *Handbook of Research Methods and Applications in Heterodox Economics*. Edward Elgar Publishing.
- Colander, David, Peter Howitt, Alan Kirman, Axel Leijonhufvud, and Perry Mehrling (2008). "Beyond DSGE models: toward an empirically based macroeconomics". In: *American Economic Review* 98.2, pp. 236–40.
- Collingridge, Dave S (2013). "A primer on quantitized data analysis and permutation testing". In: *Journal of Mixed Methods Research* 7.1, pp. 81–97.
- Cox, Brian (2018). "Mobility and the energy transition: a life cycle assessment of swiss passenger transport technologies including developments until 2050". PhD thesis. ETH Zurich.
- Crawford, Ian (2010). "Habits revealed". In: *The Review of Economic Studies* 77.4, pp. 1382–1402.
- Dawid, Herbert (2006). "Agent-based models of innovation and technological change". In: *Handbook of computational economics* 2, pp. 1235–1272.
- Dawid, Herbert, Simon Gemkow, Philipp Harting, Sander Van der Hoog, and Michael Neugart (2012). "The eurace@ unibi model: An agent-based macroeconomic model for economic policy analysis". In: *Bielefeld working papers in economics and management*.
- Dawid, Herbert, Simon Gemkow, Philipp Harting, Sander van der Hoog, and Michael Neugart (2014). "Agent-based macroeconomic modeling and policy

- analysis: The eurace@ unibi model". In: *Bielefeld Working Papers in Economics and Management*.
- Dawid, Herbert and Andreas Pyka (2018). "Introduction: Special Issue on Evolutionary Dynamics and Agent-Based Modeling in Economics". In: *Computational Economics* 52.3, pp. 707–710.
- De Ona, Juan, Rocio de Oña, Laura Eboli, and Gabriella Mazzulla (2015). "Heterogeneity in perceptions of service quality among groups of railway passengers". In: *International Journal of Sustainable Transportation* 9.8, pp. 612–626.
- Delli Gatti, Domenico and Severin Reissl (2020). *ABC: An Agent Based Exploration of the Macroeconomic Effects of Covid-19*.
- Di, Xuan and Henry X Liu (2016). "Boundedly rational route choice behavior: A review of models and methodologies". In: *Transportation Research Part B: Methodological* 85, pp. 142–179.
- Diana, Marco and Cristina Pronello (2010). "Traveler segmentation strategy with nominal variables through correspondence analysis". In: *Transport Policy* 17.3, pp. 183–190.
- Disdier, Anne-Célia and Frank van Tongeren (2010). "Non-tariff measures in agri-food trade: What do the data tell us? Evidence from a cluster analysis on OECD imports". In: *Applied Economic Perspectives and Policy* 32.3, pp. 436–455.
- Dosi, Giovanni, Giorgio Fagiolo, Mauro Napoletano, Andrea Roventini, and Tania Treibich (2015). "Fiscal and monetary policies in complex evolving economies". In: *Journal of Economic Dynamics and Control* 52, pp. 166–189.
- Duffy, John (2006). *Chapter 19 agent-based models and human subject experiments. Volume 2 of Handbook of Computational Economics*.

- Dynan, Karen E (2000). "Habit formation in consumer preferences: Evidence from panel data". In: *American Economic Review* 90.3, pp. 391–406.
- Edenhofer, O, R Pichs-Madruga, Y Sokona, E Farahani, S Kadner, K Seyboth, A Adler, I Baum, S Brunner, P Eickemeier, et al. (2014). "Summary for policy-makers climate change 2014, mitigation of climate change". In: *IPCC 2014, Climate Change 2014: Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.
- EEA (2015). *Evaluating 15 years of transport and environmental policy integration*. Tech. rep. doi:10.2800/214970. ISSN 1977-8449, ISBN 978-92-9213-713-7. European Environment Agency.
- Elhorst, J Paul, Jean-Loup Madre, and Alain Pirotte (2020). "Car traffic, habit persistence, cross-sectional dependence, and spatial heterogeneity: new insights using French departmental data". In: *Transportation Research Part A: Policy and Practice* 132, pp. 614–632.
- Epstein, Joshua M (1999). "Agent-based computational models and generative social science". In: *Complexity* 4.5, pp. 41–60.
- Eriksson, Karin Hederos, Randi Hjalmarsson, Matthew J Lindquist, and Anna Sandberg (2016). "The importance of family background and neighborhood effects as determinants of crime". In: *Journal of Population Economics* 29.1, pp. 219–262.
- Evans, George W and Seppo Honkapohja (2005). "An interview with thomas j. sargent". In: *Macroeconomic Dynamics* 9.4, pp. 561–583.
- Fagiolo, Giorgio, Mattia Guerini, Francesco Lamperti, Alessio Moneta, and Andrea Roventini (2019). "Validation of agent-based models in economics and finance". In: *Computer Simulation Validation*. Springer, pp. 763–787.

- Fagiolo, Giorgio, Alessio Moneta, and Paul Windrum (2007). "A critical guide to empirical validation of agent-based models in economics: Methodologies, procedures, and open problems". In: *Computational Economics* 30.3, pp. 195–226.
- Fagiolo, Giorgio and Andrea Roventini (2012). "Macroeconomic policy in DSGE and agent-based models". In: *Revue de l'OFCE* 5, pp. 67–116.
- Faraway, Julian James (2002). *Practical regression and ANOVA using R*. Vol. 168. University of Bath Bath.
- Farmer, J Doyne and Duncan Foley (2009). "The economy needs agent-based modelling". In: *Nature* 460.7256, p. 685.
- Farrar, Donald E and Robert R Glauber (1967). "Multicollinearity in regression analysis: the problem revisited". In: *The Review of Economic and Statistics*, pp. 92–107.
- Ferraro, Paul J and Juan José Miranda (2013). "Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment". In: *Resource and Energy Economics* 35.3, pp. 356–379.
- Franke, Reiner (2009). "Applying the method of simulated moments to estimate a small agent-based asset pricing model". In: *Journal of Empirical Finance* 16.5, pp. 804–815.
- Friedrichsmeier, Thomas, Ellen Matthies, and Christian A Klöckner (2013). "Explaining stability in travel mode choice: An empirical comparison of two concepts of habit". In: *Transportation research part F: traffic psychology and behaviour* 16, pp. 1–13.
- Gallegati, Mauro, Antonio Palestrini, and Alberto Russo (2017). *Introduction to agent-based economics*. Academic Press.

- Gardner, Benjamin (2009). "Modelling motivation and habit in stable travel mode contexts". In: *Transportation Research Part F: Traffic Psychology and Behaviour* 12.1, pp. 68–76.
- Gärling, Tommy and Kay W Axhausen (2003). "Introduction: Habitual travel choice". In: *Transportation* 30.1, pp. 1–11.
- Garmendia, Maddi, Cecilia Ribalaygua, and José María Ureña (2012). "High speed rail: implication for cities". In: *Cities* 29, S26–S31.
- Gatti, Domenico Delli, Edoardo Gaffeo, Mauro Gallegati, and Antonio Palestrini (2005). "The apprentice wizard: Monetary policy, complexity and learning". In: *New Mathematics and Natural Computation* 1.01, pp. 109–128.
- Geanakoplos, John, Robert Axtell, J Doyne Farmer, Peter Howitt, Benjamin Conlee, Jonathan Goldstein, Matthew Hendrey, Nathan M Palmer, and Chun-Yi Yang (2012). "Getting at systemic risk via an agent-based model of the housing market". In: *American Economic Review* 102.3, pp. 53–58.
- Gerhard, Patrick, Joe J Gladstone, and Arvid OI Hoffmann (2018). "Psychological characteristics and household savings behavior: The importance of accounting for latent heterogeneity". In: *Journal of Economic Behavior & Organization* 148, pp. 66–82.
- Giri, Federico (2017). "Heterogeneity in Macroeconomics: DSGE and Agent-Based Model Approach". In: *Introduction to Agent-Based Economics*. Elsevier, pp. 87–103.
- Gower, John C (1971). "A general coefficient of similarity and some of its properties". In: *Biometrics*, pp. 857–871.
- Graham, Michael H (2003). "Confronting multicollinearity in ecological multiple regression". In: *Ecology* 84.11, pp. 2809–2815.

- Grazzini, Jakob and Matteo Richiardi (2015). "Estimation of ergodic agent-based models by simulated minimum distance". In: *Journal of Economic Dynamics and Control* 51, pp. 148–165.
- Gregory, Allan W. and Gregor W. Smith (1990). "Calibration as estimation". In: *Econometric Reviews* 9.1, pp. 57–89. DOI: [10.1080/07474939008800178](https://doi.org/10.1080/07474939008800178). eprint: <https://doi.org/10.1080/07474939008800178>. URL: <https://doi.org/10.1080/07474939008800178>.
- Grünwald, Peter D and Abhijit Grunwald (2007). *The minimum description length principle*. MIT press.
- Gsottbauer, Elisabeth and Jeroen CJM Van den Bergh (2011). "Environmental policy theory given bounded rationality and other-regarding preferences". In: *Environmental and Resource Economics* 49.2, pp. 263–304.
- Gualdi, Stanislao, Marco Tarzia, Francesco Zamponi, and Jean-Philippe Bouchaud (2015). "Tipping points in macroeconomic agent-based models". In: *Journal of Economic Dynamics and Control* 50, pp. 29–61.
- Guerini, Mattia and Alessio Moneta (2017). "A method for agent-based models validation". In: *Journal of Economic Dynamics and Control* 82, pp. 125–141.
- Hall, George and John Rust (2003). "Simulated minimum distance estimation of a model of optimal commodity price speculation with endogenously sampled prices". In: *manuscript, Yale University*.
- Hamill, Lynne and G Nigel Gilbert (2016). *Agent-based modelling in economics*. Wiley Online Library.
- Han, Jiawei, Jian Pei, and Micheline Kamber (2011). *Data mining: concepts and techniques*. Elsevier.
- Harrell Jr, Frank E (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

- Hartley, James E and James E Hartley (2002). *The representative agent in macroeconomics*. Routledge.
- Haustein, Sonja (2012). "Mobility behavior of the elderly: an attitude-based segmentation approach for a heterogeneous target group". In: *Transportation* 39.6, pp. 1079–1103.
- Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante (2009). "Quantitative macroeconomics with heterogeneous households". In: *Annu. Rev. Econ.* 1.1, pp. 319–354.
- Heckbert, Scott, Tim Baynes, and Andrew Reeson (2010). "Agent-based modeling in ecological economics". In: *Annals of the New York Academy of Sciences* 1185.1, pp. 39–53.
- Heckman, James J (2000). "Causal parameters and policy analysis in economics: A twentieth century retrospective". In: *The Quarterly Journal of Economics* 115.1, pp. 45–97.
- (2001). "Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture". In: *Journal of political Economy* 109.4, pp. 673–748.
- Hidrue, Michael K, George R Parsons, Willett Kempton, and Meryl P Gardner (2011). "Willingness to pay for electric vehicles and their attributes". In: *Resource and energy economics* 33.3, pp. 686–705.
- Hoffmann, Christin, Charles Abraham, Mathew P White, Susan Ball, and Stephen M Skippon (2017). "What cognitive mechanisms predict travel mode choice? A systematic review with meta-analysis". In: *Transport reviews* 37.5, pp. 631–652.
- Holland, John H (2006). "Studying complex adaptive systems". In: *Journal of systems science and complexity* 19.1, pp. 1–8.

- Hommes, Cars (2021). "Behavioral and experimental macroeconomics and policy analysis: A complex systems approach". In: *Journal of Economic Literature* 59.1, pp. 149–219.
- Hoog, Sander van der (2019). "Surrogate modelling in (and of) agent-based models: A prospectus". In: *Computational Economics* 53.3, pp. 1245–1263.
- Hosmer Jr, David W, Stanley Lemeshow, and Rodney X Sturdivant (2013). *Applied logistic regression*. Vol. 398. John Wiley & Sons, pp. 31–47.
- Huang, Sheng and Anjan V Thakor (2013). "Investor heterogeneity, investor-management disagreement and share repurchases". In: *The Review of Financial Studies* 26.10, pp. 2453–2491.
- Hunecke, Marcel, Sonja Haustein, Susanne Böhler, and Sylvie Grischkat (2010). "Attitude-based target groups to reduce the ecological impact of daily mobility behavior". In: *Environment and behavior* 42.1, pp. 3–43.
- Hurtado, Samuel (2014). "DSGE models and the Lucas critique". In: *Economic Modelling* 44, S12–S19.
- Idris, Ahmed Osman, Khandker M Nurul Habib, and Amer Shalaby (2015). "An investigation on the performances of mode shift models in transit ridership forecasting". In: *Transportation Research Part A: Policy and Practice* 78, pp. 551–565.
- IPCC (2018). *Global Warming of 1.5° C, an IPCC special report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty special report*. Tech. rep.
- Janzen, Maxim and Kay W Axhausen (2018). "Measuring long-distance travel". In: *18th Swiss Transport Research Conference (STRC 2018)*.

- Jenkins, Jesse D. (2014). "Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design?" In: *Energy Policy* 69, pp. 467–477. ISSN: 0301-4215.
- Jermann, Urban J (1998). "Asset pricing in production economies". In: *Journal of monetary Economics* 41.2, pp. 257–275.
- Kassambara, Alboukadel (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning*. Vol. 1.
- Kaufman, Leonard and Peter J Rousseeuw (2009). *Finding groups in data: an introduction to cluster analysis*. Vol. 344. John Wiley & Sons.
- Kirman, Alan (2006). "Heterogeneity in economics". In: *Journal of Economic Interaction and Coordination* 1.1, pp. 89–117.
- Kiuila, Olga (2018). "Decarbonisation perspectives for the Polish economy". In: *Energy Policy* 118, pp. 69–76.
- Kleven, Henrik Jacobsen and Claus Thustrup Kreiner (2006). "The marginal cost of public funds: Hours of work versus labor force participation". In: *Journal of Public Economics* 90.10-11, pp. 1955–1973.
- Klößner, Christian A and Ellen Matthies (2004). "How habits interfere with norm-directed behaviour: A normative decision-making model for travel mode choice". In: *Journal of environmental psychology* 24.3, pp. 319–327.
- Klügl, Franziska (2008). "A validation methodology for agent-based simulations". In: *Proceedings of the 2008 ACM symposium on Applied computing*. ACM, pp. 39–43.
- Krusell, Per and Anthony A Smith Jr (1998). "Income and wealth heterogeneity in the macroeconomy". In: *Journal of political Economy* 106.5, pp. 867–896.

- Kydland, Finn E and Edward C Prescott (1982). "Time to build and aggregate fluctuations". In: *Econometrica: Journal of the Econometric Society*, pp. 1345–1370.
- Lamperti, F, Giovanni Dosi, Mauro Napoletano, Andrea Roventini, and A Sapi (2020). "Climate change and green transitions in an agent-based integrated assessment model". In: *Technological Forecasting and Social Change* 153, p. 119806.
- Lamperti, Francesco (2018a). "An information theoretic criterion for empirical validation of simulation models". In: *Econometrics and Statistics* 5, pp. 83–106.
- (2018b). "Empirical validation of simulated models through the GSL-div: an illustrative application". In: *Journal of Economic Interaction and Coordination* 13.1, pp. 143–171.
- LeBaron, Blake (2012). "Heterogeneous gain learning and the dynamics of asset prices". In: *Journal of Economic Behavior & Organization* 83.3, pp. 424–445.
- Lee, Cheng F, John C Lee, Alice C Lee, et al. (2000). *Statistics for business and financial economics*. Vol. 1. Springer.
- Levine, Paul (2020). "The State of DSGE Modeling". In: *Oxford Research Encyclopedia of Economics and Finance*.
- Li, Dawei, Tomio Miwa, Takayuki Morikawa, and Pan Liu (2016). "Incorporating observed and unobserved heterogeneity in route choice analysis with sampled choice sets". In: *Transportation research part C: emerging technologies* 67, pp. 31–46.
- Li, Shengxiao Alex, Xiaodong Guan, and Donggen Wang (2022). "How do constrained car ownership and car use influence travel and life satisfaction?" In: *Transportation Research Part A: Policy and Practice* 155, pp. 202–218.

- Li, Zhibin, Wei Wang, Chen Yang, and David R Ragland (2013). "Bicycle commuting market analysis using attitudinal market segmentation approach". In: *Transportation Research Part A: Policy and Practice* 47, pp. 56–68.
- Loder, Allister and Kay W Axhausen (2018). "Mobility tools and use". In: *Journal of Transport and Land Use* 11.1, pp. 367–385.
- Lopes, Fernando and Helder Coelho (2018). *Electricity markets with increasing levels of renewable generation: structure, operation, agent-based simulation, and emerging designs*. Vol. 144. Springer.
- Lucas Jr, Robert E (1976). "Econometric policy evaluation: A critique". In: *Carnegie-Rochester conference series on public policy*. Vol. 1. North-Holland, pp. 19–46.
- Lutz, Christoph and Gemma Newlands (2018). "Consumer segmentation within the sharing economy: The case of Airbnb". In: *Journal of Business Research* 88, pp. 187–196.
- Ma, Tai-Yu, Philippe Gerber, Samuel Carpentier, and Sylvain Klein (2015). "Mode choice with latent preference heterogeneity: a case study for employees of the EU institutions in Luxembourg". In: *Transportmetrica A: transport science* 11.5, pp. 441–463.
- Macal, Charles and Michael North (2014). "Introductory tutorial: Agent-based modeling and simulation". In: *Proceedings of the Winter Simulation Conference 2014*. IEEE, pp. 6–20.
- Maréchal, Kevin (2009). "An evolutionary perspective on the economics of energy consumption: the crucial role of habits". In: *Journal of Economic Issues* 43.1, pp. 69–88.
- (2010). "Not irrational but habitual: The importance of "behavioural lock-in" in energy consumption". In: *Ecological Economics* 69.5, pp. 1104–1114.

- Massaro, Domenico (2013). "Heterogeneous expectations in monetary DSGE models". In: *Journal of Economic Dynamics and Control* 37.3, pp. 680–692.
- Maulik, Ujjwal and Sanghamitra Bandyopadhyay (2002). "Performance evaluation of some clustering algorithms and validity indices". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.12, pp. 1650–1654.
- McFadden, Daniel (1989). "A method of simulated moments for estimation of discrete response models without numerical integration". In: *Econometrica: Journal of the Econometric Society*, pp. 995–1026.
- Meghir, Costas and Guglielmo Weber (1996). "Intertemporal nonseparability or borrowing restrictions? A disaggregate analysis using a US consumption panel". In: *Econometrica: Journal of the Econometric Society*, pp. 1151–1181.
- Mifsud, Deborah, Maria Attard, and Stephen Ison (2017). "To drive or to use the bus? An exploratory study of older people in Malta". In: *Journal of transport geography* 64, pp. 23–32.
- Morton, Craig, Jillian Anable, and John D Nelson (2017). "Consumer structure in the emerging market for electric vehicles: Identifying market segments using cluster analysis". In: *International Journal of Sustainable Transportation* 11.6, pp. 443–459.
- Moser, Corinne, Yann Blumer, and Stefanie Lena Hille (2018). "E-bike trials' potential to promote sustained changes in car owners mobility habits". In: *Environmental research letters* 13.4.
- Mosselmans, Bert (2005). "Adolphe Quetelet, the average man and the development of economic methodology". In: *The European Journal of the History of Economic Thought* 12.4, pp. 565–582.
- Murray-Smith, David J (2015). "Testing and validation of computer simulation models". In: *Cham: Springer*. doi 10, pp. 978–3.

- Napoletano, Mauro (2018). "A short walk on the wild side: Agent-based models and their implications for macroeconomic analysis". In: *Revue de l'OFCE* 3, pp. 257–281.
- Nguyen, Khoa and René Schumann (2019). "On developing a more comprehensive decision-making architecture for empirical social research". In: *The 20th International Workshop on Multi-Agent-Based Simulation (MABS 2019)*. 13 May 2019.
- O'Brien, Margaret and Deborah Jones (1999). "Children, parental employment and educational attainment: an English case study". In: *Cambridge Journal of Economics* 23.5, pp. 599–621.
- Osborne, Matthew (2005). "Consumer Learning, Habit Formation and Heterogeneity: A Structural Examination". In: *Manuscript, Univ. Chicago*.
- Page, Daniel R (2012). "Generalized algorithm for restricted weak composition generation". In: *Journal of Mathematical Modelling and Algorithms in Operations Research* 12.4, pp. 345–372.
- Park, Hae-Sang and Chi-Hyuck Jun (2009). "A simple and fast algorithm for K-medoids clustering". In: *Expert systems with applications* 36.2, pp. 3336–3341.
- Parker, Jonathan A, Nicholas S Souleles, David S Johnson, and Robert McClelland (2013). "Consumer spending and the economic stimulus payments of 2008". In: *American Economic Review* 103.6, pp. 2530–53.
- Pennings, Joost ME, Philip Garcia, and Scott H Irwin (2005). *Heterogeneity in Economic Behavior*.
- Petrovic, Marko, Bulent Ozel, Andrea Teglio, Marco Raberto, Silvano Cincotti, et al. (2017). *Eurace Open: An agent-based multi-country model*. Tech. rep.
- Piana, Valentino, Alperen Bektas, and Nguyen Khoa (2020). "Temporal morphogenesis". In: *Economics Web Institute EWI Essay series*.

- Piana, Valentino, Andrea Del Duce, and René Schumann (2019). *Spanning the space of scenarios for future mobility in Switzerland*. Deliverable D.2.2a. SCCER CREST Joint-Activity “Evolution of Mobility: A Socio-Economic Analysis”.
- Piana, Valentino, Beat Hintermann, Paul Burger, Mehdi Farsi, Andrea Del Duce, Kannan Ramachandran, Ali Darudi, Hannes Weigt, and René Schumann (2020). *JA Mobility: Deliverable DX.1 Scenarios and Pathways for the Future Swiss Mobility and Energy Perspective*. Deliverable. SCCER CREST Joint-Activity “Evolution of Mobility: A Socio-Economic Analysis”.
- Prada, María F and Sergio Urzúa (2017). “One size does not fit all: Multiple dimensions of ability, college attendance, and earnings”. In: *Journal of Labor Economics* 35.4, pp. 953–991.
- Pruna, Radu T, Maria Polukarov, and Nicholas R Jennings (2020). “Loss aversion in an agent-based asset pricing model”. In: *Quantitative Finance* 20.2, pp. 275–290.
- Pyka, Andreas and Giorgio Fagiolo (2007). “29 Agent-based modelling: a methodology for neo-Schumpeterian economics”. In: *Elgar companion to neo-schumpeterian economics* 467.
- Queiroz, Mariza Motta, Carlos Roque, and Filipe Moura (2020). “Shifting from private to public transport using duration-based modeling of a school-based intervention”. In: *Transportation research record* 2674.7, pp. 540–554.
- Ramachandran, Kannan and Hal Turton (2014). *Switzerland Energy Transition Scenarios—Development and Application of the Swiss TIMES Energy System Model (STEM)*.
- Reichert, Alexander and Christian Holz-Rau (2015). “Mode use in long-distance travel”. In: *Journal of Transport and Land Use* 8.2, pp. 87–105.

- Ringler, Philipp, Dogan Keles, and Wolf Fichtner (2016). "Agent-based modelling and simulation of smart electricity grids and markets—a literature review". In: *Renewable and Sustainable Energy Reviews* 57, pp. 205–215.
- Ritchie, Hannah and Max Roser (2020). *CO₂ and Greenhouse Gas Emissions*. <https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions>.
- Roberto, Christina A and Ichiro Kawachi (2015). *Behavioral economics and public health*. Oxford University Press.
- Rousseeuw, Peter J (1987). "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis". In: *Journal of computational and applied mathematics* 20, pp. 53–65.
- Royer, Heather, Mark Stehr, and Justin Sydnor (2015). "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company". In: *American Economic Journal: Applied Economics* 7.3, pp. 51–84.
- Rozen, Kareen (2010). "Foundations of intrinsic habit formation". In: *Econometrica* 78.4, pp. 1341–1373.
- Russell, Stuart and Peter Norvig (2002). "Artificial intelligence: a modern approach (3rd Edition)". In: pp. 694–695.
- Saunders, John A (1980). "Cluster analysis for market segmentation". In: *European Journal of marketing*.
- Sbordone, Argia M, Andrea Tambalotti, Krishna Rao, and Kieran James Walsh (2010). "Policy analysis using DSGE models: an introduction". In: *Economic policy review* 16.2.
- Schlecht, Ingmar and Hannes Weigt (2014). "Swissmod-a model of the Swiss electricity market". In: *Available at SSRN 2446807*.

- Schubert, Erich and Peter J Rousseeuw (2019). "Faster k-medoids clustering: improving the PAM, CLARA, and CLARANS algorithms". In: *International conference on similarity search and applications*. Springer, pp. 171–187.
- Schubert, Iljana, Sylvain Weber, Adan L Martinez-Cruz, Paul Burger, and Mehdi Farsi (2022). "Structural Equation Modeling as a Route to Inform Sustainable Policies: The Case of Private Transportation". In: *Frontiers in Sustainability* 3, p. 837427.
- Schumann, René, Hannes Weigt, and Tom Kober (2020). *Report on coherent and quantified scenarios and pathways, including contribution of behavior change to energy use, CO2 emissions and welfare*. Deliverable D.2.2 and D.2.3. SCCER CREST.
- Schwanen, Tim, David Banister, and Jillian Anable (2012). "Rethinking habits and their role in behaviour change: the case of low-carbon mobility". In: *Journal of Transport Geography* 24, pp. 522–532.
- Sharmeen, Fariya and Harry Timmermans (2014). "Walking down the habitual lane: analyzing path dependence effects of mode choice for social trips". In: *Journal of Transport Geography* 39, pp. 222–227.
- Simon, Herbert A (1972). "Theories of bounded rationality". In: *Decision and organization* 1.1, pp. 161–176.
- Şimşekoğlu, Özlem, Trond Nordfjærn, and Torbjørn Rundmo (2015). "The role of attitudes, transport priorities, and car use habit for travel mode use and intentions to use public transportation in an urban Norwegian public". In: *Transport Policy* 42, pp. 113–120.
- Soto, Jose J, Victor Cantillo, and Julián Arellana (2021). "Market segmentation for incentivising sustainable transport policies". In: *Transportation Research Part D: Transport and Environment* 99.

- Stern, Nicholas (2007). *The economics of climate change: the Stern review*. Cambridge University Press.
- Stiglitz, Joseph E (2018). “Where modern macroeconomics went wrong”. In: *Oxford Review of Economic Policy* 34.1-2, pp. 70–106.
- Sütterlin, Bernadette, Thomas A Brunner, and Michael Siegrist (2011). “Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics”. In: *Energy Policy* 39.12, pp. 8137–8152.
- Tesfatsion, Leigh (2018). “Electric power markets in transition: Agent-based modeling tools for transactive energy support”. In: *Handbook of computational economics*. Vol. 4. Elsevier, pp. 715–766.
- (2021). *Agent-Based Computational Economics: Overview and Brief History*. ISU General Staff Papers.
- Tesfatsion, Leigh and Kenneth L Judd (2006). *Handbook of computational economics: agent-based computational economics*. Vol. 2. Elsevier.
- Thaler, Richard H and LJ Ganser (2015). *Misbehaving: The making of behavioral economics*. WW Norton New York.
- Triandis, Harry C (1979). “Values, attitudes, and interpersonal behavior.” In: *Nebraska symposium on motivation*. University of Nebraska Press.
- Trichet, Jean-Claude (2010). *Reflections on the nature of monetary policy non-standard measures and finance theory*. <https://www.ecb.europa.eu/press/key/date/2010/html/sp101118.en.html>. (Accessed on 11/11/2021).
- Turrell, Arthur (2016). “Agent-based models: understanding the economy from the bottom up”. In: *Bank of England Quarterly Bulletin*, Q4.
- Urdan, Timothy C (2016). *Statistics in plain English*. Taylor & Francis.

- Veit, Daniel J, Anke Weidlich, and Jacob A Krafft (2009). "An agent-based analysis of the German electricity market with transmission capacity constraints". In: *Energy Policy* 37.10, pp. 4132–4144.
- Verplanken, Bas, Henk Aarts, Ad Van Knippenberg, and Carina van Knippenberg (1994). "Attitude Versus General Habit: Antecedents of Travel Mode Choice 1". In: *Journal of applied social psychology* 24.4, pp. 285–300.
- Verplanken, Bas, Vemund Myrbakk, and Erik Rudi (2005). "The measurement of habit". In: *The routines of decision making* 231, p. 247.
- Verplanken, Bas and Wendy Wood (2006). "Interventions to break and create consumer habits". In: *Journal of Public Policy & Marketing* 25.1, pp. 90–103.
- W Axhausen, Kay, Andreas Horni, and Kai Nagel (2016). *The multi-agent transport simulation MATSim*. Ubiquity Press.
- Wangsness, Paal Brevik, Stef Proost, and Kenneth Løvold Rødseth (2020). "Vehicle choices and urban transport externalities. Are Norwegian policy makers getting it right?" In: *Transportation Research Part D: Transport and Environment* 86.
- Wedel, Michel and Wagner A Kamakura (2012). *Market segmentation: Conceptual and methodological foundations*. Vol. 8. Springer Science & Business Media.
- Weidlich, Anke and Daniel Veit (2008). "A critical survey of agent-based wholesale electricity market models". In: *Energy Economics* 30.4, pp. 1728–1759.
- Wen, Wei-Jen and Tsung-Chi Cheng (2013). "Performing arts attendance in Taiwan: who and how often?" In: *Journal of Cultural Economics* 37.2, pp. 309–325.
- Wickens, Michael (2012). *Macroeconomic theory: a dynamic general equilibrium approach*. Princeton University Press.

- Windrum, Paul, Giorgio Fagiolo, and Alessio Moneta (2007). "Empirical validation of agent-based models: Alternatives and prospects". In: *Journal of Artificial Societies and Social Simulation* 10.2, p. 8.
- Xiang, Xiaorong, Ryan Kennedy, Gregory Madey, and Steve Cabaniss (2005). "Verification and validation of agent-based scientific simulation models". In: *Agent-directed simulation conference*. Vol. 47, p. 55.
- Yakovlev, Evgeny (2018). "Demand for alcohol consumption in Russia and its implication for mortality". In: *American Economic Journal: Applied Economics* 10.1, pp. 106–49.
- Zhang, Hongru and Yang Yang (2019). "Prescribing for the tourism-induced Dutch disease: A DSGE analysis of subsidy policies". In: *Tourism Economics* 25.6, pp. 942–963.
- Zhao, Wei, Yi Lu, and Genfu Feng (2019). "How many agents are rational in China's economy? Evidence from a heterogeneous agent-based new Keynesian model". In: *Computational Economics* 54.2, pp. 575–611.
- Zhou, Fan, Zuduo Zheng, Jake Whitehead, Simon Washington, Robert K Perrons, and Lionel Page (2020). "Preference heterogeneity in mode choice for car-sharing and shared automated vehicles". In: *Transportation Research Part A: Policy and Practice* 132, pp. 633–650.