UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

THE SPATIAL DISTRIBUTION OF INCOME IN SWITZERLAND

The role of natural amenities, the persistence of inequality and the spatial mismatch between income and production

PhD thesis submitted to the Faculty of Economics and Business

Institute of Economic Research

University of Neuchâtel

For the PhD degree in Economics

by

Joséphine LEUBA

Approved by the dissertation committee:

Prof. Jean-Marie GRETHER, University of Neuchâtel, thesis director

Prof. Claude JEANRENAUD, University of Neuchâtel

Prof. Raphaël PARCHET, University of Lugano

Dr. Jörg SCHLÄPFER, Wüest Partner

Defended on November 26th, 2019

Faculté des sciences économiques
Avenue du 1er-Mars 26
CH-2000 Neuchâtel
www.unine.ch/seco

UNIVERSITÉ DE NEUCHÂTEL

IMPRIMATUR POUR LA THÈSE

The spatial distribution of income in Switzerland

The role of natural amenities, the persistence of inequality and the spatial mismatch between income and production

Joséphine LEUBA

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES ÉCONOMIQUES

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

Prof. Jean-Marie Grether (directeur de thèse, Université de Neuchâtel)
Prof. emeritus Claude Jeanrenaud (président du jury, Université de Neuchâtel)
Prof. Raphaël Parchet (Università della Svizzera italiana)
Dr. Jörg Schläpfer (Head of macroeconomics, Wüest Partner)

Autorise l'impression de la présente thèse.

Neuchâtel, le 10 février 2020

Annik Dubied

La doyenne
Annik Dubied

Résumé

Cette thèse contient trois chapitres portant sur la distribution spatiale du revenu en Suisse. D'abord, nous explorons la relation entre les attributs naturels et les inégalités. Ensuite, nous étudions la persistance des disparités entre les communes suisses entre 1971 et 2015. Finalement, nous explorons le décalage entre la localisation géographique du revenu et celle de la production. Nous trouvons que l'environnement naturel est un déterminant clé des inégalités de revenu entre les communes, mais aussi en leur sein. L'importance des lacs et de l'altitude est dominante. Dans le deuxième chapitre, nous montrons que les disparités spatiales entre les communes sont persistantes depuis le début des années septante, et qu'elles sont même en augmentation dans la dernière décennie. Nous simulons la distribution du revenu en 2040, qui se trouve être très similaire à celle de 2014/15. Le troisième chapitre s'occupe du différentiel entre la distribution du revenu et celle de la valeur ajoutée. Nous trouvons que les ménages riches ne se trouvent pas nécessairement dans les mêmes communes que les entreprises les plus productives. Une spécialisation s'opère au niveau communal. Nous détectons en particulier des "ceintures résidentielles" autours des centres productifs. Toutes les méthodes empiriques sont empruntées à la littérature récente sur l'analyse spatiale exploratoire, sur l'économétrie spatiale de panel et sur la convergence des revenus. Nous utilisons abondamment les sources de données mises à disposition par l'administration fédérale suisse au niveau communal. L'essentiel de nos résultats est que l'inégalité spatiale des revenus et l'inadéquation géographique entre production et revenu sont des phénomènes importants et persistants. Ils devraient être pris en compte dans le débat public, notamment en termes de développement régional, d'utilisation du territoire, de compétition fiscale, de péréquation financière et d'infrastructure de transport.

Mots-clés: distribution spatiale du revenu, attributs naturels, tri sélectif des revenus, persistance, inégalité, décalage spatial, communes suisses

Abstract

This thesis consists of three chapters on the spatial distribution of income in Switzerland. First, we explore the relationship between natural amenities and inequalities. Then, we study the persistence of spatial inequality between Swiss municipalities between 1971 and 2015. Finally, we explore the mismatch between the geographical location of income and of value-added. We find that the natural environment is a key determinant of income inequalities between, but also within municipalities. The importance of lakes and altitude is prevalent. In the second chapter, we show that spatial disparities between Swiss municipalities are persistent since the early seventies and even increasing in the last decade. We simulate the distribution of 2040, which turns out to be very similar to the one of 2014/15. The third chapter deals with the differential between income and value-added distribution. We find that the richest households do not necessarily locate in the same municipalities as the most productive firms. Some specialisation takes place at the municipal level. In particular, we detect residential belts around productive centres. All empirical methods are borrowed from the recent literature on exploratory spatial analysis, spatial panel econometrics and income convergence. We make an extensive use of recent data sources made available by the Swiss Federal administration at the level of municipalities. The bottom line is that spatial income inequality and mismatch between production and income are important and persistent phenomena. They should be accounted for in the public debate, notably in terms of regional development, land use, fiscal competition, financial equalisation and transport infrastructures.

Keywords: spatial income distribution, natural amenities, income sorting, persistence, inequality, spatial mismatch, Swiss municipalities

Table of contents

List of Figures

List of Tables

Remerciements

Je remercie bien sûr en premier lieu mon directeur de thèse Jean-Marie Grether pour son soutien, sa bienveillance et ses nombreux conseils. Ce travail n'aurait simplement pas débuté sans lui, je n'aurais jamais autant apprécié ces années de recherche si n'importe qui d'autre avait été à sa place. Je lui dois également de belles expériences sur le terrain de l'enseignement, qui grâce à sa confiance ont été extrêmement formatrices. En deuxième lieu, je suis redevable à Claude Jeanrenaud, qui a accepté de présider mon jury de thèse, dont les conseils de lecture et l'expertise sur l'économie suisse m'ont été infiniment utiles. J'exprime aussi ma gratitude à Jörg Schläpfer et Raphaël Parchet pour leur intérêt, leurs critiques constructives et leur humanité. Ma reconnaissance s'adresse ensuite à Nicole Mathys et Brahim Boualam, qui m'ont accordé leur temps avec beaucoup de générosité, et dont les commentaires contribuent pour beaucoup à la cohérence de mon propos. Je suis reconnaissante également envers Manuel Lehner et la firme Fahrländer Partner, qui ont partagé leurs données avec beaucoup de gentillesse. Merci aussi à tous ceux et toutes celles qui ont rendu mon quotidien agréable à l'Institut de Recherches économiques, dont mes collègues doctorant-e-s Sandra, Géraud, Ivan, Cécile, Ghislaine, Jeremy, Laurent, Juliette, Evert, Nathan, Teo et Niccolò. Je remercie en particulier Luciano, Alexandra et Martin pour m'avoir rendu les tâches d'assistanat plus faciles lorsque partagées. Enfin, merci à Benjamin, non seulement pour sa productivité lors de la rédaction de notre chapitre commun, mais aussi pour son esprit et sa bonne humeur de tous les jours. Pour finir, comme bien sûr mes préoccupations doctorales ne sont pas restées enfermées dans l'enceinte de l'Université, j'adresse ici ma reconnaissance à mes proches. À Camille, Mélodie, Tania et Virginie, pour leur bon sens et leur optimisme lorsqu'il a fallu parler de ma lassitude. Leur présence a été importante et continuera de l'être. À Evin et Sofiane, car quand il s'agit de trouver de l'aide, la vie a parfois plus d'imagination que nous. À mes cousins, cousines, tontons, tatas, belle-soeur et beau-frère: merci pour ce si bel entourage. À mes parents, grande sœur, petit frère et grands-parents: merci pour ce soutien sans faille et cette intelligence. On ne choisit pas sa famille, mais pour moi qu'importe: j'aurais choisi précisément celle-là.

Acknowledgements

First of all, of course, I would like to thank my thesis director Jean-Marie Grether for his support, his kindness and his many advices. This work simply wouldn't have started without him, I would never have appreciated these years of research so much if anyone else had been in his place. I also owe him great experiences in the field of teaching, which thanks to his confidence have been extremely formative. Secondly, I am indebted to Claude Jeanrenaud, who has accepted to chair my thesis jury, whose reading advice and expertise on the Swiss economy have been infinitely useful to me. Thank you to Jörg Schläpfer and Raphaël Parchet, for their interest, their constructive critics and the humanity. My thanks then go to Nicole Mathys and Brahim Boualam, who generously gave me their time, and whose comments contribute greatly to the coherence of my work. I am also grateful to Manuel Lehner and to the firm Fahrländer Partner, who shared their data with great kindness. I would also like to express my gratitude to all those who have made my daily life at the Institute of economic research pleasant, among them my doctoral colleagues Sandra, Géraud, Ivan, Cécile, Ghislaine, Jeremy, Laurent, Juliette, Evert, Nathan, Teo and Niccolò. A special thanks to Luciano, Alexandra and Martin for having made assistant tasks easier when shared. Finally, thank you to Benjamin, not only for his productivity in writing our common chapter, but also for the wit he showed every day. To conclude, since my doctoral concerns have not remained locked up in the University, I would like to express my gratitude to my family and friends. To Camille, Mélodie, Tania and Virginie, for their good sense and optimism when it came to talking about my fatigue. Their presence has been and will continue to be important. To Evin and Sofiane, because when it comes to finding help, life sometimes has more imagination than we do. To my cousins, uncles, aunts, brother and sister-in-law: thank you for such a beautiful entourage. To my parents, sister, brother and grandparents: thank you for your unfailing support and intelligence. No one chooses their family, but it does not matter in my case: I would have chosen precisely that one.

General Introduction

People and economic activity are not evenly distributed across space, which has important consequences in terms of transport, land use and social equity. Documenting and understanding the forces behind income dispersion or concentration is therefore crucial to design and implement sustainable development policies.

Models of location choices and income sorting (e.g. Tiebout (1956), Oates (1969), Roback (1982)) address the question from an individual point of view: rich people wish to live where taxes are low and where they enjoy cultural, natural and urban amenities (considered as luxury goods, see Brueckner et al. (1999)). In contrast, poor people prefer areas where fiscal policies are highly redistributive and where housing costs are low. In a free-market context, rental amounts, local wages and commodities prices should adjust to reach a spatial equilibrium where all people are happy to live where they live.

However, the assumption of a frictionless world is debatable, since households face mobility barriers and local complementarities in consumption. Moreover, spatial externalities matter: the location choice of one household affects other agents, which distorts the equilibrium and may call for government intervention.

Collectively, welfare losses due to suboptimal sorting may be particularly persistent since the spatial income structure is largely dependent of its own past: differences in fiscal regimes, housing market and job opportunities reinforce the income sorting process. In addition to economic factors, the psychosocial tendency of households to locate near their peers ("homophily", according to Lazarsfeld et al. (1954)) exacerbates the social divergence between geographical areas. Network effects and preferences for homogeneity of neighbourhoods strengthen the persistence of a spatially uneven distribution of income. The over-concentration of wealth may lead to several undesirable economic consequences such as over-commuting and congestion costs, undervaluation of environmental amenities, overuse of land and urban sprawl. Social consequences also matter. The geographical concentration of poverty has dramatic effects on the local government's ability to spend money. This leads to poor access to school, health and safety services in these areas (Cytron et al., 2011).

Switzerland is an interesting case to explore these issues, for a variety of reasons. It is one of the richest country in the world, but with strong income differences between cantons, and between municipalities of the same canton. Cantons and municipalities also differ in terms of language and public policies, in particular taxation.

While the impact of taxation on income sorting in Switzerland has been widely explored, this is much less the case for the role of natural amenities and the long run forces that drive the spatial income distribution. The landscape diversity is also remarkable given the small size of the territory. Altitude goes from 193m (Laggo Maggiore, TI) to 4634m (Pointe Dufour, VS). With numerous lakes, river, plains and forests, we can find a wide variety of natural amenities in Switzerland, which may be a divergence force on income if they are concentrated, or a convergence force if they are spread. Finally, distances are short. It is relatively easy to move from one municipality - or even from a canton - to another, especially if they are in the same language region. Low migration costs should typically lead to a more sorted equilibrium.

We will explore three dimensions. Income sorting is driven by many factors (such as tax rates and housing prices), but one of them is exogenous and primary: the natural environment. When we look at a Swiss map representing income by municipalities, we inevitably notice the importance of lakes and mountains. Besides, we study the time dimension, that is the convergence among Swiss municipalities and the persistence of spatial disparities. There are indeed few changes between the income map in 1971/72 and the one in 2014/15. Finally, we ask how income location relates to value-added location. While many economic geography models assume that people live where they work, this is not the case for Switzerland where distances are short and commuting costs low. This leads to a "spatial mismatch" between the location of income and the location of production.

The central interrogation that guides the present thesis is: where does income concentrate and why? The question breaks down into three parts corresponding to the three chapters of the thesis. First, how does the presence of natural amenities affect the income distribution within and between municipalities? Second, how spatial income inequalities relate to their own past and what will be the most likely income distribution in the future if recent trends are extrapolated? Third, what is the extent and what are the patterns of the spatial mismatch between income and value-added across Swiss municipalities?

The first chapter describes how the spatial distribution of income is related to natural amenities in Switzerland between 2003 and 2015. We explore how income inequalities between and within Swiss municipalities relate to the presence of lakes, rivers, mountains, accessibility and green amenities. It turns out that a larger variety of landscape is associated with larger income differentials between municipalities. We also show that income inequalities within municipalities are more pronounced at the extremes of the amenities ladder that is either in a particularly nice (e.g. next to a lake) or a particularly unwelcoming environment (e.g. in the mountains).

In the second chapter, we explore the persistence of income disparities between Swiss municipalities from 1971/72 to 2014/15. The relative standard deviation of income and the Theil index have increased in the recent past. In our regressions, we detect positive spatial autocorrelation and reject the Solow hypothesis of absolute convergence at the communal level. Instead, a permanent and specific effect pushes each municipality on its own growth path. We also simulate the future distribution of income based on an autoregressive model that includes spatial autocorrelation and individual heterogeneity. We find little prospect for a reduction of inequality between Swiss municipalities during the next 25 years.

In the third chapter, written in collaboration with Benjamin Tissot-Daguette, we explore the spatial distribution of value-added and of income between Swiss municipalities over the 2011-2015 period. Theil index decompositions indicate that value-added disparities are more pronounced than income ones and that spatial dispersion comes mainly from within-cantons variability rather than inter-cantonal inequality. By analysing the distribution of the ratio between value-added and income, and using Moran scatterplots to capture neighbourhood effects, we characterise a Swiss landscape where "productive" centres are surrounded by "residential" belts.

All in all, our results imply that the wealth of one area crucially depends on initial conditions and that convergence between municipalities is not a natural process. Instead, we observe a process of specialisation, during the course of which income and value-added concentrate, but not necessarily at the same place. Thanks to low commuting costs, rich households are able to settle in rich residential areas, near productive centres. Fiscal decentralisation and adjustment on the housing markets tend to reinforce the income sorting process, making spatial disparities increasing.

Two major policy implications follow. First, in terms of redistribution, financial equalisation mechanisms will probably be needed for many years to come, especially if fiscal differentials tend to increase (as predicted by the Tiebout model). Second, in terms of regional development and land use policies, a delicate balance must be found between equity and efficiency concerns. In the former case, the target should be to improve the variables entering the "initial conditions" of the poorest municipalities (e.g. accessibility, air quality or green spaces). In the latter case, the goal is to maximise benefits from specialisation and agglomeration economies (e.g. urban densification), while minimising the resulting congestion and pollution costs.

Chapter 1

Natural amenities and the spatial distribution of Swiss income *

1.1 Introduction

In the recent past, income inequalities have been increasing in developed countries, generating growing concerns among economists as to the policy impacts of this trend (e.g. Stiglitz (2012) or Bourguignon (2018)). In this context, fully characterising income disparities becomes central. In this chapter, we are interested in the spatial dimension of income distribution. Some places are richer than others and these spatial disparities generate a number of key policy questions. The literature on selective migration and income sorting have well established that rich households concentrate in nice places where housing rents are high and where tax rates are low. However, the sorting process is not perfect, since important social inequalities remain within small areas.

In this chapter, we focus on one particular characteristic of a region: natural amenities. Taxes and housing prices are largely endogenous, whereas the natural environment may be considered as part of the initial conditions. Exploring the relationship between natural amenities and inequality across and within living places is therefore particularly interesting.

Switzerland is well suited to explore this issue. Its federalist structure and its fractioning into small entities makes it an ideal case to study spatial income distribution. While the impact of taxation on income sorting has been widely explored, this is much less the case for the link with natural amenities. Within this small country, moving from one municipality to another

*. We thank Brahim Boualam for his precious comments and suggestions on this chapter.

is relatively easy, especially if one stays in the same language region. Lower internal migration costs should theoretically lead to a more sorted equilibrium. Moreover, the landscape diversity is remarkable, so that we can exploit the variety of mountains, lakes, rivers, plains and forests.

Because of the income sorting process, we obviously expect a positive correlation between the environmental endowment and the average income. The relationship between natural amenities and inequalities within municipalities is less clear. We briefly present the implications of the imperfect sorting model of Fretz et al. (2017), which is a general spatial equilibrium model that provides the theoretical intuition behind our results.

In the empirical part, we conduct panel analyses on Swiss data at the municipal level between 2003 and 2015. We estimate a combination of the spatial Durbin and the Mundlak model, with average income and the Gini index as dependent variables. We also run seemingly unrelated regressions to explore the link between natural amenities and the share of taxpayers in five different income brackets.

We find that income inequalities between and within municipalities are positively related to the disparity of natural endowments. For instance, the shorter the distance to a lake, the higher the average income *and* the Gini index in the municipality. These results are not only driven by the right tail of the income distribution.

The chapter is constructed as follows. We first give an overview of the existing literature (section 1.2). We next turn to the presentation of our data (section 1.3) and the econometric identification strategy (section 1.4). We finally discuss the results (section 1.5) and conclude (section 1.7).

1.2 Literature review

Selective migration and income sorting

Why do people live where they live? Economists usually answer with a standard maximisation program in which agents compare their situation in different regions and decide to move if the utility differential overcomes the cost of migration. This utility differential has been soon approximated by the wage differential. In this respect, the first paper to be cited is the one by Sjaastad (1962). The author presents a model in which migration is an investment in human capital: workers move where they get the highest return on their skills, net of migration costs.

Spatial income sorting arises essentially because regions reward differently the skills of workers, who self-select themselves into the best location. Borjas (1987) formalises this idea, which is an application of the classical Roy model (Roy, 1951) to migration.

In a federalist framework, income sorting is also explained by the redistribution system. Tiebout (1956) develops a model in which each jurisdiction provides a different combination for financing and providing public goods. When the taxation is progressive and the public goods considered as inferior, the Tiebout model predicts that rich households concentrate in areas where taxes and redistribution are low.

Many other variables can be suspected to drive household location decision at the individual level.[1] None of them seems neither definitely preponderant nor completely exogenous. Maybe the best summary of what microeconomics has to say about spatial income sorting is the model of Roback (1982): households and firms select mutually exclusive areas on the basis of expected local wages, rents and consumption or amenities. If local wages and prices can freely adjust, a general perfect sorting equilibrium arises where nobody wishes to move. However, location choices have undoubtedly a collective dimension. Going beyond the individual level provides important complements.

Spatial distribution of income and the role of amenities

The New Economic Geography (NEG) has developed as a very fertile framework to explore the collective dimension of location choices. According to the seminal paper of Krugman (1991), we should look at households and firms together within a general equilibrium model. If people follow job opportunities, it becomes central to know where firms prefer to settle. The central point of the NEG framework is the hypothesis of agglomeration economies: because of Marshallian externalities, firms move where other firms are already located. The article by Behrens and Robert-Nicoud (2014) is a perfect recent example. The authors present a model which combines natural advantage, agglomeration economies and firm selection to explain why both productivity and inequality increase with the size of a city.

The so-called supply-side approach looks at the exogenous factors behind location decisions

1. For example life-cycle dimensions (Mincer, 1978), housing price (Oates (1969) and Helpman (1998)), travel time to work, population density, school quality, distance to supermarket (Kim et al., 2005), accessibility (Zondag and Pieters, 2005), air quality (Banzhaf and Walsh, 2008), housing attributes and business location (Schirmer et al., 2013).

of firms and households: amenities. In a central publication, Brueckner et al. (1999) present the amenity-based theory, which aims to explain why Paris is richer than Detroit. The authors distinguish between natural, historical and modern amenities.

About “Natural amenities” [2], Cheshire et al. (2003) highlight the role of water and Marcouiller et al. (2004) explore the relationship between natural amenities and income distribution through the channel of tourism. More recently, Sinha and Cropper (2015) study climate amenities and Schaeffer et al. (2016) distinguish the roles of “blue” and “green” amenities. In general, researchers conclude that natural amenities are a main driver of income sorting and segregation. The water-based amenities have a preponderant impact. However, little is known about the role of natural amenities in influencing income distributions *within* small entities such as municipalities. As far as we know, Lee and Lin (2017) are the only ones to treat this question. They study the spatial distribution of income in the long run and show that variation in natural endowments within American cities is determinant to explain the persistence of disparities between neighbourhoods. The income heterogeneity is therefore more pronounced and the spatial distribution is less fluctuating in Los Angeles (coastal and hilly) than in Dallas (flat and naturally homogeneous).

Switzerland

Switzerland is a federal state with substantial variation in tax rates and very small costs of migration. This country is therefore a natural laboratory to study the effect of tax on income sorting. This is evidenced by the SNF project “The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization”. [3] The project led by Marius Brülhart, Monika Bütler, Mario Jametti and Kurt Schmidheiny was funded by the SNF from 2010 to 2016. [4] The impact of taxation on migration and income sorting in Switzerland has also been explored by other authors. Liebig and Sousa-Poza (2006) explore the individual responsiveness of tax variations and Schmidheiny (2006) uses data from the metropolitan area

2. About the two other types of amenities, see Koster et al. (2014), van Duijn and Rouwendal (2015) and Falck et al. (2015) who explore the role of “Historical amenities” like monuments, conservation areas and historical sites. “Modern amenities” such as arts production, culture and urban facilities are the focus of Throsby (1994), Boualam (2014) and Albouy (2016). The main conclusion is that the income elasticity of demand for historical and modern amenities is larger than one, which is reinforcing the income sorting process.

3. http://fiscalfederalism.ch for further information.

4. In the list of publications, we find Eeckhout et al. (2014) on spatial sorting and Brülhart et al. (2015) on tax competition models. See Schmidheiny (2017) for more details.

of Basel to estimate the impact of income on residence choice probabilities. Morger (2017) finds that the capitalisation [5] of lower tax rates into higher housing rents is not full. Moreover, the degree of capitalisation varies depending on income level (approximated by the quality of the apartments). Basten et al. (2017) take advantage of the language frontier to implement a boundary discontinuity design and estimate the effect of tax rates on housing rents. Using the same strategy, Eugster and Parchet (2019) quantify tax competition thanks to the cultural difference between French-speaking and German-speaking Swiss municipalities. They find that income sorting is reduced because tax rate choices are constrained by tax competition. Kübler and Rochat (2017) and Feld et al. (2018) study the role of the Swiss tax system on overall income inequalities. In a nutshell, these articles show that rich people are more likely to choose low tax jurisdictions, that the tax differentials are partly capitalised into land values and that tax decentralisation tends to lower pre-tax income inequality and to increase post-tax inequality.

To the best of our knowledge, amenities have not been the focus of any study on income sorting in Switzerland. Portnov et al. (2011) explore the role of accessibility [6] on population growth, but do not consider income-related questions. The hedonic pricing literature inspects the explanatory power of natural amenities on housing prices. For instance, Baranzini and Schaerer (2011) estimate the value of lake view in the region of Geneva. They find that the rent of dwellings with lake view is 57% higher than those without. Waltert et al. (2011) explore the role of amenities on regional development, in terms of population and employment. None of these does explore the implications of their findings on the spatial distribution of income.

We conclude this literature review by mentioning the contributions of Segessemann and Crevoisier (2016), who develop the notion of the "residential economy". Economists usually focus on production to explain growth. In this logic, wealth is concentrated in environments that are favourable to the implementation of firms. What is less often included in spatial equilibrium models is the following process: the simple fact that people live where they live calls for additional economic activity. Goods and services such as retails, childcare, housecleaning, gardening or haircuts have to be provided near residents. Therefore, in some regions, economic development is mainly driven by the fact that households - and in particular rich households - live

5. Capitalisation is the mechanism through which the price increases on the housing market in response to a decrease of the income tax rate. See (Oates, 1969).

6. Accessibility is measured as distances to the closest major city, river, border and road.

there. Authors call these areas "residential economies". Compared to "productive economies", they tend to be richer in income terms, but also more unequal.

Our contribution differs from the ones cited above, first because we study the effect of natural amenities not only on inequalities across areas, but also within them. Second, evidence are based on very small entities that are heterogeneous in terms of natural environment, urbanisation, size and wealth. Lastly, the present study takes advantage of recently available data derived from tax returns of Swiss individual taxpayers, aggregated at the municipal (communal) level. These data have three advantages. First, they are reliable. Compared to surveys, tax returns do not have problems of non-response and inaccurate self-reporting. The only limitation concerns tax evasion. Second, data are disaggregated down to a small geographical unit, which is central to study spatial issues. Third, they are based on taxable income, which corresponds better to the earnings of residents than gross product.

Theoretical model

Our empirical analysis is closely related to the theoretical spatial equilibrium model of Fretz et al. (2017), which predicts imperfect sorting. This spatial equilibrium model includes trade and commuting costs, heterogeneity of skills and idiosyncratic preferences: workers must choose the location of residence (i) and the workplace (j). They are characterised by a certain endowment of skills ($\ell \in [0, 1]$) and preferences over the consumption of final goods (C) and housing services (H). Finals goods are aggregated in C assuming a constant elasticity of substitution. The utility function is scaled by the component B_{ji}, which captures the *joint* level of amenities in workplace and residence. The natural amenities in the residence place i therefore enter B_{ji}, as well as the utility cost of commuting between i and j. Some income sorting occurs because each location is endowed with a fixed amount of land and because preferences are idiosyncratic (at the equilibrium, workers spend therefore a fraction $1 - \alpha$ on housing). Firms are in a Dixit-Stiglitz monopolistic competition situation, with a cost function specific to each location. Trade of goods among locations is possible but costly. In this framework, the model predicts that:

> " Desirable residential locations attract a large number of residents [...], and disproportionately the high earners among them, both leading to high housing prices, which then act as a stabilizing push factor." (Fretz et al. (2017), p.9).

Note that a decrease in commuting costs acts as a sorting facilitator. In addition, the idea of "residential economies" of Segessemann and Crevoisier (2016) can be expressed through the fact that local services are characterised by high trade costs. Thus, attractive places with a lot of rich residents are also attractive places for workers, including low-skilled. Provided that commuting costs are positive, this is an pulling force for low-skilled residents as well. Hence, in the context of Swiss municipalities, these mechanisms could explain why we find higher income inequality either within nice natural environments or in remote areas.

1.3 Data

The dataset has been constructed from five sources. Fiscal data come from the Federal Tax Administration (see AFC (2013)). They are calculated on the basis of tax returns for the Federal Direct Tax (FDT). The statistics are available from 2003 to 2015.[7] These precious data have been gathered within the SNF project "The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization".[8]

Additional control variables are available on the Federal Statistical Office (FSO) website. They come from four surveys: the Population and Households Statistics (STATPOP), the Buildings and Dwellings statistic (StatBL), the Swiss Film and Cinema Statistics and the Land Use Statistics. Regarding geographical variables, Swisstopo provides the geographical coordinates of the boundaries of the municipalities, lakes and rivers, as well as the altitude at several spot elevations (Geodpoints). Swisstopo uses the LV03 projection system. To capture accessibility, we use the travel time to the closest agglomeration. The Federal Office for Spatial Development (ARE) calculated it in 2011. Finally, the firm Fahrländer Partner AG provides us the rating of the "Exposition", based on the slope and the orientation of each municipality. We have it as is in 2018.

Fiscal data and controls from the FSO are available from 2003 to 2015 and the geographical database contains the municipal boundaries of 2016. One difficulty arises because of the mergers of more than 500 entities between 2003 and 2016. Any drop would induce an important selection bias, because these mergers are non random political choices. To overcome this obstacle, we treat

7. Some cantons are missing before 2003 because of the transition from *prae-* to *postnumerando* taxation system. The cantons of Ticino, Valais and Vaud were the last to implement the reform in 2003.

8. See section 1.2.

the municipalities as if they had always been like in 2016. In other terms, we artificially put together the composing entities of a new municipality in the years preceding the effective merger. This manipulation relies on the exhaustive list of mutations provided by the Federal Office (OFS, 2018b).

Table 1.1 summarises the availability, the source and the definition of the variables. Moreover, the summary statistics for 2003, 2008 and 2015 are available in Appendix 1.8 (tables 1.10 - 1.13).

1.4 Identification Strategy

The Mundlak model

We work with a typical panel containing N observations over T periods, with N larger than T (2294 municipalities, 13 years). In order to avoid the correlation between the independent variables and the time-invariant part of the error term, we use the Mundlak (1978) specification, which is a mixture of random and fixed effect. Wooldridge (2010) [9] and Baltagi (2013) [10] show how to disentangle within and between effects by multiplying the equation by a time-demeaning matrix. The within effect will be the same as in a fixed effect estimation. Facing the trade-off between unbiasedness and efficiency (Debarsy, 2012), we pragmatically choose unbiasedness given our relatively important number of observations.

Practically, we write the model following Allison (2009) and Schunck (2013): instead of estimating the effect of X_{it}, we use its deviation from the mean $X_{it} - \bar{X}_i$. We then add $\bar{X}_i$ to obtain the between effect. This specification is called the hybrid model, estimated under the random-effect assumptions:

$$Y_{it} = \beta_0 + \beta_1(X_{it} - \bar{X}_i) + \beta_2 A_i + \beta_3 \bar{X}_i + \alpha_i + \epsilon_{it} \tag{1.1}$$

where Y_{it} is either the average income or the Gini index in municipality i at time t. The income is expressed in logarithm and the Gini index in logistic transformation. The derivation of this index is detailed in Appendix (section 1.8). X_{it} is a set of time variant controls, such as tax

9. In section 10.5.
10. In chapter 2.

Table 1.1: List of variables

Variable	Years	N	Source	Description
Income	2003 - 2015	2294[a]	Federal Tax Administration	The dataset contains the median and the mean, net and taxable in equivalent terms (adjusted for the number of people in the household). If not specified, we use the mean net equivalent income.
Gini index	2003 - 2015	2294	Federal Tax Administration	The index is calculated according to formula in Appendix 1.8. If not specified, we use the Gini on net equivalent income.
Share of taxpayers in income class j	2003 - 2015	2294	Federal Tax Administration	The 5 classes of annual taxable income are (in thousands of CHF): $0-30$; $30-40$; $40-50$; $50-75$; more than 75.
Tax burden	2003 - 2007 2008 - 2015	804[b] 2294	Federal Tax Administration	Share of cantonal and communal levies on the labour gross income. We consider the tax burden on married couples with children earning an annual taxable income of 50 000 CHF.
Number of taxpayers	2003 - 2015	2294	Federal Tax Administration	Number of households. Special and normal cases are included.
Population	2003 - 2015	2294	Federal Statistical Office (STATPOP)	Total permanent resident population (mid-August).
Share of foreigners	2003 - 2015	2294	Federal Statistical Office (STATPOP)	Ratio between the number of foreign nationals[c] and the permanent resident population (mid-August).
Housing vacation rate	2003 - 2015	2294	Federal Statistical Office (StatBL)	Percentage of flats to sale or rent (1st June) on the total number of dwellings (1st January).
Number of cinemas	2003 - 2015	2294	Federal Statistical Office (Swiss Film and Cinema Statistics)	Number of cinemas in the municipality.
Share of forest	2007	2294	Federal Statistical Office (Land Use Stat.)	Share of the area of the municipality used as forest.[d]
Share of pasture	2007	2294	Federal Statistical Office (Land Use Stat.)	Share of the area of the municipality used as meadows and farm pastures.[d]
Distance to the closest lake	2016	2294	Swisstopo (VECTOR200)	Distance from the centroid of the municipality to the boundary of the closest lake of more than 100 ha.[e]
Distance to the closest river	2016	2294	Swisstopo (VECTOR200)	Distance from the centroid of the municipality to the axis of the closest river (importance classes 1-3).[e]
Altitude	2016	7427[f]	Swisstopo (VECTOR200)	Altitude in meters at a spot elevation.
Visibility of Alpin peaks	2016	2294	Swisstopo (VECTOR200)	A summit (>3500m) is visible if the vector reaching the centroid of the municipality is not obstructed by any other object.[g]
Accessibility	2011	2294	Federal Office for Spatial Development (ARE)	Travel time (min) by individual motorized transport to the closest agglomeration or isolated town (ARE typology).
Exposition	2018	2222	Fahrländer Partner AG	Rating (1-5) of sunshine exposition, calculated from the slope and orientation of the inhabited parts of the municipality. See Fahrländer and Lehner (2014), p.5.

a: 2294 is the number of municipalities existing in 2016. All the municipalities that have merged since 2003 are treated *as if* they had merged in 2003. To obtain the variables of interest in these "pseudo-municipalities", we calculate either the sum (number of taxpayers, population, surfaces) or the average (weighted by population).

b: Only the municipalities with more than 2000 inhabitants (according to the Census 2000) are in the dataset before 2008.

c: According to the FSO definition, a foreign national is "anyone residing in Switzerland at a given time, but who does not have Swiss nationality. The permanent foreign resident population is the reference population in population statistics. It includes all foreign nationals who hold a residence permit for a minimum duration of 12 months or who have resided in Switzerland for 12 months (Permit B/C/L/F or FDFA permit - international civil servants, diplomats and members of their family)".

d: Standard Nomenclature NOAS04: Basic categories and aggregations.

e: Euclidian distances calculated on the basis of X, Y and Z coordinates of polygons (LV03 projection system) in kilometres.

f: We attribute to each municipality the value measured at the closest spot elevation (from its centroid, Euclidian distance).

g: The vector is calculated from the XY coordinates of the municipality and of the mountain peak. The algorithm looks for a point which is in the neighbourhood (1km) of this segment *and* which has a higher elevation than the linear combination of the two extremities. If it does not find any and if the distance is less than 120 km, we can see the peak from the municipality. See Müri Leupp et al. (2011), p.29.

burden and housing vacation rate, but also indicators of urbanisation (population, share of foreigners) and availability of modern and cultural amenities (number of cinemas). A_i contains time-invariant variables of particular interest, among which the distance to lakes and rivers, altitude and travel time to the closest agglomeration. They are exogenous with respect to the dependent variable. α_i is random and ϵ_{it} are independent and identically distributed. β_1 is the within effect (over time). The estimator would be the same in a fixed-effect estimation.[11] The between-effect β_3 gives the role of X_{it} across observations.

In our empirical setting, housing rents and tax rates are not exogenous with respect to income. We approximate the housing price by the vacation rate and make the assumption that the variation in the excess supply is the cause of the variation in prices, not its consequence, as presented in Rosen and Smith (1983). Moreover, the overall housing vacation rate should not be affected by per capita variables in the short run. Regarding tax burden, we take the five closest entities for which the value is known and calculate the average among them. We then use it as a proxy or as an instrument. If anything, the within coefficients are underestimated compared to the real effect of rental costs and taxes. We can live with it, since we are mainly interested in β_2.

Spatial autocorrelation

Whatever happens in a given municipality is likely to happen also in the neighbouring municipalities. If the event is unobserved, the errors terms α_i and ϵ_{it} are no longer independent. Following Elhorst (2014), we write a model that takes the spatial interaction among the dependent and independent variables into account. In matrix notation, we have:

$$Y_t = \beta_0 + \rho \mathbf{W} Y_t + (\mathbf{X_t} - \bar{\mathbf{X}})\beta_1 + \mathbf{A}\beta_2 + \bar{\mathbf{X}}\beta_3 + (\mathbf{W}\mathbf{X_t} - \mathbf{W}\bar{\mathbf{X}})\beta_4 + \mathbf{W}\mathbf{A}\beta_5 + \mathbf{W}\bar{\mathbf{X}}\beta_6 + \alpha + \epsilon_t \quad (1.2)$$

where Y_t is a vector of dimension $N \times 1$, N being the number of municipalities. Y_t represents either the logarithm of income or the logistic transformation of the Gini index. $\mathbf{X_t}$ is the matrix of time-dependent variables of dimension $N \times M$, M being the number of variables. β_1 is the vector of coefficients of $M \times 1$ dimension. A is the matrix of amenities $(N \times K)$, constant over time. $\bar{\mathbf{X}}$ is the time-average of $\mathbf{X_t}$, of dimension $N \times M$. $\mathbf{W}$ is a matrix $N \times N$ row-standardised

11. Except that the FE estimator has a smaller variance.

contiguity matrix. The non standardised matrix contains 1 in w_{kl} if municipalities k and l are contiguous [12], zero otherwise. The contiguity matrix is then normalised such that the sum of each row is equal to 1. The scalar ρ represents the endogenous interaction of the dependent variable, whereas β_4 to β_6 are the exogenous interaction effects among the independent variables. α and ϵ_t are the usual error terms, satisfying the random effect assumptions.

In practice, we estimate two models: the spatial autoregressive model (SAR) that forces β_4, β_5 and β_6 to be zero and the spatial Durbin model (SDM) that does not.

Seemingly unrelated Regressions

The mean income and the Gini index are partial indicators of the distribution of income. In particular, it does not indicate to what extent inequality comes from the top of the distribution. [13] Additional information could be obtained by looking at the density of the distribution. We know the number of taxpayers belonging to five mutually exclusive net income categories, namely 0-30; 30-40; 40-50; 50-75 and more than 75 000 CHF per year. Similarly to Feld and Kirchgässner (2001), we are therefore able to estimate the following system of equations:

$$\begin{pmatrix} S_{1it} \\ S_{2it} \\ S_{3it} \\ S_{4it} \\ S_{5it} \end{pmatrix} = \begin{pmatrix} (X_{it} - \bar{X}_i) & ... & 0 \\ ... & ... & ... \\ 0 & ... & (X_{it} - \bar{X}_i) \end{pmatrix} \cdot \begin{pmatrix} \delta_{11} \\ \delta_{12} \\ \delta_{13} \\ \delta_{14} \\ \delta_{15} \end{pmatrix} + \tag{1.3}$$

$$\begin{pmatrix} A_i & ... & 0 \\ ... & ... & ... \\ 0 & ... & A_i \end{pmatrix} \cdot \begin{pmatrix} \delta_{21} \\ \delta_{22} \\ \delta_{23} \\ \delta_{24} \\ \delta_{25} \end{pmatrix} + \begin{pmatrix} \bar{X}_i & ... & 0 \\ ... & ... & ... \\ 0 & ... & \bar{X}_i \end{pmatrix} \cdot \begin{pmatrix} \delta_{31} \\ \delta_{32} \\ \delta_{33} \\ \delta_{34} \\ \delta_{35} \end{pmatrix} + \begin{pmatrix} \theta_{1i} \\ \theta_{2i} \\ \theta_{3i} \\ \theta_{4i} \\ \theta_{5i} \end{pmatrix} + \begin{pmatrix} \nu_{1it} \\ \nu_{2it} \\ \nu_{3it} \\ \nu_{4it} \\ \nu_{5it} \end{pmatrix}$$

S_{jit} holds for the cube root transformation of the share of taxpayers. [14] We jointly estimate

12. According to the Queen definition, two areas are contiguous if they share a common vertex.

13. The increase of disparities in developed countries is mainly driven by the elongation of the right tail of the distribution. See for instance Piketty and Saez (2006).

14. $S_{jit} = [\frac{Share_{jit}}{1 - Share_{jit}}]^{1/3}$ where $Share_{jit}$ is the fraction of taxpayers in income class j in municipality i at time t. Since some shares are equal to zero, we use this transformation rather than the logistic one.

the five equations by the Zellner-Aitken Seemingly unrelated regressions method (Zellner, 1962). The joint estimation allows the error terms $\theta_{ji} + \nu_{jit}$ to be correlated across equations js.

The last category is unfortunately very wide: households earning more than 75 000 CHF by year represent one third of the total number of taxpayers. Thus, the regression does not tell what happens with the richest people, that are suspected to drive inequalities. Rather, it aims to show whether the relationship with natural amenity is still present in the low and middle income categories.

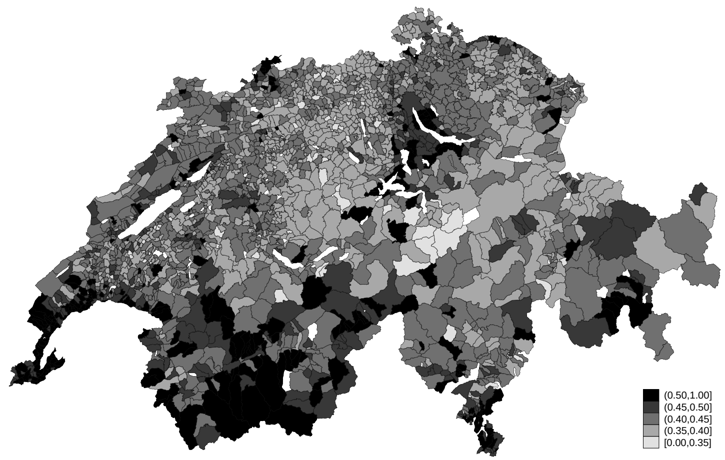
1.5 Results

Descriptive statistics

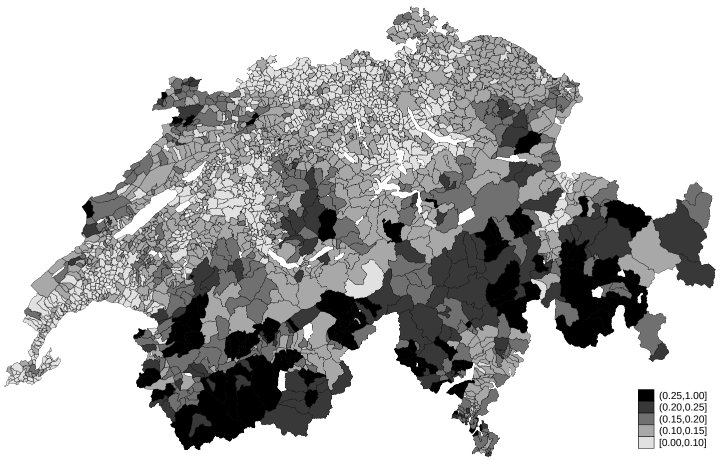
We present here some stylised facts about the distribution of income across and within Swiss municipalities. We focus on the last year available in the dataset, namely 2015. The maps show the average income, the Gini index and the share of taxpayers at the beginning and at the end of the distribution.

Income distribution has a visible link with mountains and lakes. On the maps showing average income and the share of taxpayers earning more than 75 000 CHF, we clearly see dark spots on the Lemanic Arc, around Lakes Lucerne, Lugano and Zurich. In contrast, the Jura and the Alps are lighter. At first sight, within inequalities are not evenly distributed, neither. The Gini index seems larger the closer the lakes and the higher the altitude.

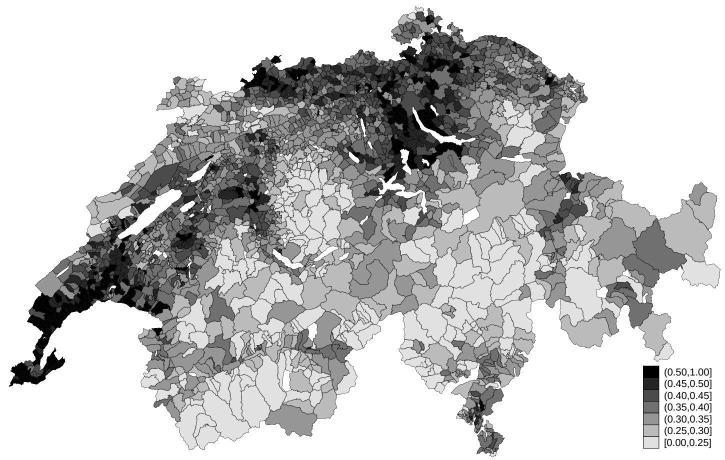
(a) Average annual net income (in thousands of CHF), 2015



(b) Gini Index (on net income), 2015



(c) Share of taxpayers with net income lower than 30 000 CHF, 2015



(d) Share of taxpayers with net income higher than 75 000 CHF, 2015

Figure 1.1: Swiss spatial income distribution - 2015

Turning to summary statistics, Table 1.2 shows the averages of fiscal variables, weighted by the number of taxpayers. The municipalities were split between lakeside and non-lakeside.[15] In lakeside municipalities, the annual income is 17,77% larger and the Gini index is 13,10% larger than in non-lakeside areas. The share of the richest households is larger in municipalities next to water. Note also that the average number of taxpayers is more than twice larger in lakeside areas, because cities have developed next to blue amenities.

Table 1.2: Summary statistics 2015, by distance to lake

	Non-lakeside	Lakeside	T-test	P-value
Average net income	40323.049	47489.481	-7.040	0.000
Median net income	32991.427	33907.404	-5.867	0.000
Number of taxpayers	1796.091	4096.163	-5.519	0.000
Gini index (net)	0.435	0.492	-16.466	0.000
Share under 30 000 CHF	0.125	0.123	1.392	0.164
Share 30-40 000 CHF	0.108	0.101	7.816	0.000
Share 40-50 000 CHF	0.138	0.124	8.397	0.000
Share 50-75 000 CHF	0.255	0.243	7.826	0.000
Share over 75 000 CHF	0.375	0.410	-9.041	0.000
Tax burden (married, 50 000 CHF)	20.057	19.348	8.308	0.000
Number of obs.	1882.000	412.000	.	.

The statistics are weighted by the number of taxpayers, except this variable itself

Table 1.3 shows the information when the sample is divided between low and high altitude. The critical value is set at 610m, the median. The table indicates that average income is lower and Gini index is higher in the second group. The figures on the different shares suggest that the income distribution has more density on the left tail in the mountains. The municipalities at high altitude have a smaller population and charge more taxes than their low-altitude counterparts.

Table 1.3: Summary statistics 2015, by altitude

	Lower than median	Higher than median	T-test	P-value
Average net income	43582.120	41190.728	7.933	0.000
Median net income	34157.835	31797.235	17.624	0.000
Number of taxpayers	2803.657	1613.669	3.706	0.000
Gini index (net)	0.452	0.457	-3.188	0.001
Share under 30 000 CHF	0.112	0.146	-18.318	0.000
Share 30-40 000 CHF	0.102	0.111	-9.243	0.000
Share 40-50 000 CHF	0.131	0.137	-6.545	0.000
Share 50-75 000 CHF	0.254	0.244	4.035	0.000
Share over 75 000 CHF	0.400	0.363	14.923	0.000
Tax burden (married, 50 000 CHF)	19.730	19.980	7.892	0.000
Number of obs.	1148.000	1146.000	.	.

The statistics are weighted by the number of taxpayers, except this variable itself

15. “Lakeside” is defined such that the distance between the centroid of the municipality and the closest border of a lake is smaller than 3km.

Regressions

Table 1.4 presents the models with the logarithm of average annual net income as the dependent variable. Distances to blue amenities are in kilometres and log forms. Altitude is expressed in kilometres to limit the number of digits in the coefficients. The first column shows the fixed effect estimation, the second is the hybrid model with standard errors clustered by municipalities (equation 1.1), the third and the fourth are hybrid models corrected for spatial autocorrelation (equation 1.2). In model (3), β_4, β_5 and β_6 are assumed to be zero (spatial autoregressive model).

A battery of tests have been conducted. We do not detect multicollinearity in the variance inflation factor. The robust Lagrange multiplier test confirms the presence of spatial lag dependence in income and Gini index. A Wald test shows that the null hypothesis $\beta_4 = \beta_5 = \beta_6 = 0$ can be rejected. Our preferred specification is therefore the spatial Durbin model (fourth column). The coefficients on spatially lagged explanatory variables are not reported in the table for readability concerns. Interested readers find them in the Appendix (table 1.14). Table 1.5 shows the same models estimated with the logistic transformation of the Gini index as the dependent variable, in which the average income is included as a control.

As expected, income is positively correlated with all the attractive natural amenities (lake, river, proximity to an agglomeration [16]). It has a negative link with altitude and forest. The former may be considered as a natural dis-amenity (because of climate and slope for instance), the latter as an indicator for rurality. The Gini index is higher in lakeside municipalities, in the mountains and in areas with a large share of green amenities. The closer the next urban centre, the lower the indicator of inequality.

In line with Behrens et al. (2014), the most populated municipalities are also the richest one (between coefficient). Regarding the Gini index, the between coefficient of population is negative. It indicates that income is more evenly distributed in large municipalities than in small villages. This confirms the observation of Castells-Quintana and Royuela (2015), according to which "In already developed or unequal countries, i.e., countries in which concentration of resources is already high, urbanisation in small and medium-sized cities appears to be associated with decreasing inequality." (p.306). Income and share of foreigners are positively linked across

16. The distance variables (lake, river, travel time) must be reversed in the interpretation.

Table 1.4: Dependent variable: Log(income)

	(1) OLS	(2) FE	(3) Hybrid model	(4) Spatial AR	(5) Spatial Durbin
Log(distance to the closest lake - km)	-0.0345***	0	-0.0329***	-0.0298***	-0.0466***
	(0.00126)	(.)	(0.00469)	(0.00432)	(0.00536)
Log(distance to the closest river - km)	-0.0578***	0	-0.0587***	-0.0400***	-0.0360***
	(0.00186)	(.)	(0.00701)	(0.00636)	(0.00908)
Altitude - km	-0.122***	0	-0.123***	-0.0875***	-0.0821***
	(0.00246)	(.)	(0.00806)	(0.00750)	(0.00794)
Share of forest	-0.176***	0	-0.173***	-0.110***	-0.0751***
	(0.00816)	(.)	(0.0297)	(0.0277)	(0.0275)
Share of pasture	-0.0673***	0	-0.0731**	-0.0899***	-0.0539
	(0.0108)	(.)	(0.0326)	(0.0315)	(0.0405)
Travel time to the closest agglomeration (min)	-0.00331***	0	-0.00312***	-0.00213***	-0.00291***
	(0.0000912)	(.)	(0.000327)	(0.000287)	(0.000366)
Log(population)	0.0110***				
	(0.00130)				
Share of foreign population	0.370***				
	(0.0184)				
Housing vacation rate (%)	-0.0202***				
	(0.000913)				
Tax burden in the 5 closest municipalities	-0.0279***				
	(0.000554)				
Number of cinemas	-0.0260***				
	(0.00183)				
Log(population) (within)		0.187***	0.187***	0.0971***	0.0285
		(0.00906)	(0.0208)	(0.0176)	(0.0188)
Share of foreign population (within)		0.958***	0.958***	0.562***	0.299***
		(0.0272)	(0.0779)	(0.0718)	(0.0705)
Housing vacation rate (within)		0.00108**	0.00108*	0.000144	-0.000349
		(0.000438)	(0.000650)	(0.000555)	(0.000543)
Tax burden in the 5 closest municipalities (within)		-0.0243***	-0.0243***	-0.0123***	-0.00155
		(0.000530)	(0.00112)	(0.00102)	(0.00175)
Number of cinemas (within)		-0.00846**	-0.00846	-0.00316	0.000321
		(0.00404)	(0.00517)	(0.00429)	(0.00426)
Log(population) (between)			0.0136**	0.0163***	0.0184***
			(0.00529)	(0.00490)	(0.00487)
Share of foreign population (between)			0.326***	0.101	0.0197
			(0.0803)	(0.0714)	(0.0714)
Housing vacation rate (between)			-0.0417***	-0.0277***	-0.0237***
			(0.00422)	(0.00389)	(0.00404)
Tax burden in the 5 closest municipalities (between)			-0.0266***	-0.0188***	-0.0322***
			(0.00230)	(0.00221)	(0.00345)
Number of cinemas (between)			-0.0271***	-0.0210***	-0.0221***
			(0.00681)	(0.00573)	(0.00549)
ρ				0.475***	0.402***
				(0.0161)	(0.0168)
θ				-2.225***	-2.203***
				(0.0575)	(0.0581)
σ_e				0.00423***	0.00417***
				(0.000467)	(0.000463)
Observations	29822	29822	29822	29822	29822
R^2	0.371	0.258	0.388	0.408	0.465

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Table 1.5: Dependent variable: Gini index[a]

	(1) OLS	(2) FE	(3) Hybrid model	(4) Spatial AR	(5) Spatial Durbin
Log(distance to the closest lake - km)	-0.0426***	0	-0.0430***	-0.0363***	-0.0286***
	(0.00129)	(.)	(0.00437)	(0.00413)	(0.00461)
Log(distance to the closest river - km)	-0.00265	0	-0.00130	0.00683	0.0000887
	(0.00190)	(.)	(0.00698)	(0.00661)	(0.00789)
Altitude - km	0.180***	0	0.182***	0.159***	0.123***
	(0.00258)	(.)	(0.0122)	(0.0115)	(0.00799)
Share of forest	0.0638***	0	0.0654**	0.0872***	0.0811***
	(0.00828)	(.)	(0.0307)	(0.0291)	(0.0252)
Share of pasture	-0.0462***	0	-0.0410	-0.00929	0.0760**
	(0.0109)	(.)	(0.0313)	(0.0295)	(0.0378)
Travel time to the closest agglomeration (min)	0.00184***	0	0.00184***	0.00190***	0.00249***
	(0.0000939)	(.)	(0.000380)	(0.000362)	(0.000367)
Log(population)	-0.0143***				
	(0.00131)				
Share of foreign population	0.445***				
	(0.0187)				
Housing vacation rate (%)	0.00771***				
	(0.000928)				
Tax burden in the 5 closest municipalities	0.0198***				
	(0.000581)				
Number of cinemas	0.0213***				
	(0.00185)				
Log(net income)	0.672***				
	(0.00583)				
Log(population) (within)		0.0834***	0.0834***	0.0191	0.0196*
		(0.0115)	(0.0309)	(0.0297)	(0.0116)
Share of foreign population (within)		-0.106***	-0.106	-0.182**	0.0703**
		(0.0350)	(0.0935)	(0.0910)	(0.0347)
Housing vacation rate (within)		-0.000165	-0.000165	-0.000409	0.000211
		(0.000552)	(0.000872)	(0.000845)	(0.000511)
Tax burden in the 5 closest municipalities (within)		0.0151***	0.0151***	0.0151***	0.000928
		(0.000693)	(0.00178)	(0.00176)	(0.000982)
Number of cinemas (within)		-0.000683	-0.000683	0.000573	-0.00341
		(0.00508)	(0.00575)	(0.00524)	(0.00469)
Log(income)(within)		0.657***	0.657***	0.638***	0.810***
		(0.00759)	(0.0528)	(0.0534)	(0.00795)
Log(population) (between)			-0.0156***	-0.0148***	-0.0213***
			(0.00565)	(0.00544)	(0.00403)
Share of foreign population (between)			0.463***	0.371***	0.505***
			(0.0895)	(0.0859)	(0.0591)
Housing vacation rate (between)			0.0158**	0.0184***	0.0121***
			(0.00621)	(0.00543)	(0.00397)
Tax burden in the 5 closest municipalities (between)			0.0203***	0.0182***	0.00654**
			(0.00219)	(0.00208)	(0.00301)
Number of cinemas (between)			0.0224***	0.0237***	0.0230***
			(0.00591)	(0.00614)	(0.00540)
Log(income)(between)			0.684***	0.653***	0.764***
			(0.0401)	(0.0391)	(0.0201)
ρ				0.288***	0.383***
				(0.0126)	(0.00720)
θ				-1.883***	-1.857***
				(0.0369)	(0.0178)
σ_e				0.00779***	0.00706***
				(0.000289)	(0.0000607)
Observations	29822	29822	29822	29822	29822
R^2	0.445	0.246	0.447	0.433	0.509

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01
[a]: logistic transformation

time. The Gini index and the share of foreigners are positively related within and between municipalities.

The coefficient of vacation between municipalities rate is significantly negative on income and positive on the Gini index. Poor and unequal municipalities have high vacation rates (hence probably low housing prices). The mean income is negatively related to the tax burden between municipalities. This result indicates that rich people self-select themselves into areas where taxes are low (Tiebout, 1956). We also find that the tax burden and the Gini index (calculated on pre-taxed incomes) are positively correlated between municipalities. One explanation would be that the tax burden is considered here on married couples with an annual income of 50 000 CHF, which are not rich households. Probably the higher this rate, the less progressive is the taxation, so that taxes fall proportionally more on the middle of income distribution.

The number of cinemas has a negative between-coefficient in the regression on income and Gini. This is quite surprising, since we a priori would have considered this variable as a positive cultural amenity. This result may come from the fact that cinema is not a luxury good. As in Dewenter and Westermann (2005), it would suggest that the income-elasticity of the demand for cinema is smaller than one. In addition, cinemas are usually only located in urban areas and they attract consumers who live in other municipalities, adding some noise to the estimations. Finally, inequality is positively related to the mean income, both across time and space. This reflects the fact that distribution of income is typically more spread on the right than on the left. If there is some sorting, it mechanically implies that richer areas are more unequal.

The seemingly unrelated regression (equation 1.3) is partially reported in Table 1.6. For the sake of space, only δ_{2j} coefficients are shown. Table 1.15 list the other coefficients in the Appendix. The share of households in the two extreme categories is positively related to the proximity of lake or river, contrary to the middle class (30 000 - 75 000 CHF). The proportion of the poorest fringe of the population ($<$ 40 000 CHF) is positively linked to altitude, while the share of households earning over 50 000 CHF is lower in the mountains. The reverse is true concerning accessibility. Forest (and pasture) have a positive coefficient on all the categories but the (two) highest, suggesting again that the share of green amenities is an indicator of rurality, not a positively valued amenity. All in all, these seemingly unrelated regressions show that the link between income distribution and natural amenities is not only driven by the richest households. We find significant and consistent coefficients on the left of the distribution as well.

Table 1.6: Panel Seemingly Unrelated Regression

	(1) SUR
Share < 30 000 CHF[a]	
Log(distance to the closest lake - km)	-0.00467***
	(0.00163)
Log(distance to the closest river - km)	0.00539**
	(0.00238)
Altitude - km	0.0780***
	(0.00313)
Share of forest	0.0644***
	(0.0104)
Share of pasture	0.0366***
	(0.0138)
Travel time to the closest agglomeration (min)	0.00178***
	(0.000117)
Share 30 000 - 40 000 CHF[a]	
Log(distance to the closest lake - km)	0.000377
	(0.000859)
Log(distance to the closest river - km)	0.0119***
	(0.00126)
Altitude - km	0.00826***
	(0.00165)
Share of forest	0.0469***
	(0.00549)
Share of pasture	0.0549***
	(0.00728)
Travel time to the closest agglomeration (min)	0.000527***
	(0.0000618)
Share 40 000 - 50 000 CHF[a]	
Log(distance to the closest lake - km)	0.00372***
	(0.000835)
Log(distance to the closest river - km)	0.0141***
	(0.00122)
Altitude - km	-0.00118
	(0.00161)
Share of forest	0.0350***
	(0.00534)
Share of pasture	0.0552***
	(0.00708)
Travel time to the closest agglomeration (min)	0.000274***
	(0.0000601)

	(1) (continued)
Share 50 00 - 75 000 CHF[a]	
Log(distance to the closest lake - km)	0.00733***
	(0.000858)
Log(distance to the closest river - km)	0.00857***
	(0.00126)
Altitude - km	-0.0114***
	(0.00165)
Share of forest	0.0159***
	(0.00548)
Share of pasture	-0.00684
	(0.00727)
Travel time to the closest agglomeration (min)	-0.000144**
	(0.0000617)
Share > 75 000 CHF[a]	
Log(distance to the closest lake - km)	-0.00767***
	(0.00226)
Log(distance to the closest river - km)	-0.0342***
	(0.00330)
Altitude - km	-0.0631***
	(0.00434)
Share of forest	-0.127***
	(0.0144)
Share of pasture	-0.0941***
	(0.0191)
Travel time to the closest agglomeration (min)	-0.00200***
	(0.000163)
Observations	29822

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01
[a]: cubic root transformation: $S_{jit} = [\frac{Share_{jit}}{1 - Share_{jit}}]^{1/3}$

To conclude, all the regressions in this section show that the valuation of natural amenities is positively related to the mean income. The more the landscape is variate, the larger inequalities between municipalities. The results remain valid when we control for taxes and pressure on the housing market, indicating that natural amenities are factors of income sorting by themselves. Within inequality is also related to the environmental context. The Gini index is larger next to blue amenities and in the mountains, indicating that disparities are larger in particularly nice natural environments or in remote areas. This mechanism is not only driven by the richest households.

1.6 Robustness

We first change the definition of variables. That is, we replace the population by the number of taxpayers, we reduce the minimal size of a "lake" at 50 ha and we add the variables "Exposition" [17] and "Visibility of Alpin peaks". We also use the taxable instead of the net income. Different specifications of the spatial Durbin model are shown in table 1.7. The most robust coefficients are lake and altitude, which remain significant in all four specifications. When significant, travel time and distance are of the same sign as in previous section. Share of forest and pasture switch signs in the income regression when exposition is included (the observations numbers vary), which is positively linked to Gini and income. Visibility of Alpin peaks is, if anything, negatively related to income. We suspect multicollinearity between these variables.

The econometric modelling may also be changed. Instead of using a proxy of tax burden, we use a 2SLS regression. We also introduce time fixed effects, we lag the explanatory factors and estimate the model with first-difference of the time-variant variables. Finally, we run the model only with the 1242 municipalities that have not changed their borders between 2003 and 2016. Note that doing so, we cannot use the contiguity matrix any more and must settle for the hybrid model specification. Tables 1.8 and 1.9 show a synthesis of these results. The sign of the coefficients on amenities remains similar in the vast majority of cases. One variable is sensitive and looses its significance, namely the share of pasture. In general, the correlations remain in the same order of magnitude.

17. This inclusion implies the use of the state of municipalities in 2018, hence the reduction of N from 2294 to 2222 observations per year.

Table 1.7: Robustness: definition of variables - Spatial Durbin model

	(1) Log(taxable income)	(2) Log(taxable income)	(3) Gini index (taxable)	(4) Gini index (taxable)
Distance to the closest lake (50 ha)	-0.0376***	-0.0583***	-0.0683***	-0.0659***
	(0.00485)	(0.0115)	(0.00530)	(0.0137)
Distance to the closest river	-0.0658***	0.0440	-0.0459***	0.0360
	(0.00740)	(0.0317)	(0.00868)	(0.0359)
Altitude	-0.0480***	-0.0216*	0.0917***	0.0462***
	(0.0131)	(0.0128)	(0.0164)	(0.0162)
Number of visible Alpin peaks	-0.00241***	-0.000228	-0.000205	-0.000638
	(0.000660)	(0.000559)	(0.000892)	(0.000771)
Exposition		0.0164***		0.0141***
		(0.00381)		(0.00484)
Share of forest	-0.161***	0.0121**	-0.0473	0.0226***
	(0.0336)	(0.00549)	(0.0382)	(0.00675)
Share of pasture	-0.0982***	0.0115**	-0.0223	0.0167**
	(0.0364)	(0.00515)	(0.0458)	(0.00677)
Travel time to the closest agglomeration	-0.00275***	-0.00117*	-0.000338	0.000618
	(0.000357)	(0.000656)	(0.000415)	(0.000748)
Log(nb of taxpayers) (within)	-0.120***	-0.108***		
	(0.0201)	(0.0269)		
Log(population) (within)			0.162***	0.0800**
			(0.0267)	(0.0318)
Share of foreign population (within)	0.380***	0.0186***	0.392***	0.0209***
	(0.0633)	(0.00432)	(0.0886)	(0.00568)
Housing vacation rate (within)	-0.000188	0.000435	-0.000816	-0.0000154
	(0.000583)	(0.000474)	(0.000867)	(0.000676)
Tax burden in the 5 closest municipalities (within)	-0.00473***	0.104*	0.00400**	0.101
	(0.00115)	(0.0535)	(0.00156)	(0.0680)
Number of cinemas (within)	0.00412	-0.00198	-0.000921	-0.00923**
	(0.00413)	(0.00323)	(0.00570)	(0.00371)
Log(Taxable income)(within)			-0.763***	-0.689***
			(0.0284)	(0.0324)
Log(nb of taxpayers) (between)	0.0368***	0.00598		
	(0.00620)	(0.00523)		
Log(population) (between)			-0.00607	-0.00743
			(0.00767)	(0.00677)
Share of foreign population (between)	0.237***	0.0210***	0.715***	0.0355***
	(0.0872)	(0.00744)	(0.103)	(0.00871)
Housing vacation rate (between)	-0.0442***	-0.0119***	-0.0187***	-0.00794**
	(0.00443)	(0.00324)	(0.00621)	(0.00377)
Tax burden in the 5 closest municipalities (between)	-0.0277***	0.612***	-0.0000751	0.843***
	(0.00242)	(0.193)	(0.00283)	(0.239)
Number of cinemas (between)	-0.0225***	-0.0155***	0.00451	0.00916*
	(0.00601)	(0.00460)	(0.00452)	(0.00533)
Log(Taxable income)(between)			-0.0844	-0.176***
			(0.0601)	(0.0534)
ρ	0.110***	0.379***	0.0638***	0.318***
	(0.0145)	(0.0300)	(0.0106)	(0.0180)
θ	-2.299***	-2.059***	-2.056***	-1.857***
	(0.0534)	(0.0548)	(0.0432)	(0.0495)
σ_e	0.00477***	0.00423***	0.00984***	0.00918***
	(0.000499)	(0.000455)	(0.000750)	(0.000751)
Observations	29822	26664	29822	26664
R^2	0.430	0.528	0.225	0.287

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Table 1.8: Robustness: income model

	(1) 2SLS Log(taxable income)	(2) Time FE Log(income)	(3) Lagged Log(income)	(4) 1st diff. Δ Log(income)	(5) Non-merged mun. Log(income)
[1em] Distance to the closest lake (100 ha)	-0.0505*** (0.00528)	-0.0313*** (0.00475)	-0.0308*** (0.00467)	-0.000811*** (0.000218)	-0.0265*** (0.00648)
Distance to the closest river	-0.0542*** (0.00800)	-0.0487*** (0.00698)	-0.0596*** (0.00694)	0.000436 (0.000295)	-0.0326*** (0.00884)
Altitude	-0.126*** (0.00913)	-0.0826*** (0.00885)	-0.0794*** (0.00881)	-0.000783* (0.000462)	-0.158*** (0.00893)
Share of forest	-0.158*** (0.0335)	-0.122*** (0.0304)	-0.144*** (0.0307)	-0.00589*** (0.00156)	-0.207*** (0.0396)
Share of pasture	-0.0407 (0.0392)	-0.0388 (0.0330)	-0.0528 (0.0325)	0.00124 (0.00182)	0.0242 (0.0490)
Travel time to the closest agglomeration	-0.00209*** (0.000344)	-0.00200*** (0.000326)	-0.00244*** (0.000333)	-0.0000497*** (0.0000187)	-0.00137*** (0.000415)
Log(nb of taxpayers) (within)	0.0882*** (0.0216)	-0.167*** (0.0193)			0.0819*** (0.0194)
Log(nb of taxpayers) (1st diff)				-0.318*** (0.0477)	
Log(nb of taxpayers) (t-1) (within)			-0.0411** (0.0176)		
Share of foreign population (within)	0.740*** (0.0738)	0.281*** (0.0588)			0.977*** (0.0942)
Share of foreign population (t-1) (within)			0.335*** (0.0646)		
Share of foreign population (1st diff)				-0.0448 (0.0560)	
Housing vacation rate (within)	0.000402 (0.000573)	-0.000284 (0.000540)			0.000806 (0.000844)
Housing vacation rate (t-1) (within)			-0.000434 (0.000585)		
Housing vacation rate (1st diff)				-0.000454 (0.000335)	
Tax burden (within)	-0.0153*** (0.00158)				
Tax burden in the 5 closest municipalities (within)		0.000779 (0.00112)			-0.0258*** (0.00174)
Tax burden in the 5 closest municipalities (t-1) (within)			-0.00380*** (0.00101)		
Tax burden in the 5 closest municipalities (1st diff)				-0.000547 (0.000810)	
Number of cinemas (within)	-0.00764* (0.00460)	0.00572 (0.00353)			-0.0140 (0.00970)
Number of cinemas (t-1) (within)			0.00232 (0.00381)		
Number of cinemas (1st diff)				-0.00240 (0.00235)	
Log(nb of taxpayers) (between)	0.0165*** (0.00599)	0.0373*** (0.00575)	0.0315*** (0.00578)	-0.00164*** (0.000298)	0.0277*** (0.00707)
Share of foreign population (between)	0.441*** (0.0901)	0.268*** (0.0846)	0.201** (0.0817)	0.00161 (0.00341)	0.200** (0.0978)
Housing vacation rate (between)	-0.0457*** (0.00469)	-0.0411*** (0.00423)	-0.0413*** (0.00421)	-0.000914*** (0.000205)	-0.0496*** (0.00494)
Number of cinemas (between)	-0.0277*** (0.00764)	-0.0272*** (0.00681)	-0.0234*** (0.00607)	0.000170 (0.000178)	-0.0286*** (0.00824)
Tax burden in the 5 closest municipalities (between)		-0.0179*** (0.00220)	-0.0260*** (0.00223)	-0.000647*** (0.000104)	-0.0247*** (0.00291)
ρ		0.133*** (0.0194)	0.186*** (0.0205)	0.0635*** (0.00980)	
θ		-2.339*** (0.0560)	-2.263*** (0.0574)	20.34*** (0.0576)	
σ_e		0.00399*** (0.000446)	0.00417*** (0.000462)	0.00416*** (0.000599)	
Observations	20861	29822	27528	27528	16634
R^2	0.299	0.406	0.419	0.0274	0.396

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

Table 1.9: Robustness: Gini index model

	(1) 2SLS Gini index (taxable)	(2) Time FE Gini index	(3) Lagged Gini index	(4) 1st diff. Δ Gini index	(5) Non-merged mun. Gini index
Distance to the closest lake (100 ha)	-0.0680***	-0.0441***	-0.0454***	-0.000533*	-0.0588***
	(0.00563)	(0.00436)	(0.00438)	(0.000313)	(0.00623)
Distance to the closest river	-0.0465***	-0.00361	-0.00556	-0.00435***	-0.0274***
	(0.00891)	(0.00704)	(0.00707)	(0.000443)	(0.00924)
Altitude	0.0889***	0.161***	0.165***	0.00210***	0.194***
	(0.0137)	(0.0127)	(0.0128)	(0.000750)	(0.0155)
Share of forest	-0.0703*	0.0651**	0.0663**	0.00848***	0.0309
	(0.0383)	(0.0316)	(0.0316)	(0.00215)	(0.0430)
Share of pasture	-0.0748*	-0.0483	-0.0354	0.00771***	-0.0663
	(0.0432)	(0.0316)	(0.0316)	(0.00240)	(0.0465)
Travel time to the closest agglomeration	-0.000315	0.00143***	0.00131***	-0.0000746***	0.00190***
	(0.000386)	(0.000394)	(0.000396)	(0.0000250)	(0.000536)
Log(Taxable income)(within)	-0.557***				
	(0.0339)				
Log(income)(within)		0.801***			0.677***
		(0.0529)			(0.0621)
Log(income)(t-1) (within)			0.183***		
			(0.0494)		
Log(income)(1st diff)				1.076***	
				(0.0386)	
Log(nb of taxpayers) (within)	0.416***	0.626***			0.510***
	(0.0281)	(0.0311)			(0.0299)
Log(nb of taxpayers) (t-1) (within)			0.367***		
			(0.0280)		
Log(nb of taxpayers) (1st diff)				1.009***	
				(0.0671)	
Share of foreign population (within)	0.503***	-0.168**			-0.668***
	(0.0882)	(0.0819)			(0.111)
Share of foreign population (t-1) (within)			0.0148		
			(0.0736)		
Share of foreign population (1st diff)				-0.0386	
				(0.0747)	
Housing vacation rate (within)	-0.00127	-0.00134			-0.00209*
	(0.000878)	(0.000863)			(0.00107)
Housing vacation rate (1st diff)				0.000360	
				(0.000576)	
Tax burden (within)	-0.000558				
	(0.00207)				
Tax burden in the 5 closest municipalities (within)		0.00757***			0.0233***
		(0.00138)			(0.00236)
Tax burden in the 5 closest municipalities (t-1) (within)			0.00912***		
			(0.00139)		
Tax burden in the 5 closest municipalities (1st diff)				0.00200*	
				(0.00105)	
Number of cinemas (within)	-0.0139**	-0.0102*			-0.000786
	(0.00626)	(0.00536)			(0.00823)
Number of cinemas (t-1) (within)			-0.00231		
			(0.00537)		
Number of cinemas (1st diff)				0.00211	
				(0.00398)	
Log(Taxable income)(between)	-0.117**				
	(0.0580)				
Log(nb of taxpayers) (between)	-0.00577	-0.0144**	-0.0135**	0.00140***	-0.00869
	(0.00676)	(0.00619)	(0.00620)	(0.000444)	(0.00875)
Share of foreign population (between)	0.689***	0.436***	0.446***	0.0159***	0.462***
	(0.101)	(0.0882)	(0.0882)	(0.00499)	(0.131)
Housing vacation rate (between)	-0.0188***	0.0164***	0.0165***	0.00124***	0.0216**
	(0.00652)	(0.00629)	(0.00634)	(0.000337)	(0.00948)
Number of cinemas (between)	0.00115	0.0162***	0.0162***	0.000266	0.0173**
	(0.00466)	(0.00544)	(0.00532)	(0.000256)	(0.00715)
Log(income)(between)		0.703***	0.707***	-0.00814***	0.643***
		(0.0416)	(0.0424)	(0.00188)	(0.0539)
Tax burden in the 5 closest municipalities (between)		0.0212***	0.0213***	0.000402***	0.0185***
		(0.00218)	(0.00220)	(0.000154)	(0.00267)
[1em] Housing vacation rate (t-1) (within)			-0.00190**		
			(0.000935)		
ρ		0.00703	0.0166*	0.103***	
		(0.0123)	(0.00961)	(0.00799)	
[1em] θ		-1.998***	-1.801***	21.22***	
		(0.0380)	(0.0465)	(0.0466)	
σ_e		0.00718***	0.00941***	0.00596***	
		(0.000233)	(0.000685)	(0.000197)	
Observations	20861	29822	27528	27528	16634
R^2	0.208	0.461	0.439	0.467	0.414

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01

1.7 Conclusion

In this chapter, we study the link between the Swiss spatial distribution of income and natural amenities. We find that income and inequalities are high in municipalities that are close to water, well exposed to the sunshine and from where we can see the Alps. In addition, the higher the elevation, the travel time to the next agglomeration and the share of forest, the lower the average income. Inequalities are also more pronounced in the mountains and far away from the closest agglomeration.

In general, controlling for other factors, income inequalities both between *and* within municipalities are positively related to the existence of natural amenities. The richer the households, the more they value natural amenities, so that income tends to concentrate next to positive-valued amenities, which explains spatial inequality. Less obviously, nice places are also more unequal, and this fact is not only driven by the richest households. According to Fretz et al. (2017), imperfect sorting comes from commuting and trade costs, while preferences over consumption and housing are idiosyncratic. If that interpretation is correct, then any decrease of distance-related costs implies a reinforcement of income sorting, hence an increase of inequality between municipalities over time.

Our analysis is based on the mean income, the Gini index, and income classes at the municipality level. Other indicators would have been instructive, such as the Theil index (which is additive and decomposable by subgroups) and the share of income held by the richest households (top 1% or 10%). However, these variables are not available at the municipal level. In the same vein, additional controls could be added in the regressions to refine our understanding of the role of amenities. For instance, the attraction effect of water is not necessarily the same from one lakeside municipality to another - it may depend on the characteristics of the lake, on the topography and on the orientation.

Having these several caveats in mind, we can anyway conclude that the link between the studied natural amenities and income distribution exists, that these amenities are permanent and that the income sorting process is self-reinforcing. We thus expect high persistence of inequalities between and within municipalities in a heterogeneous landscape like Switzerland. This issue is further explored in chapter 2.

1.8 Appendix

Derivation of the Gini index

The index is calculated as follows: for a given time t and municipality i, let $Income_n$ be the income of individual $n = 1, 2, ...N$. $Income_n$ is indexed in a non-decreasing order, such that $Income_1 \leq Income_2 \leq ... \leq Income_N$. The Gini coefficient is given by:

$$Gini = \frac{1}{N}[N + 1 - 2\frac{\sum_{n=1}^{N}(N+1-n)Income_n}{\sum_{n=1}^{N} Income_n}] = \frac{2\sum_{n=1}^{N} n \cdot Income_n}{N\sum_{n=1}^{N} Income_n} - \frac{N+1}{N} \quad (1.4)$$

The Gini index is bounded between 0 and 1, we therefore use its logistic transformation:

$$G_{it} = \log\left(\frac{Gini_{it}}{1 - Gini_{it}}\right) \quad (1.5)$$

Descriptive statistics

Table 1.10: Summary statistics 2003

Variable	Obs	Mean	Std. Dev.	Min	Max
Average net income (CHF/year)	2 294	36 898.69	9 314.55	10 917	193 795
Median net income (CHF/year)	2 294	30 563.66	4 292.385	4 600	53 600
Gini index (net equivalent)	2 294	.427	.058	.258	.858
Tax burden (%)	800	21.436	2.918	5.89	26.9
Tax burden in the 5 closest municipalities (%)	2 294	21.221	2.534	11.044	25.846
Share of taxpayers under 30 000 CHF (net)	2 294	.136	.052	.008	.72
Share of taxpayers 30 - 40 000 CHF (net)	2 294	.126	.025	.036	.368
Share of taxpayers 40 - 50 000 CHF (net)	2 294	.144	.022	0	.296
Share of taxpayers 50 - 75 000 CHF (net)	2 294	.28	.026	0	.507
Share of taxpayers over 75 000 CHF (net)	2 294	.313	.081	0	.685
Number of taxpayers	2 294	1 892.673	6 846.994	17	230 202
Population	2 294	3 188.253	10 596.52	24	342 116
Share of foreign population	2 294	.202	.106	0	.538
Housing vacation rate (%)	2 294	.901	1.134	0	19.108
Number of cinemas	2 294	2.727	5.439	0	18

The statistics are weighted by the population (except for population and number of taxpayers)

Table 1.11: Summary statistics 2008

Variable	Obs	Mean	Std. Dev.	Min	Max
Average net income (CHF/year)	2 294	41 037.4	13 364.51	13 081	276 359
Median net income (CHF/year)	2 294	32 416.56	48 44.692	4 800	54 700
Gini index (net equivalent)	2 294	.442	.065	.287	.88
Tax burden (%)	804	20.532	2.938	6.843	26.477
Tax burden in the 5 closest municipalities (%)	2 294	20.409	2.569	10.17	25.443
Share of taxpayers under 30 000 CHF (net)	2 294	.123	.05	0	.635
Share of taxpayers 30 - 40 000 CHF (net)	2 294	.109	.022	0	.357
Share of taxpayers 40 - 50 000 CHF (net)	2 294	.135	.024	.03	.458
Share of taxpayers 50 - 75 000 CHF (net)	2 294	.274	.031	.083	.458
Share of taxpayers over 75 000 CHF (net)	2 294	.359	.084	.036	.706
Number of taxpayers	2 294	2 012.986	7 148.246	17	242 480
Population	2 294	3 310.154	10 940.09	19	358 540
Share of foreign population	2 294	.211	.106	0	.568
Housing vacation rate (%)	2 294	.98	1.137	0	14.596
Number of cinemas	2 294	2.422	4.998	0	19

The statistics are weighted by the population (except for population and number of taxpayers)

Table 1.12: Summary statistics 2015

Variable	Obs	Mean	Std. Dev.	Min	Max
Average net income (CHF/year)	2 294	42 850.58	14 773.31	16 603	478 065
Median net income (CHF/year)	2 294	33 467.16	4 788.394	5 850	55 200
Gini index (net equivalent)	2 294	.452	.069	.264	.936
Tax burden (%)	2 294	19.81	2.933	7.205	25.737
Tax burden in the 5 closest municipalities (%)	2 294	19.731	2.676	10.534	24.823
Share of taxpayers under 30 000 CHF (net)	2 294	.122	.048	.022	.611
Share of taxpayers 30 - 40 000 CHF (net)	2 294	.105	.022	0	.333
Share of taxpayers 40 - 50 000 CHF (net)	2 294	.133	.027	0	.417
Share of taxpayers 50 - 75 000 CHF (net)	2 294	.251	.031	.036	.412
Share of taxpayers over 75 000 CHF (net)	2 294	.389	.087	0	.747
Share of foreign population	2 294	.243	.111	0	.606
Number of taxpayers	2 294	2 209.182	7710.732	18	263 358
Population	2 294	3 590.962	11 856.01	13	391 359
Housing vacation rate	2 294	1.166	1.241	0	11.523
Number of cinemas	2 294	1.821	4.261	0	19

The statistics are weighted by the population (except for population and number of taxpayers)

Table 1.13: Summary statistics - constant variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Distance to the closest lake (km)	2 294	12.135	10.027	.068	51.19
Distance to the closest river (km)	2 294	28.598	17.533	.541	75.882
Altitude (m)	2 294	823.822	579.883	197	3451
Visibility of Alpin peaks	2 294	5.919	10.869	0	112
Share of forest	2 294	.302	.158	0	.88
Share of pasture	2 294	.16	.129	0	.687
Travel time to the closest agglomeration (min)	2 294	33.093	15.456	0	161
Exposition	2 222	2.274	1.165	1	5

Regressions

Table 1.14: Spatial Durbin model: spatial dependence of explanatory variables

		(1) Log(income)	(2) Gini index
β_4	Log(population) (within)	0.0597**	0.267***
		(0.101)	(0.0205)
	Share of foreign population (within)	0.764***	-0.243***
		(0.000776)	(0.0657)
	Housing vacation rate (%) (within)	0.00275**	-0.00234**
		(0.00109)	(0.00108)
	Tax burden in the 5 closest municipalities (within)	-0.0128***	0.00776***
		(0.00196)	(0.00126)
	Number of cinemas (within)	-0.0000906	0.00230
		(0.0149)	(0.0107)
	Log(income)(within)		-0.599***
			(0.0134)
β_5 :	Log(distance to the closest lake - km)	0.0527***	-0.00397
		(0.00739)	(0.00698)
	Log(distance to the closest river - km)	0.0224*	-0.0179*
		(0.0129)	(0.0108)
	Altitude - km	-0.0432***	0.0715***
		(0.0122)	(0.0144)
	Share of forest	-0.165***	-0.114**
		(0.0548)	(0.0465)
	Share of pasture	-0.165***	-0.157***
		(0.0531)	(0.0571)
	Travel time to the closest agglomeration (min)	0.00177***	-0.00266***
		(0.000506)	(0.000518)
β_6 :	Log(population) (between)	-0.0166**	0.0339***
		(0.00813)	(0.00721)
	Share of foreign population (between)	0.700***	-0.353***
		(0.124)	(0.110)
	Housing vacation rate (between)	-0.0224***	0.00423
		(0.00689)	(0.00730))
	Tax burden in the 5 closest municipalities (between)	0.0188***	0.00891**
		(0.00404)	(0.00367)
	Number of cinemas (between)	0.00153	-0.00342
		(0.0133)	(0.0127)
	Log(income)(between)		-0.380***
			(0.0352)
Observations		29822	29822

Complement to column (4) of Tables 1.4 and 1.5. It shows the coefficient on spatially lagged explanatory variables in equation (1.2)

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

Table 1.15: Seemingly unrelated regression

Dep. var:	Share < 30	Share 40-50	Share > 75
Log(population) (within)	-0.0209***	-0.0960***	0.121***
	(0.00421)	(0.00408)	(0.00397)
Share of foreign population (within)	0.0440***	0.0374***	0.00656
	(0.0133)	(0.0129)	(0.0125)
Housing vacation rate (within)	0.000341*	-0.000785***	0.000283
	(0.000196)	(0.000190)	(0.000185)
Tax burden in the 5 closest municipalities (within)	0.000659**	-0.000915***	0.00282***
	(0.000278)	(0.000270)	(0.000262)
Number of cinemas (within)	-0.000733	0.00146	0.00220
	(0.00180)	(0.00174)	(0.00170)
Log(population) (between)	-0.0158***	0.00325***	0.0108***
	(0.00168)	(0.000864)	(0.00234)
Share of foreign population (between)	0.159***	-0.0694***	0.0364
	(0.0244)	(0.0125)	(0.0338)
Housing vacation rate (between)	0.00957***	0.00765***	-0.0255***
	(0.00164)	(0.000840)	(0.00227)
Tax burden in the 5 closest municipalities (between)	0.00478***	0.00315***	-0.00853***
	(0.000763)	(0.000392)	(0.00106)
Number of cinemas (between)	0.0118***	0.00404***	-0.0208***
	(0.00236)	(0.00121)	(0.00327)
Year FE	yes	yes	yes

Dep. var:	Share 30-40	Share 50-75
Log(population) (within)	-0.00512	-0.0326***
	(0.00407)	(0.00391)
Share of foreign population (within)	-0.0376***	-0.0414***
	(0.0128)	(0.0123)
Housing vacation rate (within)	-0.000522***	0.0000306
	(0.000190)	(0.000182)
Tax burden in the 5 closest municipalities (within)	-0.000975***	-0.00202***
	(0.000269)	(0.000258)
Number of cinemas (within)	0.000128	-0.00261
	(0.00174)	(0.00167)
Log(population) (between)	-0.00118	0.00743***
	(0.000888)	(0.000887)
Share of foreign population (between)	-0.0552***	-0.103***
	(0.0129)	(0.0128)
Housing vacation rate (between)	0.00891***	0.00537***
	(0.000864)	(0.000863)
Tax burden in the 5 closest municipalities (between)	0.00461***	-0.000721*
	(0.000403)	(0.000402)
Number of cinemas (between)	0.00620***	0.00148
	(0.00124)	(0.00124)
Year FE	yes	yes

Observations: 29 822

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

Chapter 2

The persistence of spatial income inequalities between Swiss municipalities

2.1 Introduction

Income inequalities are increasing within developed countries. Joseph E. Stiglitz raises the alarm on the high price the world could pay because of inequality (Stiglitz, 2012). He states that stability, efficiency and growth of the economic system are threatened, as well as democracy and social cohesion. Switzerland is not an exception compared to the developed countries described in Piketty (2013) and Bourguignon (2018). The usual indicators of inequality have increased during the last decades. According to the data of the Federal Tax Administration [1], the overall Gini index calculated on net equivalent income increased from 0.347 in 1971/72 to 0.392 in 2015. [2]

In any federalist structure, the question of spatial inequality between jurisdictions is inescapable. As in the USA and in Europe, the issue is central in Switzerland, especially because taxes and redistribution policies are largely implemented at the cantonal and municipal level. Brueckner (2000) mentions several examples of externalities which may arise in a spatial equi-

1. Precisions about the methodology can be found in AFC (2013).
2. Two other sources are concordant: the Gini index calculated on the basis of data from the old-age and survivor's insurance (OASI) has increased from 0.4 to 0.43 between 1980 and 2010 (see Martínez (2017)). Moreover, according to the world inequality database (WID, 2019), the share of national taxable income held by the top 1% (5%) has increased from 8.7% (21.51%) in 1971 to 11.3% (24.08%) in 2014.

librium with substantial geographical disparities: congestion costs and excessive commuting, undervaluation of environmental amenities and open spaces, or non-internalisation of the infrastructural costs of development. Moreover, social negative consequences of income sorting are well recognized. Cytron et al. (2011) list the severe effects of the geographical concentration of poverty on the ability of local governments to spend money on school, health and safety.

Furthermore, spatial disparities are self-reinforcing over time, for at least three reasons. First, as shown by Brueckner et al. (1999), local amenities are luxury goods, so that rich households tend to concentrate in nice natural and cultural environments. Second, wealthy municipalities are able to set lower tax rates, without unbalancing their budget. This, in turn, attracts households, especially those with high income. Income sorting in response to taxation may be so high that the effective country-wide average tax rate becomes regressive. According to Roller and Schmidheiny (2016), this is the case in Switzerland for very high levels of income. Third, housing rents increase in response to population inflows. This tends to crowd out the poorest people from attractive places.

In this paper, we explore how spatial disparities among Swiss municipalities have evolved since the seventies building upon the most recent investigative techniques and proposing some refinements. Economic activity is prone to important spatial spillovers across municipalities. We must therefore carefully consider the role of spatial autocorrelation in our investigations. We do so by using exploratory spatial analysis tools to describe the income distribution in Switzerland from the 1970s. We raise the question of whether or not there is hope of convergence among communes. We analyse convergence among municipalities using spatial dynamic panel specifications. Finally, we build an autoregressive model and simulate the spatial distribution of income in 2040.

The relative standard deviation of income and the Theil index reveal that spatial inequality has increased particularly rapidly since 2004/05. At the same time, the strength of spatial spillovers has decreased. In the regressions, we find that Swiss communes fill the gap with the national mean at an annual rate of 2 to 5% when heterogeneity is ignored. However, further analysis shows that each municipality actually follows its own growth path. Thus, overall disparities increase over time because the initial heterogeneity persists and strengthens through path dependency. During the same period, spatial autocorrelation comes down, which is consistent with decreasing distance-related costs.

Finally, we simulate the future distribution of income with high in-sample fit indicators with 2004/05 - 2014/15 data (in the best case, the Pearson correlation coefficient between actual and predicted income of 2014/15 is 0.996). The model includes the lagged income, the lagged income of the contiguous municipalities and an individual fixed effect. If the income distribution follows the same process in the future, the relative standard deviation and the Theil index in 2040/41 would increase significantly compared to 2014/15.[3] This suggests that fiscal equalisation schemes and regional development policies are currently not sufficient to counterbalance income sorting.

The chapter is organised as follows: section 2.2 presents briefly the literature on income convergence and section 2.3 shows the theoretical relationship between β-convergence and σ-convergence in the presence of spatial spillovers and initial heterogeneity. Section 2.4 presents the dataset and section 2.5 the descriptive statistics, the estimates of β-convergence and the simulation of the future income distribution. Concluding remarks and policy implications can be found in section 2.6.

2.2 Literature review

Income convergence

The tendency of convergence between regions has been famously formalised by Robert Solow (Solow, 1956). In his view, given diminishing returns to capital and the fact that technological change crosses borders, the poorest regions are able to catch-up with the richer by imitation. Empirically, authors define convergence either as β-convergence or σ-convergence. The former means that the coefficient found when regressing growth on income level is negative. The latter implies that income variance across regions decreases over time. Absolute convergence (i.e. convergence in the absence of other controls) is rejected by most empirical evidence. As shown in the landmark article of Deininger and Squire (1996), growth does not automatically benefit particularly to the poor. The catching-up process is therefore only conditional, hence the existence of “convergence clubs” and the extensions to endogenous growth (Barro and Sala-i

3. The Theil index would increase from 0.052 to 0.112 and the relative standard deviation of income from 0.501 to 0.729. The 2040/41 estimations are lower bounds, since we assume the same level of national income as in 2014/15.

Martin, 2003) models. Methodologically, convergence has been studied either through Markov transition matrices (Quah, 1993), or by the estimation of the growth equation, i.e. $g_{it} = c + \beta \ln y_{it-\Delta} + \varepsilon_{it}$, where $g_{it} = \ln y_{it} - \ln y_{it-\Delta}$, with $\ln y_{it}$ the natural logarithm of per capita income in country i at time t and Δ the time gap. In this last case, the use of dynamic panels is advocated by Islam (1995). Indeed, the initial conditions make the parameters of the production function specific to each location. When the omitted variable bias is corrected, the rate of convergence is substantially higher. In terms of spatial equity though, the prospect is not better, since each area converges toward a different steady-state.

Studies on income convergence and spatial dependence have flourished in the 90s and 00s. The research has been motivated by a crucial issue in a particular region: is there hope of convergence between members of the European Union? Armstrong (1995) shows that the estimated speed of convergence between the GDP per capita of European regions between 1950 and 1990 is slower when peripheral disadvantaged regions are included in the dataset. Crozet and Koenig (2005) discuss the growth-equity trade-off across European regions, Fischer and Stumpner (2008) work on income distribution dynamics and cross-region convergence in Europe, whereas Martin (2009) describes within and across regions cohesion in the European Union. These authors conclude on the weakness of convergence and the increase of spatial inequality. The rate of absolute convergence is generally around 2% among European regions, while σ-convergence is not observed.

Spatial autocorrelation

Spatial effects play a central role in the convergence process. As stated in Egger and Pfaffermayr (2006), "it is the remoteness of a region which determines its possibility to catch up" (p.208). Spatial autocorrelation has been explored since the end of the 90s. López-Bazo et al. (1999) and Gallo and Ertur (2003) apply spatial exploratory analysis tools such as Moran's I and Getis-Ord indicators. They detect substantial spatial autocorrelation and growth clusters among European regions. Scholars introduce spatial dependence in the convergence process with two main methods. The first is to estimate Markov transition probabilities matrix conditional on the income of neighbours (e.g. Gallo (2004)). The second is to calculate β-convergence by including a spatial lag [4] of the initial income, of the growth rate and/or of the error terms

4. The vector is pre-multiplied by a spatial weight matrix.

(e.g. Moreno and Trehan (1997) and Ertur and Koch (2007) on worldwide countries; Fingleton (1999), Fingleton and López-Bazo (2006), López-Bazo et al. (2004), Egger and Pfaffermayr (2006), Badinger et al. (2004) and Bouayad-Agha and Védrine (2010) on European regions; Rey and Montouri (1999), Rey (2004b), Rey (2004a) and Yu and Lee (2012) on US states). These authors find that spatial spillovers matter. In general, positive spatial autocorrelation increases β-convergence in central regions and hampers it on the peripheral areas. The influence of spatial spillovers on the evolution of σ is not predetermined. It depends on the pattern of spatial dependence (i.e. the degree to which regions are close to each other).

Income sorting

The articles cited in the previous subsection measure convergence based on GDP per capita, which is only an approximation of income.[5] Yet, income convergence is in general more difficult to achieve than productivity convergence. The spatial distribution of wages obviously depends on where economic activity concentrates (see Eeckhout et al. (2014)) but the persistence of spatial inequalities is also explained by factors that have nothing to do with the location of production.

The role of taxation in income sorting has received particular interest in a federalist framework. In a seminal paper, Tiebout (1956) develops a model in which different regions provide different combinations of public goods. Households choose their preferred location until the achievement of an equilibrium within the decentralised state. Oates (1969) includes the role of land value in the sorting model. Since property taxes reduce property values and because the presence of public goods makes the demand of residence increase, the price of housing is negatively related to the tax rate. In other terms, low taxes are capitalised into land values. Jurisdictions with low taxes and high housing prices therefore attract the richest households and repel the poorest.

Switzerland is an ideal case to explore the Tiebout model and tax capitalisation. Fiscal federalism has been at the core of the Sinergia project "The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization".[6] Schmidheiny (2017) presents

5. Martin (2009) is an exception, he explores both income and GDP. His results are different depending on which variable is used.

6. The project was funded by the Swiss National Science Foundation from 2010 to 2016. Main applicants are Marius Brülhart, Monika Bütler, Mario Jametti and Kurt Schmidheiny. Additional details can be found on the website http://fiscalfederalism.ch

the collected data and summarises the main lessons of this project. In light of the objective of the present chapter, three papers deserve a particular mention. Brülhart and Parchet (2014) study the sensitivity of elderly people to the level of bequest tax in Swiss cantons. They do not find that bequest tax differentials affect significantly the location choices of old taxpayers. Roller and Schmidheiny (2016) measure the effective progressivity of the Swiss income tax system. They show that income sorting significantly lowers tax rates that rich households effectively face. For some high income level, the effective tax scheme becomes even regressive. Eugster and Parchet (2019) exploit the language frontier to identify interjurisdictional tax competition. Using a difference-in-differences strategy, they find that local governments anticipate income sorting and set lower taxes than they would do in the absence of tax competition within a 20 km radius. [7] Papers that directly deal with the distributional consequences of the Swiss fiscal system are however rare. Two recent articles mention the role of jurisdictional fragmentation on inequality: Kübler and Rochat (2017) point that inequality of pre-tax income is lower in Switzerland than in other OECD countries, but that the differential disappears when we compare post-tax distributions. According to them, geopolitical fragmentation therefore hampers redistribution. Feld et al. (2018) study the influence of fragmentation and decentralisation on inequality at the cantonal level. They find that decentralisation reduces pre-tax income concentration within cantons, but this result is conditional on low fragmentation.

Compared to the existing literature, the present chapter presents some particular characteristics. First, the analysis is conducted at the municipal level in Switzerland, rather than on American States, European regions or even Swiss cantons. [8] By studying the smallest possible aggregation level, we mitigate the modifiable areal unit problem (MAUP) introduced by Open-

7. In addition, the impact of taxation on migration and income sorting in Switzerland has been explored by Liebig and Sousa-Poza (2006) who explore the individual responsiveness of tax variations. They find that young, educated and nationals are more likely to migrate after a tax increase. Schmidheiny (2006) uses data from Basel to estimate a multinomial probit. He finds that rich people are more likely to choose low tax jurisdictions. Stadelmann and Billon (2012) regress land prices on fiscal and expenditure variables in the canton of Zurich. They highlight the link between the ability to capitalise fiscal variables into land value and the land scarcity. According to them, capitalisation occurs even if more land is available. On the same subject, Morger (2017) finds that capitalisation decreases when income increases, meaning that higher housing prices are more than offset by lower taxes. Martínez (2016) uses a tax reform in Obwald as a quasi-experiment; she finds that rich taxpayers are quite reactive to decreasing tax rates. The elasticity of the stock of rich with respect to the average net-of-tax rate is indeed between 2.4 and 3.5. Basten et al. (2017) take advantage of the language frontier to draw a boundary discontinuity design. They find an income tax elasticity of housing rents of 0.26. One third of this effect is driven by income sorting, the rest consists of direct capitalisation.

8. Few exception work at a more disaggregated level, such as Kakamu and Fukushige (2005) on Japanese cities and the Lee and Lin (2017) on American metropolitan areas.

shaw (1984).[9] Second, we focus on *income* data gathered by the tax administration. Abstracting from the tax evaders, these data are fairly reliable and reflect better than GDP what residents earn. Third, we consider both β- and σ-convergence. As highlighted by Fischer and Pfaffermayr (2018), the latter is unfortunately often disregarded or not distinguished from β-convergence in the existing literature. Finally and contrary to most of the studies about Switzerland, we take explicitly spatial effects into account. Indeed, many shocks and adjustments that affect one municipality cross the communal borders.

2.3 Convergence and spatial autocorrelation with heterogeneity

In this section, we formalise the process of income growth and the evolution of its variance in the presence of fixed effects and spatial autocorrelation. In this framework, we follow the argument of Young et al. (2008) to show that β-convergence is necessary but not sufficient for σ-convergence. We also introduce spatial effects using the same notation as Egger and Pfaffermayr (2006).

If the income follows a basic growth process à la Solow (1956), we can write the income at time $t+1$ in matrix notation:

$$\mathbf{Y}_{t+1} = \alpha\boldsymbol{\iota} + (\mathbf{I} + \beta\mathbf{I})\mathbf{Y}_t + \boldsymbol{\Xi}_t \tag{2.1}$$

where $\mathbf{Y}_t$ is the vector of log per capita incomes, $\ln y_{it}$, of each region i (municipality). α is a constant scalar and $\boldsymbol{\iota}$ a vector of 1 of dimension $N \times 1$. $\mathbf{I}$ is the $N \times N$ identity matrix and $\boldsymbol{\Xi}_t$ the $N \times 1$ vector of error terms ε_{it}. In this framework, the income variance converges to its steady state level, iff $-1 \leq \beta < 0$.[10] Indeed, expressing the variable of interest in deviation from the mean, we have ($\tilde{\mathbf{Y}}_{t+1}$):

$$\boldsymbol{\Omega}_{t+1} = E(\tilde{\mathbf{Y}}_{t+1}\tilde{\mathbf{Y}}'_{t+1}) \cong (\mathbf{I} + \beta\mathbf{I})(\mathbf{I} + \beta\mathbf{I})'\boldsymbol{\Omega}_t + \boldsymbol{\Omega}_{\Xi} \tag{2.2}$$

9. From a pure geographical point of view, any aggregation of spatial units is arbitrary. In combinatorial terms, there is an infinite number of ways to group the 2309 Swiss municipalities into 26 entities. As shown by Openshaw (1984), results may substantially vary across different definitions of the geographical unit.

10. $-1 \leq \beta < 0$ is also a necessary condition for stationarity.

where $\mathbf{\Omega}_t$ is the variance-covariance matrix of income at time t and $\mathbf{\Omega}_\Xi$ the variance-covariance of the error terms. The equality is perfect when the error terms are independent from $\mathbf{Y}_t$. In this context, the variance converges to its steady state $\mathbf{\Omega}^* = \mathbf{\Omega}_\Xi(\mathbf{I} - (\mathbf{I} + \beta\mathbf{I})(\mathbf{I} + \beta\mathbf{I})')^{-1}$, where $\mathbf{\Omega}^* = \sigma_*^2\mathbf{I}$ in the case of iid error terms. So, in the absence of frequent shocks on the error terms, β-convergence implies that variance converges to σ_*^2. If it does from above, then σ-convergence occurs. Moreover, the higher β in absolute value (i.e. the higher the speed of convergence), the faster the decrease in σ.

Let us extend the framework to include spatial and region-specific effects. In that case, $\mathbf{Y}_{t+1}$ follows the process:

$$\mathbf{Y}_{t+1} = \mathbf{A} + (\mathbf{I} + \beta\mathbf{I} + \theta\mathbf{W})\mathbf{Y}_t + \mathbf{\Xi}_t \tag{2.3}$$

where $\mathbf{A}$ is a vector with α_i specific to each row. The growth path is therefore different in each region. θ is the coefficient of spatial autocorrelation theoretically lying between -1 and 1. $\mathbf{W}$ is the row-standardized Queen contiguity matrix.[11] Spatial dependence is modelled as time-space recursive (see Anselin et al. (2008)): growth is conditional on the initial income of contiguous municipalities.[12] Then, the evolution of the variance becomes more complex. By developing the expression (see algebraic details in Appendix 2.7), we obtain:

$$\mathbf{\Omega}_{t+1} = E(\tilde{\mathbf{Y}}_{t+1}\tilde{\mathbf{Y}}'_{t+1}) \cong \mathbf{\Omega}_A + (\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})(\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})'\mathbf{\Omega}_t + \mathbf{\Omega}_\Xi + \mathbf{\Sigma}_{Y_t,A} \tag{2.4}$$

where $\mathbf{\Omega}_A$ is the variance-covariance matrix of the fixed effects and $\mathbf{\Sigma}_{Y_t,A}$ the covariance between $\mathbf{Y}_t$ and $\mathbf{A}$. This term is increasing in t in the presence of path dependency and spatial autocorrelation, i.e. positive elements of $(\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})$. We neglect the covariance of A and Y_t with the error terms, hence the approximation sign.

11. The non-standardized contiguity matrix $\mathbf{W}_{NS}$ is a $N \times N$ matrix which contains 1 on the ith row and the jth column if municipality i and j are contiguous in the Queen sense (i.e. the two borders share at least one common pair of XY coordinates), 0 if $i = j$ or if i and j are not neighbours. The row-standardized contiguity matrix $\mathbf{W}$ is obtained by normalising the elements of $\mathbf{W}_{NS}$ such that they sum up to 1 by row.

12. As highlighted by Fingleton (2014), dynamic stability requires that all the eigenvalues of $(\mathbf{I} + \beta\mathbf{I} + \theta\mathbf{W})$ lie inside the unit circle.

The introduction of the individual fixed effects increases the steady state variance, if it exists.[13] Besides, as $\mathbf{\Sigma}_{Y_t,A}$ is in general positive, $\mathbf{\Omega}_A$ and $\mathbf{\Sigma}_{Y_t,A}$ widen the gap between $\mathbf{\Omega}_{t+1}$ and $\mathbf{\Omega}_t$, slowing σ-convergence (or accelerating σ-divergence). Generally, the influence of spatial autocorrelation is not determined a priori. It depends on the intensity (θ) and the pattern ($\mathbf{W}$) of spatial spillovers.[14]

2.4 Data

Income data used in this study come from the Sinergia project described in footnote 6. The original data have been gathered by the Swiss Federal Tax Administration (reference document AFC (2013), website AFC (2018b)). The dataset contains communal average and median income, in net and taxable terms. "Net" income is income after removal of insurance and savings interests. "Taxable" income excludes the fiscal deductions and is about 30% lower than net income. Averages and medians are either calculated by taxpayer, or in equivalent terms. In the latter case, each household is divided by an equivalence factor [15], so that we can interpret the statistic per individual. The figures take into account any natural person who pays the Federal Direct Tax, including special cases [16], in their last municipality of legal residence.

Income data are provided every pair of year from 1971/72 to 1995/96, and every year between 2003 and 2015. We distinguish the two periods, because of the significant change in the tax system that is detailed in Foellmi and Martínez (2017). Swiss cantons (except Basel-Stadt) had applied the *praenumerando* taxation until 1998. In this system, people report the income they have earned during the last two years (assessment period) and they pay the corresponding tax during the two following years (fiscal period). The available years are the fiscal periods from 1973/74 to 1997/98, corresponding to the assessment periods from 1971/72 to 1995/96.[17] In the presentation of our results, we consider assessment periods, that is the period during which the income was earned. A reform was decided in 1995: every cantons had to switch to the *postnumerando* system, where taxes are calculated on the current income and paid during the

13. The steady state variance is given by$[\mathbf{\Omega}_A + \mathbf{\Omega}_\Xi + \mathbf{\Sigma}_{Y_t,A}][\mathbf{I} - (\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})(\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})']^{-1}$. It is finite if the eigenvalues of $(\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})$ are inside the unit circle.
14. See Badinger et al. (2004), p. 243.
15. Singles have a factor of 1 and couples of 1.5. Each child is worth 0.3.
16. Persons who are also taxed abroad (such as secondary residents) and fictive incomes of those who pay a forfeit.
17. Except the assessment period 1985/86 (i.e. fiscal period 1987/88) which is missing for an unknown reason.

same year. However, the 26 cantons adapted at different moments: Zurich and Thurgau already switched in 1999, Ticino, Vaud and Valais in 2003, all the others in 2001. The year of "tax holiday" [18] is different from one canton to another so that data are not complete between 1998 and 2002. From 2003 onwards, the 26 cantons are available again, fiscal and assessment periods coincide and data are yearly published until 2015. In our analyses, we use the moving average over two years, so that the results are comparable between the two periods. [19]

The state of municipalities constantly evolves because of mergers. We use the state of January 1st, 2015 and do as if the municipalities have always been in this state. [20] In case of changes between January 1971 and December 2014, we replace the income of the "not yet existing" commune before the year of mutation by the weighted average [21] of its future components. We rely on the list of mutations of the statistical office (OFS, 2018b). This manipulation avoids the drop of many observations that would bias our sample. Municipality mergers is indeed a political process which is not randomly distributed over space, time and income level.

In the spatial analyses, we use the geographical coordinates of municipal boundaries. The contiguity matrix was built from the geodata VEC200 of Swisstopo. [22]

2.5 Results

In this section, we expose descriptive statistics on the dispersion of income and on spatial autocorrelation. Then, we estimate spatial dynamic panel models of β-convergence. Finally, we test the in-sample fit of the predictions and suggest a simulation of the 2040/41 spatial distribution of income.

σ-divergence

A first set of graphs shows the evolution of spatial disparities in Switzerland from 1971/72 to 2014/15. Figure 2.1 shows the evolution of the relative standard deviation (variation coefficient).

18. The income earned during the switching year was never declared.

19. $y_t = \dfrac{(y^y_{t-1} + y^y_t)}{2}$ with $t = 2005, 2007, ...2015$, with y^y_t being the yearly available income. All the analyses have also been conducted on yearly data. They yield similar results, which are available upon request from the author.

20. Except Bolligen, Rubigen and Schlosswil, that are cases of splits. These three municipalities are treated as if they had never been separated.

21. Weights correspond to the number of taxpayers.

22. XY coordinates of the boundaries, LV03 projection system.

The statistics is calculated on the average net equivalent income (y) by municipality (indexed by subscript i) given by:

$$\sigma_t = \frac{1}{\bar{y}_t}\sqrt{\sum_{i=1}^{N}\omega_{it}(y_{it}-\bar{y}_t)^2} \tag{2.5}$$

$$\text{where } \omega_{it} = \frac{n_{it}}{\sum_{i=1}^{N} n_{it}} = \frac{n_{it}}{n_t^{ch}}$$

where N is the number of municipalities and $\bar{y}_t$ the national average income (per individual) at time t. Each municipality is weighted by its size: ω_{it} where n_{it} is the number of taxpayers in commune i and n_t^{ch} the total number of taxpayers in Switzerland at time t.

We use the relative rather than the absolute standard deviation, because the average income is almost three times bigger in 2014/15 than in 1971/72.[23] This increase is largely due to inflation[24] and makes the variance of income increase mechanically.

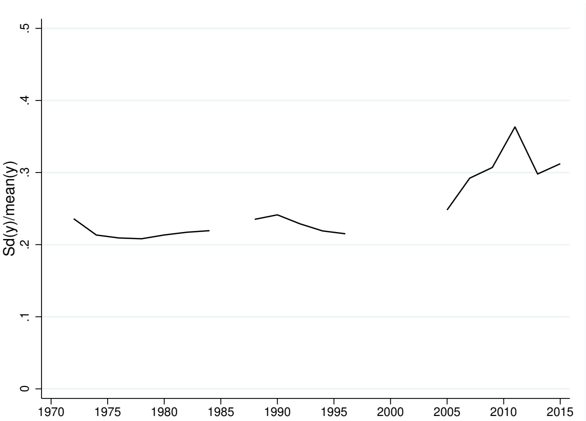


Figure 2.1: Relative standard deviation of average income 1971/72 - 2014/15

Note: the end of the assessment period is represented on the horizontal axis. For example, the standard deviation of income in assessment period 1971/72 (corresponding to the fiscal period 1973/74) is drawn in 1972. We excluded Anières (GE), of which average income reached almost 1.5 millions CHF in 2014/15 (see the article of "La Tribune de Genève" by Mabut (2015)).

The relative variance of average income has been very stable until 1996, but then it almost

23. See table 2.6, Appendix 2.7.
24. According to the Statistical office (OFS, 2019), the Consumer Price Index (reference year 1966) was 120.1 in 1971 and 331.5 in 2015.

doubled between 2004/05 and 2014/15. σ-divergence occurs in the most recent period of the sample.

The evolution of spatial disparities can also be analysed through the Theil index. Since this indicator belongs to the entropy class (see Shorrocks (1980)), it can be expressed as a weighted sum of the indices for population subgroups. We are therefore able to disentangle the within- and between-cantonal inequality. Formally, following Shorrocks and Wan (2005), we can write:

$$\begin{aligned} T_t &= \sum_{i=1}^{N} \omega_{it} \cdot \frac{y_{it}}{\bar{y}_t} \cdot \ln\left(\frac{y_{it}}{\bar{y}_t}\right) \\ &= \underbrace{\sum_{c=1}^{M} s_{ct} \cdot T_{ct}}_{within} + \underbrace{\sum_{c=1}^{M} s_{ct} \cdot \ln\left(\frac{\bar{y}_{ct}}{\bar{y}_t}\right)}_{between} \end{aligned} \tag{2.6}$$

$$\text{where } s_{ct} = \frac{n_{ct} \cdot \bar{y}_{ct}}{n_t^{ch} \cdot \bar{y}_t}$$

T_t is the Theil index measuring overall inequality across municipalities, T_{ct} the Theil index within each canton c, n_{ct} the number of taxpayers in canton c, M the number of cantons, $\bar{y}_t$ the national mean income and $\bar{y}_{ct}$ the cantonal mean income at time t.

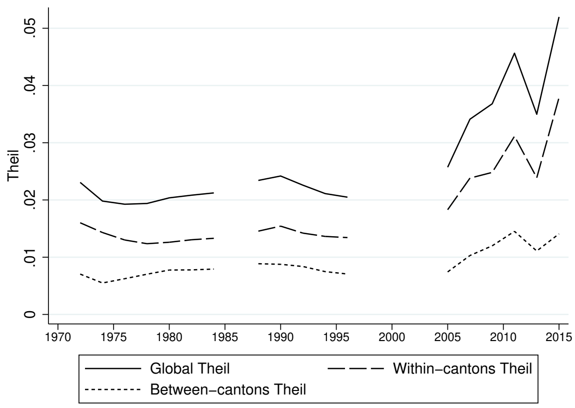


Figure 2.2: Theil Index decomposition 1971/72 - 2014/15

The Theil decomposition indicates that income inequalities between Swiss municipalities take place mainly within cantons. The increase of the Global Theil since 2004/05 follows the evolution of the within-component, while disparity between cantons remains more stable over time. As does the standard deviation, the Theil index fluctuates substantially more in the recent period.

Overall, this indicator confirms that spatial disparities are increasing since 2004/05. There are some important short term fluctuations, probably driven by the very rich municipalities.

Spatial autocorrelation

We calculate Moran's I index in order to detect spatial autocorrelation in income distribution. Following Getis and Ord (1992), we write the statistic as:

$$I_t = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}(y_{it}-\bar{y}_t)(y_{jt}-\bar{y}_t)}{\frac{1}{N}\cdot\sum_{i=1}^{N}(y_{it}-\bar{y}_t)^2\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}}, i \neq j \tag{2.7}$$

w_{ij} are the element of the i^{th} row and j^{th} column of the spatial weight matrix $\mathbf{W}$. In our case, $\mathbf{W}$ is a row-standardized Queen contiguity matrix. There is positive (negative) global spatial autocorrelation if I is significantly lower (higher) than the expectation of I under no spatial autocorrelation.[25]

The Moran's I is positive and highly significant every year as shown in figure 2.3a.[26] Two different processes lie behind a positive autocorrelation. First, neighbours share common characteristics (e.g. landscape, canton and accessibility). Second, one municipality benefits (or suffers) from spatial spillovers of economic activity. For instance, the establishment of a business in one municipality may have positive effects on income all around. The parallel is obvious with what the New Economic Geography literature defines as first and second nature factors (see Lin et al. (2015)). The first nature factors are the initial conditions, for example the proximity of a lake or a river, the altitude and the topography. The second nature factors refer to the "path dependency" effects: economic activity attracts other firms that attracts workers, who are consumers, which attracts firms, and so on. It is empirically difficult to properly disentangle these two effects, because initial conditions still influence current economic variables such as productivity and housing prices.[27] Nevertheless, the comparison of figures 2.3a (income level) and 2.3b (income deviation from cantonal mean) provides two indications about spatial autocorrelation. First, it shows how spatial spillovers evolve over time and second which part of the spatial

25. $E_0(I) = \frac{-1}{N-1}$. The p-values are derived from the normal distribution (see Anselin (1995)).

26. Corresponding figures are presented in table 2.7. In the Appendix 2.7, the reader may also find additional facts on spatial autocorrelation. In particular, the Moran scatterplots of 2014/15 (figures 2.14), as well as the Moran maps (figure 2.15).

27. The effects are separable only if the "first nature factor" has lost its advantage over time, as this is the case of portages studied by Bleakley and Lin (2012).

(a) Income level

(b) Income deviation from cantonal mean

Figure 2.3: Global Moran's I: 1971/72 - 2014/15

autocorrelation is due to the cantonal effect.

The statistic is increasing during the 70s/80s and decreasing since the early 90s. The Moran's I negatively co-evolves with standard deviation and the Theil index. The right-hand side curve is only slightly lower than the left one and follows a similar path, which suggests that spatial autocorrelation remains important within cantons.

In conclusion, positive spatial autocorrelation occurs, but is declining over time. Average income of a given municipality is less and less related to the one of its neighbours, which would advocate for a decline of the "path dependency" effect. This may come from a decrease of commuting and trade costs. The income sorting process is reinforced as workplaces and places of residence become increasingly far apart. In this context, residential motivations take priority on job-based decisions (see Bayer and McMillan (2012)).

Together with the observed σ-divergence, the fact that spatial spillovers decrease over time points toward an increasing importance of initial heterogeneity (e.g. natural environment, cultural amenities, transport infrastructure) in the income dispersion.

β-convergence

We estimate the "time-space recursive" model according to the typology of spatial dynamic panels developped by Anselin et al. (2008): [28]

$$\ln y_{it} = c + (1+\beta)\ln y_{it-1} + \theta \sum_{j\neq i}^{N} w_{ij} \ln y_{jt-1} + \gamma \ln n_{it} + \delta_t + \underbrace{\alpha_i + \nu_{it}}_{\varepsilon_{it}} \tag{2.8}$$

where $\ln y_{it}$ is the logarithm of income at time t. In case of convergence, β is negative and lies between -1 and 0. The smaller $1+\beta$, the faster the convergence. [29] $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt-1}$ is the arithmetical mean of the logarithms of neighbours' incomes [30], with w_{ij} the element of the row-standardized contiguity matrix. n_{it} holds for the number of taxpayers. If the error terms ε_{it} are randomly distributed over t and i, then the model can be estimated by an OLS. In this case, β represents absolute convergence. However, when the residuals are composed by a non-random individual-specific part, such that $\varepsilon_{it} = \alpha_i + \nu_{it}$, then each municipality follows its own growth path and the convergence is only conditional. In this case, we face a dynamic panel bias. As shown by Nickell (1981), α_i are inherently and positively [31] correlated with $\ln y_{it-1}$ which implies that the fixed effect estimators underestimate the true coefficient. The Arellano and Bond generalized method of moments (Arellano and Bond, 1991) provides an estimator for dynamic panels. The basic idea is to first differentiate the equation and then use the lags of order 2 and more as instrumental variables. The instrument is valid in the absence of serial correlation of order 2, 3..., T.

However, with persistent data, the first difference specification may result in weak instruments. Blundell and Bond (1998) suggest rather a system-GMM estimator, where the equation in level is estimated simultaneously with the one in first-difference. Lagged differences are used as instruments in the level equation, and conversely. The last step consists of checking that all the instruments are valid (i.e. non-correlated with $\ln y_{it}$). If the model does not pass the Sargan

28. The model can also be written with the growth rate as the dependent variable, as in Rey and Montouri (1999). Since $g_{it} = \ln y_{it} - \ln y_{it-1}$, the coefficient in front of $\ln y_{it-1}$ is β.

29. The rate of convergence is then given by $\dfrac{-\ln(1+\beta)}{\Delta}$, where Δ is the time lag. See Barro and Sala-i Martin (2003).

30. Equivalently, the logarithm of the geometrical mean of neighbours' incomes, given that $\sum w_j \ln(y_j) = \ln(\prod y_j^{w_j})$.

31. α_i captures the initial conditions such as the initial level of technology, the natural environment and the institutional factors, see Islam (1995).

test, we must restrict the number of instruments.[32] In the case of spatial dynamic panels, the use of system-GMM is advocated by Kukenova and Monteiro (2009), since it corrects for endogeneity of both time and spatial lag ($\ln y_{it-1}$ and $\sum_{j \neq i}^{N} w_{ij} \ln y_{jt-1}$). In addition, system-GMM does not require the inversion of the contiguity matrix, which is an important practical advantage when N is large. Table 2.1 presents the coefficients estimated by OLS (1 and 6), by panel fixed effects (2 and 7), by the Arellano-Bond GMM (3 and 8), by the two-step Blundell-Bond system GMM with all the possible instruments (4 and 9) and by the system GMM with only valid instruments (5 and 10).

Table 2.1: β-convergence, time-space recursive

	Period 1971/72 - 1995/96				
	(1)	(2)	(3)	(4)	(5)
	OLS	Within	GMM^a	Sys-GMM^b	Sys-GMM^c
L.Log(income)	0.908***	0.519***	0.399***	0.541***	0.631***
	(0.00266)	(0.00569)	(0.0954)	(0.0734)	(0.100)
L.W Log(income)	0.0846***	0.135***	-0.115**	-0.0961*	0.457***
	(0.00333)	(0.00966)	(0.0550)	(0.0525)	(0.104)
Log(nb taxpayers)	-0.0000434	0.0318***	0.00368	0.00214	0.000707
	(0.000320)	(0.00310)	(0.0216)	(0.0201)	(0.00201)
Constant	0.203***	3.294***	7.222***	5.494***	-0.815***
	(0.0215)	(0.0889)	(0.552)	(0.459)	(0.128)
Year FE	yes	yes	yes	yes	yes
N	23090	23090	18472	23090	23090
R^2	0.971	0.925	0.502	0.899	0.954

	Period 2004/05 - 2014/15				
	(6)	(7)	(8)	(9)	(10)
	OLS	Within	GMM^a	Sys-GMM^b	Sys-GMM^d
L.Log(income)	0.958***	0.269***	0.514**	0.946***	0.803***
	(0.00364)	(0.0105)	(0.236)	(0.101)	(0.0917)
L.W Log(income)	0.0573***	0.0913***	-0.0931	-0.140*	-0.158
	(0.00462)	(0.0206)	(0.0847)	(0.0844)	(0.103)
Log(nb taxpayers)	-0.00240***	-0.0148	-0.0503**	0.0193	0.0114***
	(0.000457)	(0.00915)	(0.0245)	(0.0276)	(0.00214)
Constant	-0.100***	6.945***	6.547***	1.991***	3.776***
	(0.0308)	(0.227)	(1.685)	(0.524)	(0.408)
Year FE	yes	yes	yes	yes	yes
N	11545	11545	9236	11545	11545
R^2	0.945	0.796	.630	.930	0.883

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

Notation in the indications below:: "L(k/m)" means "lags of order $k, k+1, ..., m-1, m$". The point represents the maximum possible, that is T-1.

L(2/.) must be therefore interpreted as lag of order 2, 3,... T-1). Δ is the first difference, i.e $X_t - X_{t-1}$, L.Δ is the lag of order 1 of Δ, i.e. $X_{t-1} - X_{t-2}$

^{a} : Instruments for differenced equation: L(2/.).Log(income), L.ΔWLog(income) , ΔLog(nb tp) and year fixed effects

^{b} : Instruments for differenced equation: L(2/.).Log(income), L.ΔWLog(income), ΔLog(nb tp) and year fixed effects ; instruments for level equation: L.ΔLog(income)

^{c} : Instruments for differenced equation: L(10/.).Log(income); L(12/.).WLog(income) and ΔLog(nb tp) ;

instruments for level equation: L2.ΔLog(income), L2.ΔWLog(income) and Log(nb tp)

^{d} : Instruments for differenced equation: L(4/.).Log(income), L(4/.).WLog(income), ΔLog(nb tp);

instruments for level equation: L.ΔLog(income); L.ΔWLog(income) and Log(nb tp)

The first line represents $1 + \beta$. In the OLS specification, municipal incomes converge to the national mean at an annual rate of 4.83% between 1971/72 and 1995/96. On the recent period, this rate falls at 2.15%. However, the Hausman test rejects the independence of α_i in both sub-samples[33] and $\hat{\beta}_{OLS}$ is therefore upward biased. Columns (2) and (7) present the

32. See Blundell and Bond (2000), pp. 128-129.
33. Model 1971/72 - 1995/96: H_0: differences in coefficients is not systematic. $\chi^2(12) = 6322.706$ and $Pr(> \chi^2) = 0.000$.

coefficients estimated by the fixed effect panel method. $\hat{\beta}_{within}$ is downward biased (see Nickell (1981)), such that this model provides a lower bound for the true β. The third and eighth columns show the estimation of the Arellano-Bond GMM. In the second period, $\hat{\beta}_{GMM}$ lies between $\hat{\beta}_{OLS}$ and $\hat{\beta}_{within}$, meaning that the instruments allow to recover part of the true effect of y_{it-1} (including α_i). In contrast, in model (3), $\hat{\beta}_{GMM}$ is smaller than the within estimator, which reflects the weakness of the instruments. Indeed, with persistent data, time variations are small compared to cross-country differentials, which causes an important downward bias of the Arellano-Bond GMM estimator. Thus, we estimate β by the Blundell and Bond (1998) system GMM estimator. The two-steps coefficients are presented in the fourth and ninth columns. As expected, the coefficient is larger than $\hat{\beta}_{GMM}$ and lies between $\hat{\beta}_{OLS}$ and $\hat{\beta}_{within}$. However, the Sargan test rejects the validity of over-identifying restrictions [34] (see Blundell and Bond (2000)), the instruments are therefore correlated with the dependent variable. In order to obtain a correctly specified model, we restrict the number of instruments in models 5 and 10, until it passes the Sargan test. For instance in model 10, only income and neighbours' income lags of order 4-5 and first difference of the log size are included in the difference equation. The lagged differences of income and neighbours' income as well as the log of the number of taxpayers are included in the level equation. [35] To validate models (5) and (10), we must check the absence of serial correlation. The Arellano-Bond test (Arellano and Bond, 1991) does not reject the hypothesis of absence of serial correlation of order 2 (p-value of 0.376 in the first sub-sample and of 0.227 in the second). Finally, the error terms in the three GMM models are robust to heteroscedasticity. In the last GMM specification, we find an annual rate of convergence ($-\ln(1+\beta)/2$) of 23% in the first period, that falls to 11% in the second. There has been more inertia in income distribution since 2004/05. These rates may seem high, but let us remember that convergence is defined here as conditional. As stated by Islam (1995):

> "there is probably little solace to be derived from finding that [countries in the world] are converging at a faster rate, when the points to which they are converging remain very different" (p.1162).

Model 2004/05 - 2014/15: H_0: differences in coefficients is not systematic. $\chi^2(7) = 5061.58$ and $Pr(>\chi^2) = 0.000$.

34. Model 1971/72 - 1995/96: H_0: instruments are valid; $\chi^2(48) = 692.620$; $Pr(>\chi^2) = 0.000$.
Model 2004/05 - 2015/15: H_0: instruments are valid; $\chi^2(13) = 58.708$; $Pr(>\chi^2) = 0.000$.

35. See list below table 2.1. Results of the Sargan test are the following:
Model 1971/72 - 1995/96: H_0: instruments are valid; $\chi^2(10) = 15.97$; $Pr(>\chi^2) = 0.101$.
Model 2004/05 - 2014/15: H_0: instruments are valid; $\chi^2(9) = 16.304$; $Pr(>\chi^2) = 0.0608$.

Regarding the coefficient in front of the neighbours' income, θ is positive and significant in the first sub-period. In general, it reinforces β-convergence when the municipality is surrounded by a less extreme neighbourhood (by definition, this is the case on average). However, spatial effects disappear since 2004/05. All in all, results suggest that initial heterogeneity has an increasing importance compared to spatial spillovers effects. In terms of attracting wealth, municipalities became more and more competitors rather than complements since 2004/05. First nature factors seem therefore to increasingly anchor the income distribution. This result is consistent with a context of tax decentralisation [36] and decreasing commuting costs: [37] the location of production matters less and less compared to the residence-based determinants of households' location choices such as local amenities, housing prices and taxes.

Robustness

In this section, we test several alternatives to the model in column (5) and (10). We estimate the system GMM model by using the municipal median income instead of its mean. Then, we allow time-varying spatial dependence by introducing interaction terms between spatial lags and time dummies. Indeed, the strength of spatial spillovers may vary over time due, for example, to the development of transport infrastructure. Following Anselin et al. (2008) taxonomy, we also estimate a "time-space simultaneous" model in which $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt}$ replaces $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt-1}$ in equation (2.8). The "time-space dynamic" model, in which both $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt}$ and $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt-1}$ are included, is also tested. Finally, we estimate the system GMM with data spatially filtered by the Getis procedure. The basic idea of this procedure is to decompose income between its expectation under no spatial autocorrelation and its spatial component. As detailed in Getis and Griffith (2002) and Badinger et al. (2004), the filtered income is given by:

$$\tilde{y_i} = y_i \frac{\sum_j w_{ij}(d)/(N-1)}{G_i(d)} \tag{2.9}$$

$$\text{where } G_i(d) = \frac{\sum_j w_{ij}(d) \cdot y_j}{\sum_j y_j}$$

36. Indicators of tax variability are presented in Appendix 2.7, figure 2.12. Since 2003, the average communal tax burden has decreased, while its variance has increased.
37. See ARE (2013).

$w_{ij}(d)$ is the element of the i^{th} row and j^{th} column of the row-normalised inverse-distance matrix (not contiguity). $G_i(d)$ is the Getis and Ord local indicator of spatial association.[38]

Table 2.2: β-convergence, alternative specifications (system GMM estimations)

	Period 1971/72 - 1995/96				
	(1) Median[a]	(2) Time-varying θ^b	(3) Time-space simultaneous[c]	(4) Time-space dynamic[d]	(5) Getis spatial filtering[e]
Dep. Var:	Log(med. income)	Log(income)	Log(income)	Log(income)	Log(filt. income)
L.Log(med. income)	0.942***				
	(0.0316)				
L.Log(Wmed. income)	0.142***				
	(0.0306)				
L.Log(income)		0.657***	0.782***	0.523***	
		(0.108)	(0.189)	(0.0692)	
L.Log(filt. income)					0.933***
					(0.00319)
L.WLog(income)		0.789***		-0.586**	
		(0.114)		(0.230)	
WLog(income)			0.269	1.058***	
			(0.187)	(0.210)	
L.W Log(income) * 1973/74		0.119			
		(0.271)			
L.W Log(income) * 1975/76		-0.0295			
		(0.254)			
L.W Log(income) * 1977/78		0.0323***			
		(0.00816)			
L.W Log(income) * 1979/80		0.0306***			
		(0.00739)			
L.W Log(income) * 1981/82		0.0318***			
		(0.00625)			
L.W Log(income) * 1983/84		0.0235***			
		(0.00472)			
L.W Log(income) * 1991/92		0.0180***			
		(0.00278)			
L.W Log(income) * 1993/94		0.0146***			
		(0.00167)			
L.W Log(income) * 1995/96		0.00412***			
		(0.000430)			
Log(number of taxpayers)	-0.00348***	-0.00471	0.000237	0.00837***	0.00165***
	(0.000781)	(0.0260)	(0.0250)	(0.00212)	(0.000459)
Constant	-0.740***	-4.636***	-0.454***	-0.00528	0.756***
	(0.111)	(1.311)	(0.114)	(0.198)	(0.0309)
Year FE	yes	no	yes	yes	yes
N	23090	23090	23090	23090	23090
R^2	.961	.377	.889	.957	.975

	Period 2004/05 - 2014/15				
	(6) Median[f]	(7) Time-varying θ^g	(8) Time-space simultaneous[g]	(9) Time-space dynamic[g]	(10) Getis spatial filtering[h]
Dep. Var:	Log(med. income)	Log(income)	Log(income)	Log(income)	Log(filt. income)
L.Log(med. income)	0.441***				
	(0.0574)				
L.Log(Wmed. income)	0.316***				
	(0.0537)				
L.Log(income)		0.844***	0.634***	0.790***	
		(0.0804)	(0.120)	(0.0886)	
L.Log(filt. income)					0.575***
					(0.0851)
L.WLog(income)		0.594		-0.472***	
		(0.387)		(0.133)	
WLog(income)			0.0702	0.515***	
			(0.190)	(0.186)	
L.W Log(income) * 2008/09		-0.00616			
		(0.00572)			
L.W Log(income) * 2010/11		-0.00903			
		(0.00669)			
L.W Log(income) * 2012/13		-0.0112			
		(0.00789)			
L.W Log(income) * 2014/15		-0.0135			
		(0.00884)			
Log(nb taxpayers)	0.00463***	-0.0179	0.0101***	0.00554*	0.0127***
	(0.00122)	(0.0142)	(0.00288)	(0.00291)	(0.00342)
Constant	2.557***	-4.449	3.142***	1.785**	4.530***
	(0.151)	(3.743)	(0.839)	(0.818)	(0.900)
Year FE	yes	no	yes	yes	yes
N	11545	11545	11545	11545	11545
R^2	.862	.866	0.880	.931	0.812

Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01
[a]: Instruments for differenced equation: L(9/.).Log(med. income); L(11/.).Log(Wmed. income) and ΔLog(nb tp) ; instruments for level equation: L2.ΔLog(med. income) ; L2.ΔLog(Wmed. income) and Log(nb tp)
[b]: Instruments for differenced equation: L(10/.).Log(income), L(11/.).WLog(income)and ΔLog(nb tp); instruments for level equation: L2.ΔLog(income), L2.ΔWLog(income) and Log(nb tp).
[c]: Instruments for differenced equation: L(10/.).Log(income); L(11/.).WLog(income) and ΔLog(nb tp) ; instruments for level equation: L2.ΔLog(income), L3.ΔWLog(income) and L3.ΔLog(nb tp)
[d]: Instruments for differenced equation: L(10/.).Log(income); L(11/.).WLog(income) and ΔLog(nb tp) ; instruments for level equation: L2.ΔLog(income), L2.ΔWLog(income) and L2.ΔLog(nb tp)
[e]: Instruments for differenced equation: L(9/.).Log(filt. income) ; ΔLog(nb tp) and year dummies. ; instruments for level equation: L3.ΔLog(filt. income), Log(nb tp) and year fixed effects.
[f]: Instruments for differenced equation: L(4/.).Log(med. income), L(4/.).Log(Wmed. income) and ΔLog(nb tp); instruments for level equation: L.ΔLog(med. income), L.ΔLog(Wmed. income) and Log(nb tp)
[g]: Instruments for differenced equation: L(4/.).Log(income); L(4/.).WLog(income) and ΔLog(nb tp); instruments for level equation: L.ΔLog(income); L.ΔWLog(income) and Log(nb tp).
[h]: Instruments for differenced equation: L(4/.).Log(filt. income) and ΔLog(nb tp) ; instruments for level equation: L.ΔLog(filt. income) and Log(nb tp)

Compared to the average income estimations, the conditional convergence of the median income is slower in the first period and faster in the second, while spatial autocorrelation is lower in the first and higher in the second period. This fact may be linked to the increasing trend of within inequalities. Indeed, the higher the dispersion of income within each municipality, the further away the average income is from the median. This would indicate that the increase in $1+\beta$ and the decrease of θ observed on the average income in the recent past is mainly driven by the right-hand side of the income distribution. In models (2) and (7) , $1+\beta$ are similar to the ones obtained in the baseline specification, suggesting that spatial autocorrelation (θ) is stable within one sub-period. The interaction terms are not always significant, they are even dropped because of multicollinearity in some cases. When significant, they are not increasing, so there

38. $E_0[G_i(d)] = \dfrac{\sum_j w_{ij}(d)}{N-1}$ is the expectation of the indicator in the absence of spatial autocorrelation.

is not any clear trend in the evolution of spatial autocorrelation.

$1+\beta$ remains of the same magnitude in time-space simultaneous (3 and 8) and time-space dynamic models (4 and 9). The slight variations come from multicollinearity between $\ln y_{it}$ and $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt}$. When calculated on the Getis filtered income, the rate of convergence is slower in the first period and faster in the second compared to the coefficients in table 2.1. The elements of the spatial weight matrix $w_{ij}(d)$ now translate inverse-distance instead of contiguity. This makes an important difference especially when spatial autocorrelation is high.

In all the specifications, there is some conditional β-convergence: $1+\beta$ is always significant and between 0 and 1. Its magnitude varies according to the inertia of the variables (median income and filtered income tend to be more stable than average income) and depending on the multicollinearity with other controls. Among them, neighbouring effects are often significant. θ is in general positive (especially in the first period), does not show any increasing trend within a period and is more consistent when specified in dynamic terms (i.e. on $\sum_{j\neq i}^{N} w_{ij} \ln y_{jt-1}$). In addition, municipal fixed effects are improving the model in any case, as well as the introduction of the size. In the next section, we therefore keep the model as specified in table 2.1 columns (5) and (10), estimating β, θ, α_i and γ. We will slightly modify the definition of the variables, in order to focus on the spatial distribution of income, not on its overall growth. This is a safer way of considering the future.

Simulations

In this section, we simulate the future spatial distribution of income by estimating a spatial autoregressive process (see Baltagi et al. (2012) and Fingleton (2014)). First, we make in-sample comparisons to assess the fit of our model. Then, we suggest a simulation of the spatial income distribution in 2040/41 based on the 2004/05 - 2014/15 evolution process.

From now on, we express the variables in terms of deviation from the national mean. Thus, we do not rely on any assumption on the overall nominal growth. The other advantage is that spatial disparities are directly comparable across time. Note that since the national evolution is neutralised, time fixed effects become redundant. In the previous sections, we have shown the presence of spatial dependence and individual heterogeneity. We therefore include the neighbours' past income in our previsions, as well as municipal fixed effects (α_i) and size (n_{it}, number of taxpayers). Apart from the fact that variables are expressed in deviation from the national mean and that δ_t is removed, the specification is similar to equation 2.8:

$$\ln \tilde{y}_{it} = c + (1+\beta)\ln \tilde{y}_{it-1} + \theta \sum_{j \neq i}^{N} w_{ij} \ln \tilde{y}_{jt-1} + \gamma \ln \tilde{n}_{it} + \alpha_i + \nu_{it} \tag{2.10}$$

where $\ln \tilde{y}_{it}$ is the percentage deviation from the national mean of the income logarithm, and the logarithm of size $\ln \tilde{n}_{it}$ is expressed in the same way. We estimate coefficients and fixed effects with the Blundell-Bond system-GMM using the same instruments (in deviation from the mean) as in table 2.1, columns (5) and (10).

Table 2.3: Simulation model - System GMM estimators

DV: $\tilde{\text{L}}$og(income)	(1) Period 1971/72 - 1995/96[a]	(2) Period 2004/05 - 2014/15[b]
L.$\tilde{\text{L}}$og(income)	0.644***	0.864***
	(0.115)	(0.0709)
L.W$\tilde{\text{L}}$og(income)	0.313***	0.142**
	(0.113)	(0.0566)
$\tilde{\text{L}}$og(nb of taxpayers)	0.00351**	-0.00113
	(0.00147)	(0.000986)
Constant	0.000164	-0.0000703
	(0.000274)	(0.000286)
Year FE	no	no
N	23 090	11 545
R^2	0.922	0.949

Standard errors in parentheses

* p<.1, ** p<.05, *** p<.01

[a]: Instruments for differenced equation: L(10/.).$\tilde{\text{L}}$og(income), L(12/.).$\tilde{\text{W}}$Log(income, ΔGrowth(nb tp)

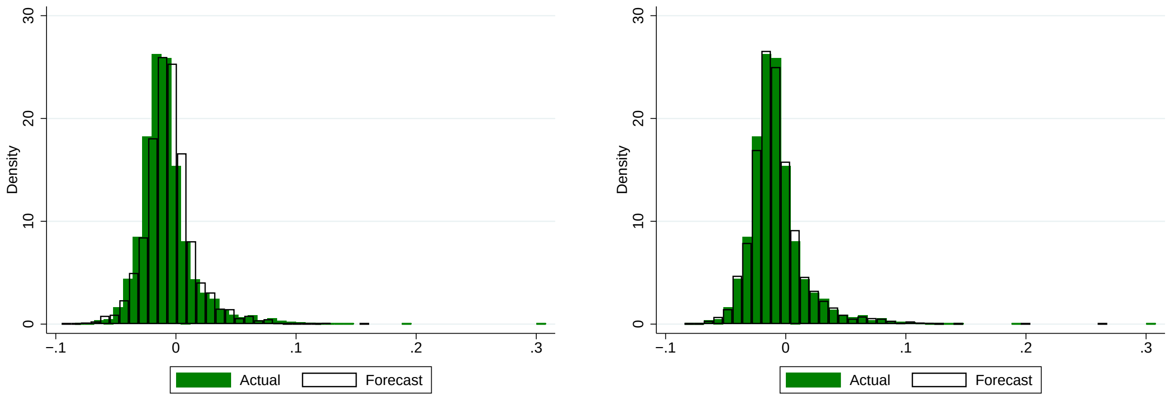
Instruments for level equation: L2.Δ$\tilde{\text{L}}$og(income), L2.Δ$\tilde{\text{W}}$Log(income) and Growth(nb tp)

[b]: Instruments for differenced equation: L(4/.).$\tilde{\text{L}}$og(income); L(4/.).$\tilde{\text{W}}$Log(income) and ΔGrowth(nb tp)

Instruments for level equation: L.Δ$\tilde{\text{L}}$og(income), L.Δ$\tilde{\text{W}}$Log(income) and Growth(nb tp)

We start by showing measures of in-sample fit. In order to obtain a forecast distribution, we estimate the autoregressive process with 1971/72 - 1995/96 data. Estimates can be found in table 2.3, column (1). Then, we apply the coefficients 5 consecutive times to the distribution of 2004/05. We obtain a simulation for the distribution of 2014/15. We add a random error term $\nu_{it} \sim N(0, \hat{\sigma}_\nu)$ at each t, where $\hat{\sigma}_\nu$ is the variance of the estimated error terms $\hat{\nu}_{it}$ in the system GMM specification. We iterate the process 500 times and average the results. The observed distribution of 2014/15 and its forecast can be compared in figure 2.4a. Figure 2.4b reports the distribution forecast when using 2004/05 - 2014/15 data (column (2) of table 2.3) to estimate the autoregressive process.

Table 2.4 reports Spearman and Pearson correlation coefficients between actual and simulated values, as well as the 5th and 95th percentile of their iterated distribution. The last column of the table contains the mean square forecast errors. For comparative purposes, the



(a) 1971/72-1995/96 process

(b) 2004/05-2014/15 process

Figure 2.4: Actual vs simulated distribution 2014/15 - 500 iterations

Note: System-GMM coefficients, time-space recursive model.

first line shows the statistics between the 2004/05 and the 2014/15 income distribution (still in log deviation from the mean).

Table 2.4: Correlations and mean squared forecast error (500 iterations) - income 2014/15

Variable	Estimation sample	Spearman coeff.	Pearson coeff.	MSE
Observed income 2004/05	-	0.9110	0.9052	.0001059
Simulated income from 2004/05 distribution	1971/72-1995/96	.86283 [.86272 - .86294]	.85576 [.85568 - .85583]	.0001620 [.0001619 - 0001621]
Simulated income from 2004/05 distribution	2004/05-2014/15	.99519 [.99516 - .99523]	.99592 [.99591 - .99593]	.00000482 [.00000480 - .00000483]

Note: the table shows the average of the obtained statistics across the 500 iterations, as well as the 5th and 95th percentiles.

Using the 1971/72 - 1995/96 process, the simple correlation between the predicted and the actual income is 0.856, less than the simple correlation between income 2004/05 and income 2014/15 (0.905). This simulation underestimates the increase of the variance. The fit improves a lot when using data from the second subperiod (i.e. 2004/05-2014/15). In this case, the Spearman and Pearson are almost 1 (0.995 and 0.996).

Table 2.5 reports the second, third and fourth moments. The skewness and the kurtosis are lower in the forecast than in reality, but the approximation is far better using the most recent subperiod.

Table 2.5: Second, third and fourth moments - income 2014/15

Distribution	Variance	Skewness	Kurtosis
Actual 2014/15	.000585	2.613	20.785
Simulated 2014/15 (from 1971/72 - 1995/96 process)	.000457	1.163	8.531
Simulated 2014/15 (from 2004/05 - 2014/15) process	.000591	2.347	16.505

Overall, the best in-sample forecast are obtained with coefficients and fixed effects estimated with 2004/05 - 2014/15 figures. High persistence and low mobility within the income ranking explain the good performance of the model. The accurate prediction of the fixed effects is critical: we tested other specifications of the model, by including the square of lagged income, by modelling simultaneous spatial dependence or by approximating the α_i by y_{i0} (see Fingleton (2014)). None of these specifications allows higher in-sample fit measures.

Turning to simulations of the 2040/41 income distribution, we take the 2014/15 log-deviation of income and enter it into the model 13 consecutive times. As for the in-sample forecast exercise, we introduce a random error term and repeat the process 500 times before averaging the results. The resulting spatial income distribution is shown on a map in figure 2.7, which can be compared with the initial income distributions of 1971/72 (figure 2.5) or 2014/15 (figure 2.6). [39] Note that the bounds of the income categories are the same in all maps. The values are set such that the eight classes contain the same number of municipalities in 2014/15.

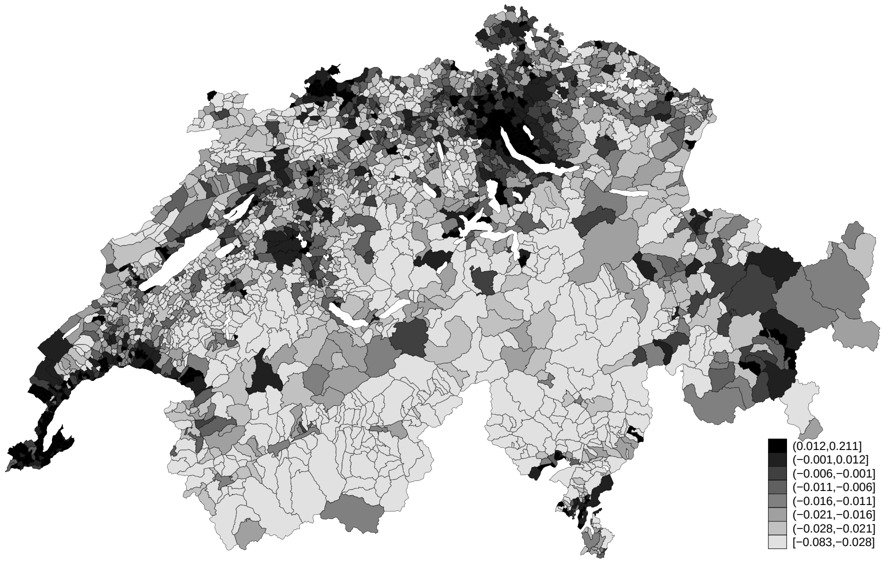


Figure 2.5: Income (Log deviation from the national mean) - 1971/72

Between 1971 and 2015, we can distinguish spatial spillovers around rich regions: the municipalities that became relatively richer over time are those that benefited from the development of the urban and peri-urban areas around Basel, Zurich, Lausanne, Geneva and Lugano. However, between 2014/15 and 2040/41, spatial spillovers are less visible, which reflects the high

39. 1995/96 and 2004/05 are also available in the Appendix, figures 2.8 and 2.9.

persistence and low θ obtained when the process is estimated on 2004/05 - 2014/15 data.

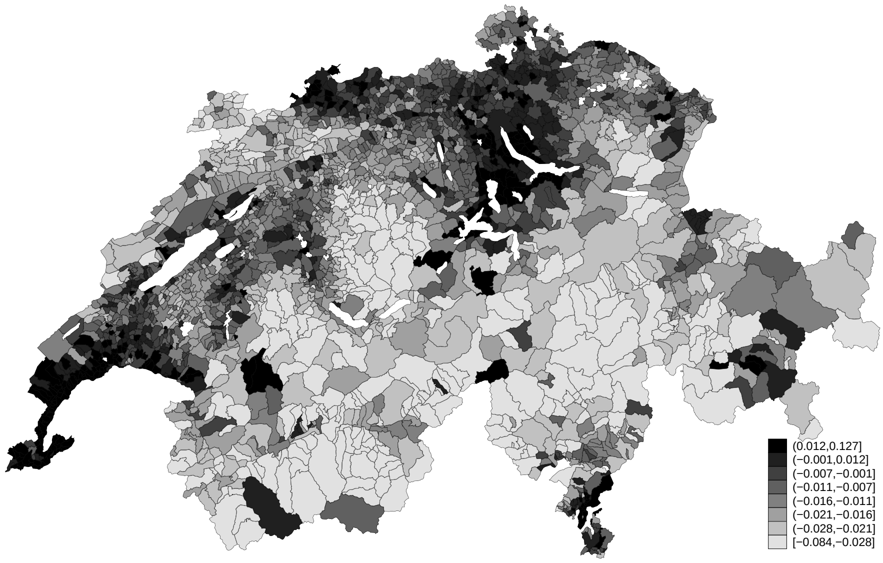


Figure 2.6: Income (Log deviation from the national mean) - 2014/15

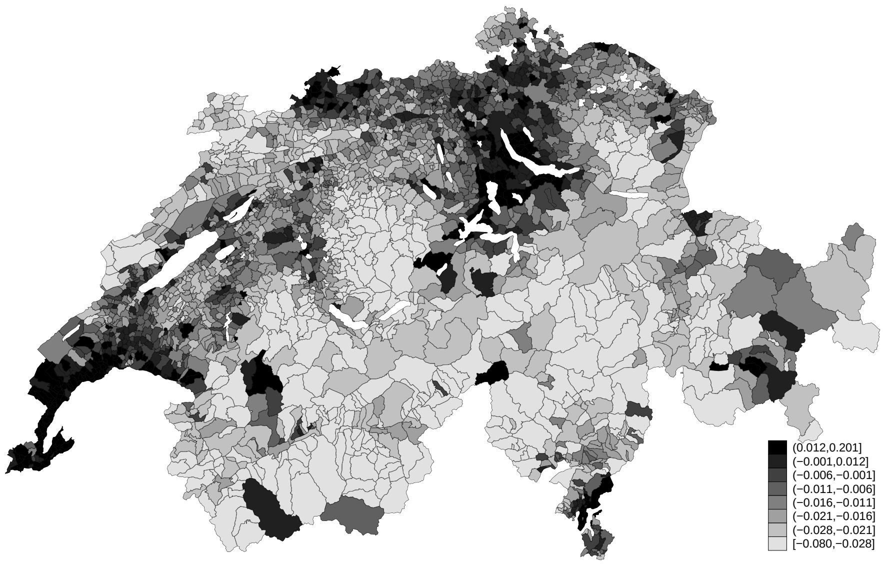


Figure 2.7: Income (Log deviation from the national mean) - 2040/41

Note: Blundell-Bond system GMM estimators, time-recursive model estimated with data 2003-2015 (column (2), table 2.3). 500 iterations.

The dark spots still lie mainly around Lake Zurich, Basel, Lemanic arc, Zug and South of Ticino. On the contrary, the remote regions such as Jura, Valais and North of Ticino seem condemned to stay at the end of the income ranking. To give an idea of the scope of global spatial inequality, we calculate the relative standard deviation and Theil index at unchanged national income. The corresponding Theil index is 0.112 and the relative standard deviation is 0.729, significantly higher than their value in 2014/15 (respectively 0.052 and 0.501).[40] So, despite (conditional) β-convergence and positive spatial autocorrelation, we find little prospects for a reduction of overall inequality (i.e. for σ-convergence) by the early 2040s.

2.6 Conclusion

This chapter presents the evolution of spatial income disparities across Swiss municipalities. Between 1971/72 and 2014/15, the relative standard deviation of income has increased, while spatial autocorrelation has decreased. β-convergence is conditional, meaning that each municipality converges towards its own steady state. Overall, spatial disparities have been persistent during the last 45 years, and especially since 2004/05. The importance of initial heterogeneity relatively to spatial autocorrelation has increased in the 21st century. Finally, according to our simulation, a reduction of spatial inequality at a 25-years horizon is unlikely. Assuming that income distribution follows the 2004-2015 trend, we obtain that income distributions of 2014/15 and 2040/41 are quite similar.

The high persistence of disparities between Swiss municipalities has several possible causes. First, the tax schedule is highly decentralised. As in the Tiebout model of public good provision, each municipality sets its own tax rate and competition among the sub-national entities becomes stronger over time. Then, the natural environment is very different from one place to another. The initial conditions - or first nature factors - are therefore very diverse and the original disparities are reinforced by "path dependency", that are second nature factors. For instance, housing prices are prohibitive in nice areas and rich municipalities are able to set low tax rates. On this last point, the spatial anchoring of income is concomitant with the increase of the variance of tax rates. Finally, distances are short, which makes commuting and moving easy. Jobs are not necessarily located in the same municipality as residences and this is all the more true in a context of decreasing commuting costs.

Our results indicate that the ranking of municipal average incomes will remain essentially the same, unless a strong political willingness to change emerges. The questioning of local fiscal sovereignty however seems a long way off. In a sense, income sorting may be seen as an optimal

40. These indicators have been calculated without Anières (GE), which is an outlier in 2014/15 (with an average net income of CHF 1.5 million).

equilibrium in which each household maximises its utility with respect to location. However, congestion costs, moving barriers, local complementarities in consumption, social spillovers and other market failures raise doubts about the efficiency of sorting, in addition to the obvious fairness issue.

In conclusion, we would like to mention some limitations and possible extensions of this work. Indeed, our analyses are conducted on *pre-tax income* coming from fiscal data, which calls for three additional comments.

First, post-tax income would be more relevant in terms of welfare. Unfortunately, this variable is unavailable at the municipal level. Although income inequalities are likely to be smaller, we expect that, overall, our core results would remain untouched for two reasons. First, the main part of the tax burden variability takes place at the cantonal level [41], while we have shown that within-cantons disparities are preponderant. [42] On the other hand, tax rates tend to be negatively correlated with income levels across municipalities: fiscal decentralisation tends to attenuate income redistribution (see Kübler and Rochat (2017)). Since income inequality comes mainly from high incomes, the inclusion of transfers should not radically change the overall picture either.

Second, the dataset reliably reflects the income of residents, with the notable exception of tax evaders. In this view, our indicators of inequality must be considered as lower bounds, since tax evasion concerns primarily top-incomes (see Alstadsæter et al. (2017)) that are concentrated in rich municipalities.

Lastly, so far, the focus has been on households and residents. Nothing has been said about firms and their location choices. A natural extension, addressed by the next chapter, is to explore how the spatial distribution of income relates to the location of production and value-added.

41. See figure 2.13 in Appendix 2.7

42. See figures 2.2 and 2.3 in section 2.5.

2.7 Appendix

σ evolution

In this subsection, we model heterogeneity (fixed effects) based on Young et al. (2008) and spatial autocorrelation by following Egger and Pfaffermayr (2006). Suppose y_{it} follows an autoregressive process with spatial and fixed effects:

$$\ln y_{it+1} = a_i + (1+\beta)\ln y_{it} + \theta \sum_{j\neq i}^{N} w_{ij} \ln y_{jt} + \varepsilon_{it} \tag{2.11}$$

where a_i is an individual fixed effect. β (θ) the convergence (spatial autocorrelation) coefficient and w_{ij} the i^{th} row and j^{th} column element of the row-standardized Queen contiguity matrix (see footnote 11). ε_{it} is the error terms, independent from y_{it} and a_i. In matrix notation, we have:

$$\mathbf{Y}_{t+1} = \mathbf{A} + (\mathbf{I} + \mathbf{I}\beta + \theta\mathbf{W})\mathbf{Y}_t + \mathbf{\Xi}_t \tag{2.12}$$

Expressing the variables in deviation from the mean, the variance-covariance matrix $E(\tilde{\mathbf{Y}}_{t+1}\tilde{\mathbf{Y}}'_{t+1})$ can be written as:

$$\begin{aligned}
E(\tilde{\mathbf{Y}}_{t+1}\tilde{\mathbf{Y}}'_{t+1}) = \mathbf{\Omega}_{t+1} =& E[(\tilde{\mathbf{A}} + (\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})\tilde{\mathbf{Y}}_t + \tilde{\mathbf{\Xi}}_t)(\tilde{\mathbf{A}} + (\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})\tilde{\mathbf{Y}}_t + \tilde{\mathbf{\Xi}}_t)'] \quad (2.13)\\
=& \overbrace{E(\tilde{\mathbf{A}}\tilde{\mathbf{A}}')}^{\mathbf{\Omega}_A} + \overbrace{(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})(\mathbf{I}-\mathbf{I}\beta+\theta\mathbf{W})'E(\tilde{\mathbf{Y}}_t\tilde{\mathbf{Y}}'_t)}^{(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})'\mathbf{\Omega}_t} + \overbrace{E(\tilde{\mathbf{\Xi}}_t\tilde{\mathbf{\Xi}}'_t)}^{\mathbf{\Omega}_\Xi}\\
&+ \underbrace{2E(\tilde{\mathbf{A}}\cdot[(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})\tilde{\mathbf{Y}}_t]')}_{\mathbf{\Sigma}_{Y_t,A}} + \underbrace{2E(\tilde{\mathbf{A}}\cdot\tilde{\mathbf{\Xi}}'_t)}_{\mathbf{\Sigma}_{A,\Xi_t}} + \underbrace{2E((\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})\tilde{\mathbf{Y}}_t\cdot\tilde{\mathbf{\Xi}}'_t)}_{\mathbf{\Sigma}_{Y_t,\Xi_t}}\\
=& \mathbf{\Omega}_A + (\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})'\mathbf{\Omega}_t + \mathbf{\Omega}_\Xi + \mathbf{\Sigma}_{Y_t,A} + \mathbf{\Sigma}_{A,\Xi_t} + \mathbf{\Sigma}_{Y_t,\Xi_t}\\
\cong& \mathbf{\Omega}_A + (\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})(\mathbf{I}+\mathbf{I}\beta+\theta\mathbf{W})'\mathbf{\Omega}_t + \mathbf{\Omega}_\Xi + \mathbf{\Sigma}_{Y_t,A}
\end{aligned}$$

where the last line is obtained by assuming that the two covariances involving the error term are negligible.

Income distribution

Some complementary statistics are shown in this section. First, the summary statistics are shown in table 2.6. Average and median income are annual, expressed in CHF, weighted by the number of taxpayers, net and in equivalent terms.

Table 2.6: Summary statistics

Ass. period	Variable	Obs	Mean	Std. Dev.	Min	Max
1971/72	Mean income	2 309	19 025.9	4 496.609	8 339	150 770
	Median income	2 309	14 626.65	1 815.626	6 811.388	27 814
	Number of taxpayers	2 309	877.429	4 595.27	3	156 747
1973/74	Mean income	2 309	21 406.93	4 573.284	9 381	89 271
	Median income	2 309	16 966.73	2 087.445	6 889	30 133
	Number of taxpayers	2 309	986.499	4 897.331	5	166 384
1975/76	Mean income	2 309	22 568.94	4 737.019	9 031	130 308
	Median income	2 309	18 326.36	2 399.682	5 000	32 202
	Number of taxpayers	2 309	1 044.229	4 960.874	7	170 716
1977/78	Mean income	2 309	23 478.66	4 910.007	8 866	97 880
	Median income	2 309	19 030.89	2 573.656	5 600	36 065
	Number of taxpayers	2 309	1 095.627	5 055.854	10	174 478
1979/80	Mean income	2 309	24 735.78	5 295.878	9 153	92 235
	Median income	2 309	20 039.34	2 837.017	5 800	38 792
	Number of taxpayers	2 309	1 149.738	5 142.745	13	177 485
1981/82	Mean income	2 309	27 414.17	5 967.138	8 717	166 811
	Median income	2 309	22 239.39	3 182.744	4050	44 476
	Number of taxpayers	2 309	1 203.112	5 296.137	10	182 107
1983/84	Mean income	2 309	29 659.54	6 537.703	9 917	176 640
	Median income	2 309	23 969.35	3 434.615	4350	49 033
	Number of taxpayers	2 309	1 252.66	5 392.084	12	185 708
1985/86	Mean income	0	-	-	-	-
	Median income	0	-	-	-	-
	Number of taxpayers	0	-	-	-	-
1987/88	Mean income	2 309	32 564.42	7 678.681	10 272	144 140
	Median income	2 309	26 071.33	3 717.322	2 200	53 533
	Number of taxpayers	2 309	1 341.641	5 517.967	12	188 632
1989/90	Mean income	2 309	35 607.86	8 635.046	11 956	171 155
	Median income	2 309	28 440.95	3 954.102	3 150	55 762
	Number of taxpayers	2 309	1 391.482	5 612.289	14	191 334
1991/92	Mean income	2 309	39 304.4	9 098.487	12 771	217 851
	Median income	2 309	31 729.1	4 427.312	4 000	60 000
	Number of taxpayers	2 309	1 410.902	5 601.249	15	191 364
1993/94	Mean income	2 309	40 445.05	8 918.005	14 541	205 637
	Median income	2 309	32 672.81	4 428.914	5 300	58 167
	Number of taxpayers	2 309	1 464.148	5 578.393	15	190 301
1995/96	Mean income	2 309	40 816.79	8 844.19	15 662	192 729
	Median income	2 309	33 095.6	4 420.437	5 600	59 576
	Number of taxpayers	2 309	1 460.004	5 501.096	6	188 337

Ass. period	Variable	Obs	Mean	Std. Dev.	Min	Max
2004/05	Mean income	2 309	45 222.7	11 324.12	18 552	213 943.5
	Median income	2 309	35 793.94	4 850.78	7 850	71 289
	Number of taxpayers	2 309	1 519.963	5 429.75	11	185 884.5
2006/07	Mean income	2 309	47739.5	14196.57	18 885	219 979
	Median income	2 309	37 438.82	5 245.218	7 700	76 867
	Number of taxpayers	2 309	1 536.942	5 457.686	11.5	187 758
2008/09	Mean income	2 309	50 121.04	15 600.61	21 269.5	404 923
	Median income	2 309	38 846.61	5 556.125	9 700	83 254
	Number of taxpayers	2 309	1 582.073	5 608.288	11.5	193 848
2010/11	Mean income	2 309	52 186.36	19 077.79	21825.5	396 393.5
	Median income	2 309	40 916.31	6 096.841	9 500	78 687
	Number of taxpayers	2 309	1 545.895	5 481.271	10.5	189 624
2012/13	Mean income	2 309	53 213.85	15 969.61	20 221	479 303.5
	Median income	2 309	41 363.03	6 175.324	9 500	79 333
	Number of taxpayers	2 309	1 509.351	5 351.768	7.5	187 010.5
2014/15	Mean income	2 309	54 616.85	27 345.53	21 486	1 428 595
	Median income	2 309	41 814.39	6 172.741	10 150	78 600
	Number of taxpayers	2 309	1 556.77	5 475.839	9.5	191 275

In complement to figures 2.5 and 2.6, figures 2.8 and 2.9 show the spatial distribution in 1995/96 and 2004/05. Income is expressed in log deviation from the mean. The bounds are the values that allocate the same number of municipalities in the eight different classes in 2014/15.

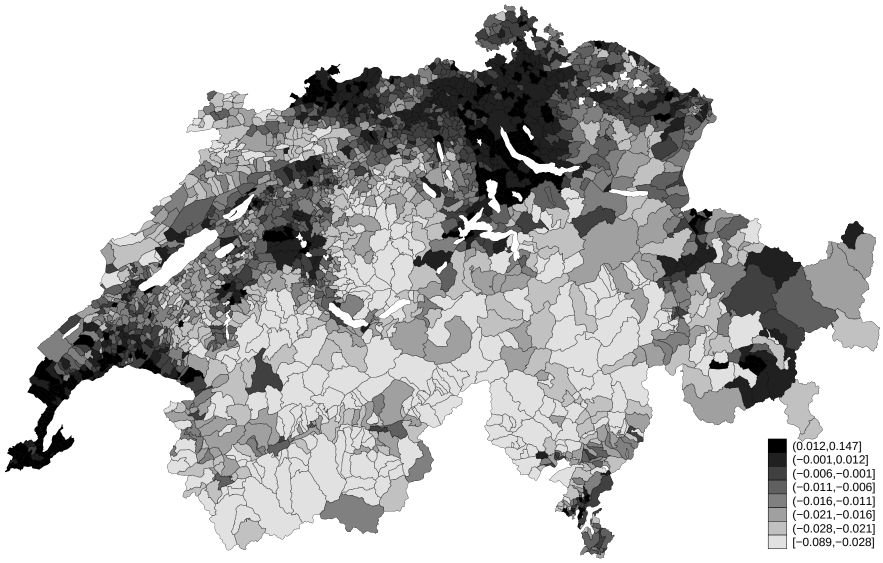


Figure 2.8: Income (Log deviation from the national mean) - 1995/96

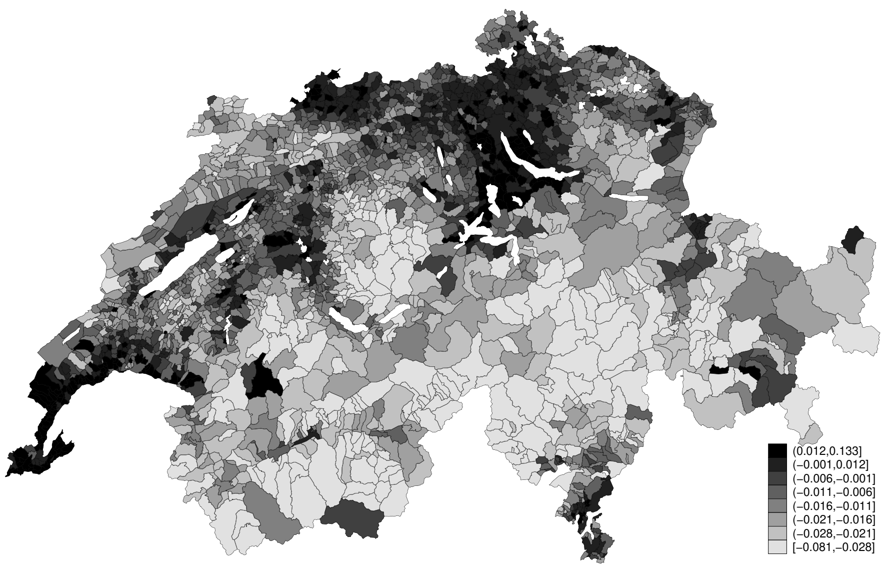
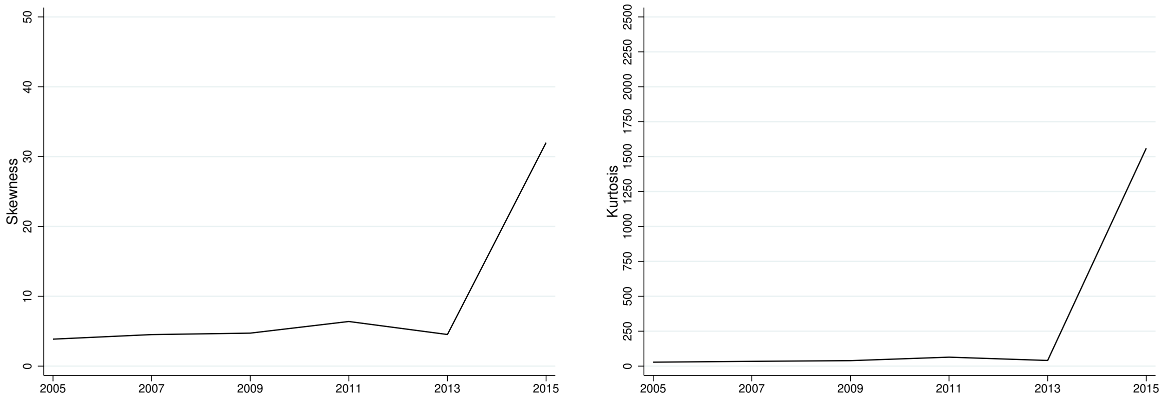


Figure 2.9: Income (Log deviation from the national mean) - 2004/05

Figure 2.10 shows the skewness and kurtosis of income distribution between 2004 and 2015. Similarly to variance, the third moment (skewness) shows peaks in 2010/11 and 2014/15, revealing that the asymmetry has been reinforced in favour of density on the left. The fourth moment also increases sharply in 2014/15, meaning that the mass around the mean has increased. Altogether, this reflects the elongation and the thickening of the right tail: rich municipalities became even richer.



(a) Skewness (b) Kurtosis

Figure 2.10: Income distribution 2004/05-2014/15

Finally, figure 2.11 shows the evolution of the relative standard deviation of median income among Swiss municipalities. The magnitude of the changes is smaller than the one observed on average income, since the influence of outliers is minimised. The indicator still increased by 20% between 1971/72 and 2014/15. Interestingly, there is an upward trend in the first decade of the period which was not observable on the standard deviation of the average income. It may be explained by an increase of inequality *within* municipalities, that has widened the gap between average and median dispersion. As shown in Levy (1997), convergence of average income is neither necessary nor sufficient to ensure median convergence.

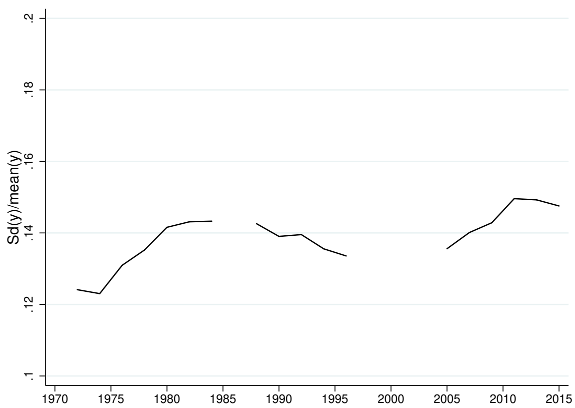
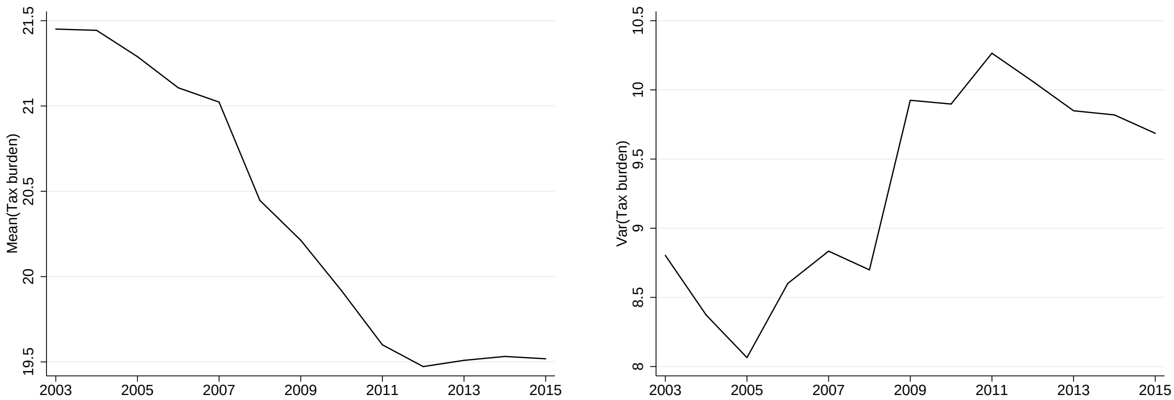


Figure 2.11: Relative standard deviation of median income 1971/72 - 2014/15

Taxes

Additional data on tax burdens used in this subsection come from the website of the Federal Tax Administration (AFC, 2018a). The tax burden is defined as the amount of cantonal, municipal and church income taxes, expressed in percentage of the gross earned income.

Figure 2.12a and 2.12b show respectively the average and the variance of the tax burden on married couples without children with gross earned income of 50 000 CHF over time in the 800 most populated Swiss communes.



(a) Average

(b) Variance

Figure 2.12: Tax burden 2003-2015

Note: the statistic is calculated on the 800 largest municipalities (no data for small municipalities until 2008).

The map of tax burdens in 2015 for a rich household (married, without children earning 50 000 CHF per year) is shown in figure 2.13. The tax burden shows substantial variation, it goes from 10 to 23%. The most important visual contrasts are inter-cantonal. To some extent, the map is the mirror of the cantonal income levels (with some exceptions like Geneva). Bern, Jura and Neuchâtel charge the highest burden (all of them are well beyond the national average income in 2015), while the two richest cantons Zoug and Schwyz are also the most advantageous in terms of taxes.

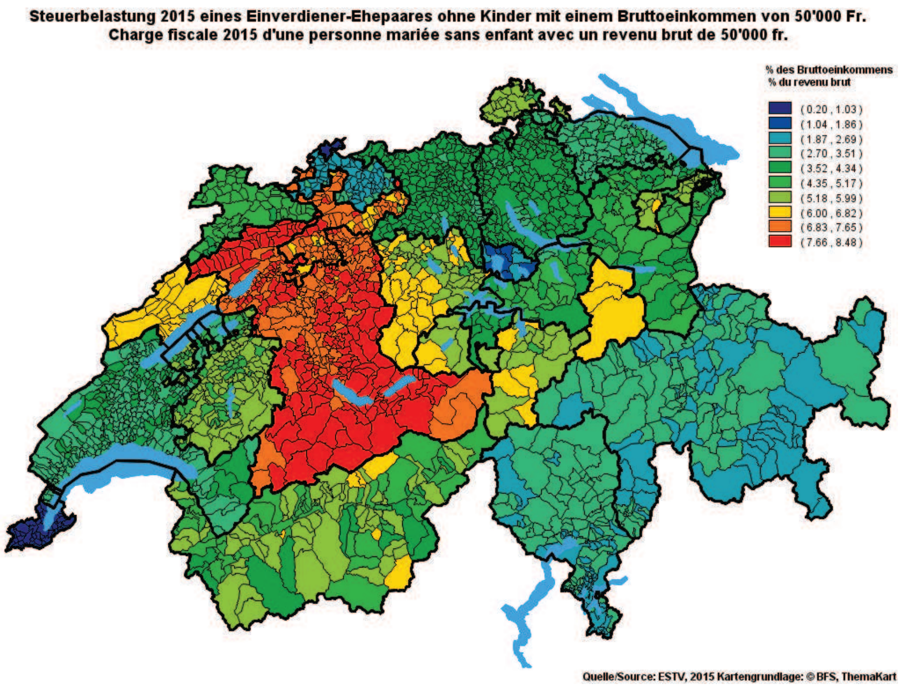
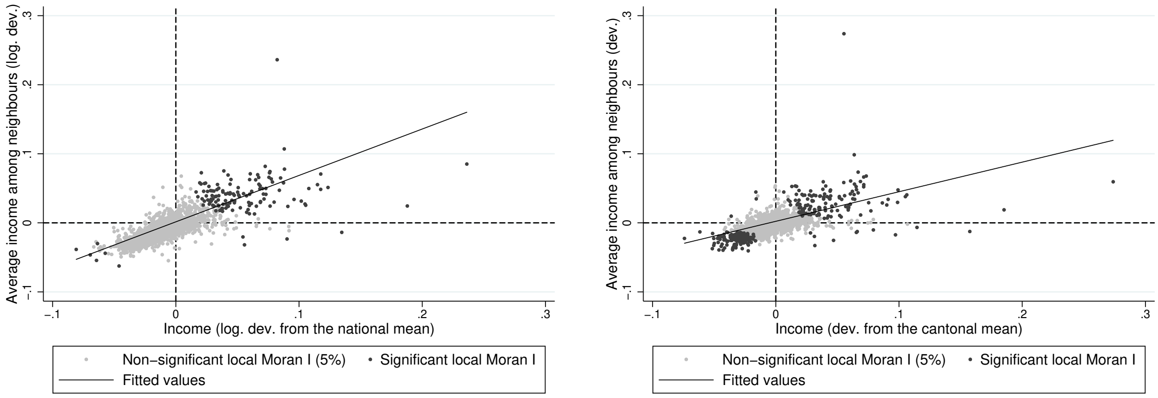


Figure 2.13: Tax burden in 2015

Spatial statistics

The Moran scatterplot shows income on the horizontal axis (in log deviation from the national mean) and the same variable averaged across the contiguous municipalities on the vertical axis. If the local Moran's I shows positive local autocorrelation [43], the municipality is represented by a black dot. The points are well aligned on the line with a positive slope of almost 1. There is clear positive spatial autocorrelation: one municipality tends to be as rich (poor) as its

43. There is local autocorrelation if the local Moran's I $I_i = \dfrac{(y_i - \bar{y})\sum_{j=1}^{N} w_{ij}(y_j - \bar{y})}{\frac{1}{N}\cdot\sum_{i=1}^{N}(y_i - \bar{y})^2}, i \neq j$ is significantly different (i.e. p-value< 0.05, normal distribution) from $E_0(I_i) = \dfrac{\sum_{j=1}^{N} w_{ij}}{N-1}$, the expectation under the absence of autocorrelation (see Anselin (1995) and Gallo and Ertur (2003)).



(a) Income relative to the national mean

(b) Income relative to the cantonal mean

Figure 2.14: Moran Scatterplot 2014/15

neighbourhood.

The right part of figure 2.14 shows the Moran scatterplot when the income is expressed as a percentage deviation from the *cantonal* average. We therefore get rid of the correlation which is due to some common geographical or institutional factors. A clear positive relationship subsists.

Figure 2.15 shows to which quadrant of the left diagram (figure 2.14a) the municipality belongs. Rich communes with rich neighbourhood are in black, poor municipalities with poor neighbourhood are in light grey. Most municipalities are in these two categories and only few are rich with a poor neighbourhood (High-low) or conversely.

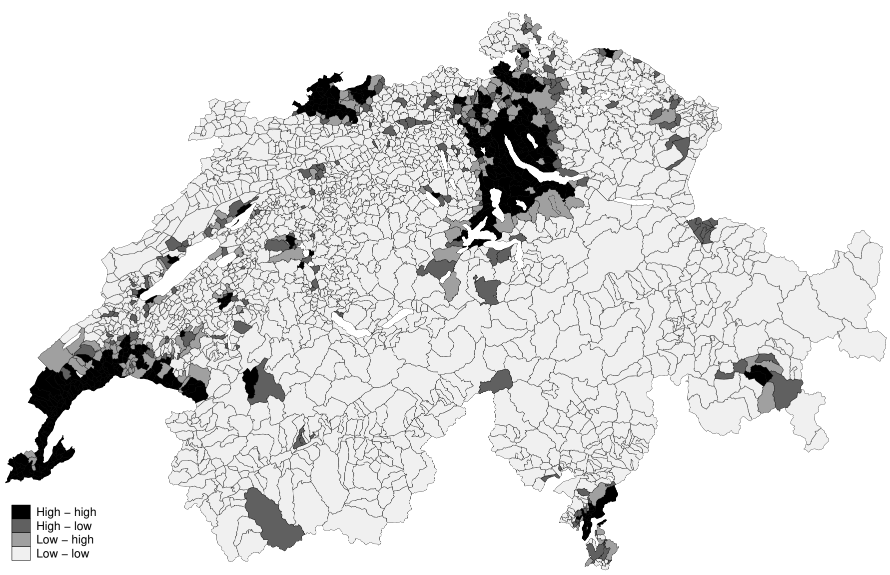


Figure 2.15: Moran map 2014/15 (Income and neighbours' income)

Table 2.7 shows the details of Moran I calculations. Each year, the indicator of global spatial autocorrelation is positive and significantly different from 0.

Table 2.7: Evolution of the Global Moran I

Year	I	E(I)	Sd(I)	Z	P-val
1971/72	.368	0	.012	29.874	0.000
1973/74	.422	0	.013	33.192	0.000
1975/76	.426	0	.013	33.851	0.000
1977/78	.479	0	.013	37.686	0.000
1979/80	.524	0	.013	41.047	0.000
1981/82	.488	0	.013	38.626	0.000
1983/84	.541	0	.013	42.784	0.000
1985/86	-	-	-	-	-
1987/88	.574	0	.013	45.113	0.000
1989/90	.578	0	.013	45.419	0.000
1991/92	.507	0	.013	40.003	0.000
1993/94	.505	0	.013	39.77	0.000
1995/96	.482	0	.013	37.878	0.000

Year	I	E(I)	Sd(I)	Z	P-val
2004/05	.531	0	.013	41.673	0
2006/07	.527	0	.013	41.385	0
2008/09	.481	0	.013	38.277	0
2010/11	.431	0	.013	34.197	0
2012/13	.437	0	.012	35.024	0
2014/15	.223	0	.009	24.184	0

Indicators of mobility association

The mobility of municipalities within the income ranking brings complementary information on spatial heterogeneity and dependency. Following Rey (2016), we calculate the global indicator of mobility association, which reflects to what extent the income ranking changes from period t to period $t+1$. The statistics are given by:

$$M_t = \frac{\tau(y_t, y_{t+1}) - 1}{-2} = \frac{d_t}{N(N-1)/2} \tag{2.14}$$

$$\text{where: } \tau(y_t, y_{t+1}) = \frac{\sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} sign(y_{it} - y_{jt}) \cdot sign(y_{it+1} - y_{jt+1})}{N(N-1)/2} = \frac{c_t - d_t}{N(N-1)/2}$$

The $sign()$ function results in -1 when the element in bracket is strictly negative, 1 when it is strictly positive and 0 when it is null. c_t is therefore the number of concordant pairs (i.e. that stay in the same order) and d_t the number of discordant pairs. $N(N-1)/2$ is the number of possible pairs of municipalities i and j, given the total of N municipalities. The correlation coefficient ($-1 \leq \tau(y_t, y_{t+1}) \leq 1$) and the mobility index ($0 \leq M_t \leq 1$) are spatially decomposable. In particular we can rewrite $\tau(y_t, y_{t+\Delta})$ in order to separate neighbours and non-neighbours:

$$\tau(y_t, y_{t+1}) = \overbrace{\frac{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1} w_{ij}^{NB} \cdot sign(y_{it} - y_{jt}) \cdot sign(y_{it+1} - y_{jt+1})}{N(N-1))/2}}^{\text{Neighbours component}}$$
$$+ \underbrace{\frac{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1}(1 - w_{ij}^{NB}) \cdot sign(y_{it} - y_{jt}) \cdot sign(y_{it+1} - y_{jt+1})}{N(N-1))/2}}_{\text{Non-neighbours component}} \tag{2.15}$$
$$= \phi\tau^{NB}(y_t, y_{t+1}) + (1-\phi)\tau^{NNB}(y_t, y_{t+1})$$

where $w_{ij}^{NB} = 1$ if i and j are neighbours and 0 otherwise. ϕ is therefore the share of the total numbers of existing pairs that are contiguous. Since M_t is additive, we also have:

$$M_t = \phi M_t^{NB} + (1-\phi)M_t^{NNB} \tag{2.16}$$

The evolution of M_t, also known as the communal general indicator of mobility association (GIMA), is shown in figure 2.16. The solid curve shows the statistic calculated among all pairs and the dashed line only on contiguous ones.[44]

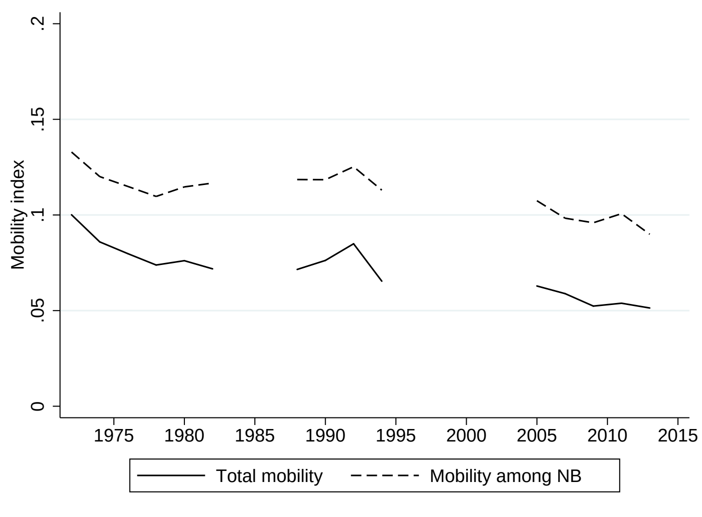


Figure 2.16: Communal mobility index 1971/72 - 2014/15

Figure 2.16 shows three important facts. First, mobility is quite low. Only 5 to 10% of the

44. Given that the total number of possible relationships is extremely high ($2.66 \cdot 10^6$) compared to the limited number of neighbours (13 010), the total GIMA corresponds also to the GIMA among non-neighbours ($\phi \cong 0$).

pairs change of relative positions in the ranking. Second, the indicator of spatial association is decreasing over time. From 1971/72 to 2012/13, the GIMA went from 0.1 to 0.051, which represents a fall of 49%. Third, mobility is substantially higher among neighbours than among non-contiguous municipalities, meaning that it is easier for one municipality to catch up (or to be caught up by) the surrounding communes.

Table 2.8 shows indicators of mobility association within each cantons. The canton of Jura and Schaffhouse show particularly high value of spatial local mobility association all over the period. Fribourg have been hotspots of mobility until 1994/95. In general, the local mobility seems higher in rural and remote cantons.

Table 2.8: Local Indicator of Mobility Association (by canton)

Ass. period	ZH	BE*	LU	UR	SZ	OW	NW	GL	ZG	FR	SO	BS	BL*
1975/76	0.056	0.088	0.072	0.058	0.083	0.190	0.055	0.000	0.055	0.123	0.087	0.000	0.070
1977/78	0.064	0.075	0.066	0.121	0.074	0.095	0.073	0.000	0.145	0.120	0.077	0.000	0.056
1979/80	0.066	0.088	0.071	0.121	0.053	0.048	0.055	0.000	0.073	0.135	0.095	0.333	0.072
1981/82	0.064	0.099	0.052	0.079	0.126	0.048	0.073	0.000	0.127	0.121	0.075	0.000	0.074
1983/84	-	-	-	-	-	-	-	-	-	-	-	-	-
1985/86	-	-	-	-	-	-	-	-	-	-	-	-	-
1987/88	0.062	0.089	0.067	0.079	0.051	0.095	0.073	0.000	0.145	0.143	0.079	0.000	0.092
1989/90	0.075	0.090	0.064	0.079	0.034	0.048	0.073	0.000	0.200	0.205	0.087	0.333	0.079
1991/92	0.067	0.086	0.065	0.074	0.055	0.095	0.036	0.000	0.164	0.190	0.083	0.000	0.089
1993/94	0.075	0.082	0.054	0.084	0.067	0.000	0.036	0.000	0.164	0.122	0.067	0.000	0.073
2004/05	0.056	0.073	0.054	0.047	0.037	0.048	0.036	0.000	0.109	0.077	0.069	0.000	0.068
2006/07	0.047	0.072	0.055	0.047	0.025	0.095	0.018	0.000	0.055	0.078	0.059	0.000	0.065
2008/09	0.047	0.077	0.060	0.079	0.025	0.143	0.018	0.000	0.036	0.088	0.057	0.000	0.052
2010/11	0.053	0.074	0.050	0.053	0.039	0.095	0.018	0.000	0.036	0.081	0.069	0.000	0.070
2012/13	0.057	0.071	0.040	0.084	0.032	0.048	0.018	0.000	0.018	0.084	0.063	0.000	0.044

Ass. period	SH	AR	AI	SG	GR	AG	TG	TI	VD	VS	NE	GE	JU
1975/76	0.175	0.126	0.067	0.087	0.090	0.075	0.076	0.078	0.098	0.104	0.102	0.074	-
1977/78	0.162	0.158	0.000	0.079	0.103	0.080	0.104	0.083	0.104	0.075	0.056	0.061	-
1979/80	0.134	0.084	0.067	0.077	0.098	0.089	0.079	0.077	0.104	0.078	0.062	0.083	0.141
1981/82	0.111	0.058	0.000	0.075	0.097	0.076	0.084	0.097	0.090	0.077	0.048	0.080	0.156
1983/84	-	-	-	-	-	-	-	-	-	-	-	-	-
1985/86	-	-	-	-	-	-	-	-	-	-	-	-	-
1987/88	0.105	0.074	0.000	0.072	0.077	0.096	0.121	0.093	0.085	0.065	0.146	0.070	0.193
1989/90	0.089	0.063	0.000	0.094	0.093	0.089	0.087	0.087	0.092	0.073	0.104	0.047	0.143
1991/92	0.114	0.100	0.067	0.083	0.085	0.101	0.109	0.089	0.092	0.058	0.099	0.046	0.157
1993/94	0.114	0.116	0.133	0.110	0.080	0.097	0.099	0.055	0.103	0.053	0.081	0.045	0.141
2004/05	0.111	0.089	0.067	0.243	0.074	0.074	0.073	0.068	0.056	0.058	0.050	0.035	0.138
2006/07	0.102	0.111	0.067	0.077	0.082	0.057	0.096	0.064	0.055	0.068	0.045	0.027	0.148
2008/09	0.114	0.079	0.133	0.069	0.089	0.062	0.073	0.058	0.060	0.071	0.054	0.045	0.165
2010/11	0.102	0.121	0.067	0.062	0.093	0.076	0.077	0.053	0.054	0.071	0.068	0.041	0.122
2012/13	0.077	0.042	0.000	0.050	0.066	0.067	0.072	0.064	0.050	0.061	0.071	0.035	0.125

The district of Laufen is treated as if it has belonged to Basel-Landschaft since 1975 though it was in canton Bern until 1994.
Indication of reading: between 1989/90 and 1991/92, one over the three pairs of municipalities in Basel-Stadt flipped in the ranking.

Table 2.9 presents the mobility indicators of cantonal and communal level. Figure 2.17 shows the General mobility indicator M_t calculated at the cantonal level. The statistics are calculated only from the assessment period 1975/76, from which the canton of Jura is in the dataset. The trend at the cantonal level is unclear. Mobility is rather small in general, it attains 0.028 between 2012/13 and 2014/15 (only 2.8% of the possible pairs of cantons are not in the same order in

Table 2.9: General Indicator of Mobility Association (Communal or cantonal level)

	Communal Level		Cantonal Level		
Ass. period	M_t^{NB}	$M_t^{NNB} = M_t$ *	M_t^{NB}	M_t^{NNB}	M_t
1971/72	0.133	0.100	-	-	-
1973/74	0.120	0.086	-	-	-
1975/76	0.115	0.080	0.071	0.063	0.065
1977/78	0.110	0.074	0.089	0.033	0.043
1979/80	0.115	0.076	0.000	0.059	0.049
1980/82	0.117	0.072	0.018	0.030	0.028
1983/84	-	-	-	-	-
1985/86	-	-	-	-	-
1987/88	0.119	0.072	0.089	0.052	0.058
1989/90	0.118	0.076	0.036	0.061	0.057
1991/92	0.125	0.085	0.125	0.121	0.122
1993/94	0.113	0.065	0.018	0.045	0.040

	Communal Level		Cantonal Level		
Ass. period	M_t^{NB}	$M_t^{NNB} = M_t$ *	M_t^{NB}	M_t^{NNB}	M_t
2004/05	0.107	0.063	0.071	0.082	0.080
2006/07	0.098	0.059	0.054	0.048	0.049
2008/09	0.096	0.052	0.036	0.033	0.034
2010/11	0.101	0.054	0.036	0.030	0.031
2012/13	0.090	0.051	0.000	0.033	0.028

*: the total number of possible relationships is extremely high ($2.66 \cdot 10^6$) compared to the limited number of neighbours (13 010), the total GIMA corresponds therefore to the GIMA among non-neighbours ($\phi \cong 0$).

2012/13 and in 2014/15). Differences of mobility between neighbours and non-neighbours are not systematic.

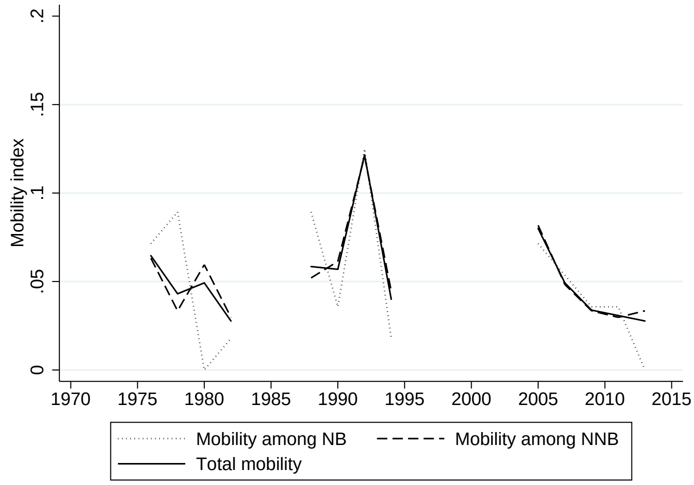


Figure 2.17: Cantonal mobility index 1976/77 - 2014/15

Chapter 3

The spatial mismatch between income and production: an analysis on Swiss municipalities *

3.1 Introduction

In a well connected area like Switzerland, with large commuting opportunities, income does not necessarily end up where it is created. Firms' and households' location choices are obviously linked, but they are motivated by different factors. Typically, firms tend to cluster due to agglomeration economies and households due to income sorting. These forces have important policy implications and have been consequently addressed by New Economic Geography models. Key parameters have been identified, such as commuting and trade costs, but as often, the final equilibrium patterns depend on the parameter values, which opens the field to empirical studies.

The present study describes the spatial mismatch between production and income in Switzerland in a novel way. We gather and combine data on income, value-added and employment at the municipal level. We uncover several empirical signals of spatial specialisation. First, Theil index decompositions show that spatial disparities exist and come mainly from within-canton inequalities, confirming that the location process takes place at the municipal level. Second, series of plots and maps show that income and value-added are differently distributed across

∗. This paper is co-authored by Benjamin Tissot-Daguette (PhD, Visiting Lecturer, École hôtelière de Lausanne).

space. Third, we identify that the Swiss economic landscape is characterised by "productive" centres being surrounded by "residential" belts. These patterns are consistent with NEG models, in particular the framework developed by Borck et al. (2010), which constitutes our theoretical benchmark.

The spatial mismatch between production and income has important implications in terms of efficiency and equity across regions. On the one hand, the issues of transport, congestion and commuting costs arise when jobs and residences are not located in the same place. This involves externalities and public goods issues that may be at the origin of substantial welfare losses. On the other hand, it is hard to avoid the question of the spatial distribution fairness, especially regarding municipalities where wealth is produced, but from where income is fleeing. The tax base and the ability to spend money at the local level crucially depend on the type of economic activity implemented in a particular area, whereas public investments influence in turn the possibility of regional development. Thus, our empirical results should be of interest for the ongoing negotiations on fiscal equalisation and geographical equity among Swiss municipalities and cantons.

After an overview of the current relevant literature, this chapter presents the data used, the empirical framework and the results. Before coming to its conclusion, the chapter presents also some robustness checks.

3.2 Literature Review

New Economic Geography (NEG)

We are interested in what makes a region rich or poor, we have therefore to explore the attraction forces behind income and production. The concentration of income arises because households sort themselves. Tiebout (1956) shows how heterogeneity of preferences and combinations of taxes and public goods lead to a sorted equilibrium in a federal state. Oates (1969) highlights the role of properties value in reinforcing the sorting process. Brueckner et al. (1999) underlines the role of amenities, arguing that a nice natural, cultural and historical environment is a luxury good. In a federalist state, we should therefore observe a concentration of rich people in nice places, where taxes are low and housing prices are high. In Switzerland, the role of taxation has been widely explored, including Schmidheiny (2006), Stadelmann and Billon (2012) and Schaltegger et al. (2011).

On the production side, since the seminal work of Marshall (1920), the empirical literature has firmly established that industries benefit from spatial concentration through productivity gains. But what hides beyond agglomeration economies? A useful preliminary step towards answering, as described by Rosenthal and Strange (2004), is to distinguish localisation economies on the one hand, resulting from the spatial concentration of a specific industry, from urbanisation economies on the other hand, resulting from the spatial concentration of economic activity as a whole. This fundamental distinction is now widely accepted in the literature.

To analyse firms' and households' locations together, the New Economic Geography (NEG) has developed very fertile frameworks. The seminal paper by Krugman (1991) presents a general equilibrium framework, where households and firms make simultaneous choices. Because of Marshallian externalities, firms locate where other firms are. Since workers follow job opportunities, and because workers are consumers, it creates an additional local demand, which attracts new firms. Muth (1971) and Mazek and Chang (1972) already pointed out this chicken or egg puzzle. Concentration arises, until congestion costs cancel out the advantage of proximity.

More recently, the NEG literature has produced several comprehensive formalisations of why economic agents tend to agglomerate (e.g. Fujita et al. (1999), Fujita and Thisse (2002), Baldwin et al. (2005), Behrens et al. (2014)). This field has also evolved towards quantitative modelling, similarly to the earlier evolution of the trade literature (see e.g. Eaton and Kortum (2002)). The most recent models aim, in particular, to a better fit with the observed data (e.g. Allen and Arkolakis (2014) or Caliendo et al. (2015)). In explaining localisation choices, this literature highlights the distinction between "first nature" factors (physical endowments) and "second nature" factors (agglomeration forces through externalities). A large empirical literature has emerged in order to deepen our understanding of these "second nature" factors.

Given our research question, an important limitation of these NEG models is the no-commuting assumption: workers usually have residence and job in the same region, for simplification concerns. The theoretical model by Borck et al. (2010) is a notable exception. The authors combine the agglomeration economies assumption of Krugman (1991) with the housing market modelling of Helpman (1998). In the case of low commuting and transport costs, they find that the equilibrium involves concentration of firms and dispersion of households at the same time. The empirical findings of this chapter can be explained by the processes involved in the model of Borck et al. (2010), which is presented in more details in the subsection "Theoretical model".

Income vs GDP

At the macroeconomic level, the spatial mismatch between value creation and earnings is nothing but the distinction between product and income. Most macroeconomic studies on spatial issues (growth, inequality or convergence between regions) focus on GDP (e.g. Gallo and Ertur (2003), Crozet (2004), López-Bazo et al. (2004)). However, since workers and capital cross borders, GDP is an imperfect indicator of the earnings of residents. The well-known Stiglitz report (Stiglitz et al., 2007) makes clear that the depreciation of capital, the balance of primary incomes and the evaluation of public expenditure should be taken into account when we evaluate the economic performance of a country. He recommends the use of national income (i.e. net value-added on a residence basis) instead of GDP (i.e. gross value-added on a territory basis).

Using income instead of GDP makes a difference regarding distributional issues. Milanovic (1999) measures world inequality by using income data rather than GDP per capita and finds larger indicators of cross-country inequality. Martin (2009) describes inequality between European regions using the two measures. He finds that in France, spatial inequalities are larger if they are measured by GDP than by income, while in the UK, this is the opposite.

Most of the evidence so far regarding differences between the spatial distribution of income and that of production has been gathered at the macro level. Within-country empirical studies are more scant (e.g. Lee and Lin (2017) for the US or Geary and Stark (2016) for Great Britain). The present chapter takes advantage of the availability of income and value-added data at the level of Swiss municipalities to perform for the first time a systematic analysis within that country.

Theoretical Model

We present here the theoretical framework of Borck et al. (2010), which may help to explain our empirical observations. These authors suggest a model which has two important features for our problematic. First, it is a general equilibrium model that makes endogenous the location of both firms and households, a common characteristics of NEG models. Second, and more innovatively, the model allows for commuting. Borck et al. (2010) recognise that agglomeration economies drive the location decisions of firms (Krugman, 1991), but they also account for the importance of the housing market in the choice of residence (Helpman, 1998).

The general setup consists of two regions, with L units of land. There are two sectors:

— Industry (X), which is characterised by increasing returns to scale, monopolistic competition and iceberg trade costs. Industry employs skilled and unskilled workers.

— Agriculture (A), which is characterised by constant returns to scale, perfect competition and tradable goods without any transport costs. Agriculture only needs unskilled labour.

Unskilled workers are mobile between sectors but immobile between regions. The reverse is true for skilled workers. Borck et al. (2010) assume a quasi-linear utility function for individuals, which depends on their consumption of X, A and L. Agents solve the maximisation problem in three stages: first workers choose the place of residence, then skilled workers choose whether they commute or not, third, individuals choose how much they work and consume. In the long run, the utility level should be the same in every regions (so that nobody wishes to migrate), population and workforce are therefore endogenously determined.

Borck et al. (2010) present their predictions in three cases: zero commuting costs, prohibitive commuting costs and positive but non-prohibitive commuting costs. For given preference parameters and available land in each region, the location choices of firms and households, as well as the commuting patterns, are only determined by the degree of trade freeness. In particular, in the (most realistic) case of non-prohibitive commuting costs, the equilibrium involves concentration of firms and dispersion of households when the degree of trade freeness is high. On the contrary, households tend to concentrate and firms to disperse when trade costs are high.

Applying these conclusions to Swiss municipalities, activities with a high degree of trade freeness like manufacturing should concentrate in municipalities that offer opportunities of agglomeration economies. The skilled and mobile workers tend to choose a residence location in a neighbouring municipality, from where commuting distance is not prohibitive and where the housing market is not congested. This should lead to the formation of productive centres with high value-added, surrounded by a belt of municipalities hosting the skilled labour force.

On the other hand, the development of sectors with high transport costs is a concentration force for households and a dispersion force for firms. This may explain why local services are spread out and represent a higher share in the creation of value in the "residential" type of municipalities. It also results in urban centres attracting households in spite of high housing rents.

Cities provide local services, that are not substitutable with what surrounding municipalities can offer.

3.3 Data

We use two data sources: the Federal Tax Administration (FTA) and the Federal Statistical Office (FSO).

Income Data

Income data come from the FTA. They are calculated on the basis of Federal Direct Tax returns (see AFC (2013)). The dataset contains annual average and total income in each municipality, in net and taxable terms.[46] The average is calculated per taxpayer (per household). In the robustness section, we also use the net benefits of legal entities aggregated by municipality. These data have been initially gathered thanks to the SNF project "The Swiss Confederation: A Natural Laboratory for Research on Fiscal and Political Decentralization", led by Marius Brülhart, Monika Bütler, Mario Jametti and Kurt Schmidheiny from 2010 to 2016.

Value-added Data

Data on value-added are derived from the WS survey database of the FSO[47] over the period 2011-2015. Due to confidentiality and feasibility constraints, we are forced to use data aggregated at the "pseudo-firm" level. A "pseudo-firm" is characterised by one municipality, one 4-digits NOGA sector and one legal form. In the process of data gathering, we need to address two major issues: first, due to the survey structure of the WS, some firms have missing values for given years. Indeed, 20% of the firms are replaced every year (this roll-over applies in particular to small firms) and some others simply do not respond. Secondly, in the context of our study, we are interested in the location of the productive unit rather than in the location of the administrative unit. In other words, plant level data are more adapted than firm level data. To tackle both issues, multiple imputation techniques are used and extensively described in Tissot-Daguette and Grether (2019). This methodology generates a total of 400 datasets, each being composed

46. *Net* income is the income after having removed insurance and savings interests. *Taxable* income is about 30% lower than the net income, as it excludes all the fiscal deductions.

47. "Statistique de la Valeur Ajoutée - OFS". This survey excludes the primary sector and the banking and financial services sector.

by the same "pseudo-firms" (whose creation relies on the same plants) but with different imputed value-added figures. In this study, we perform the analysis on each of these datasets and present the average results and the uncertainty around the imputed values.

Moreover, we use the accounting results of Swiss companies presented in the report of the FSO (OFS, 2018a) to estimate capital depreciation at the 2-digits NOGA level. These statistics are also calculated by the FSO on the basis of the WS survey.

Employment Data

In addition to the two mentioned datasets, we gather data on employment using the STATENT census[48] database from the FSO. STATENT is used in order to create the 400 datasets of value-added and to derive total employment (in full-time equivalent) in each municipality.

3.4 Empirical Framework

The Production/Income ratio

We aim to express the relative value-added created in one municipality compared to the amount earned by its residents. We need a simple statistic that translates the mismatch between production and income by municipality. The most evident indicator is the ratio between these two variables:

$$\text{ratio} = \frac{\text{VA}}{\text{INC}}$$

where VA stands for value-added and INC for income. If the ratio is higher (lower) than the Swiss average ratio, the municipality produces, relative to income, more (less) wealth compared to the national average. We are therefore able to identify easily "productive" and "residential" municipalities. The distribution of this ratio is extensively described in section 3.5.

Theil decompositions

After the description of the value-added to income ratio, we perform an analysis of its spatial dispersion. As mentioned above, disparities between municipalities may come from value-added

48. "Statistique Structurelle des Entreprises"

or income, not only in levels but also in per capita (or per household) terms. Productivity inequalities are due to Marshallian externalities: firms tend to agglomerate in productive centres. Income inequalities come from the sorting process: rich households tend to concentrate in nice places where taxes are low and housing rents are high. Moreover, both variables interact: firms are interested in locating near workers and consumers pools. Meanwhile, consumption opportunities matter in the households' choice of residence. To quantify these two effects and their interaction, we perform a Theil decomposition.[49] In particular, the Theil index of the ratio (T_r) can be decomposed into net contributions of per household value-added (v), of the inverse of average income per household (y^{-1}) and of the interaction between the two (Ω). With $\bar{x}$ being the average of any variable x, the decomposition can be expressed as follows:[50]

$$
\begin{aligned}
& T_r = \gamma T_v + \gamma T_y + \Omega \\
\text{where} \quad & r = v \cdot y^{-1} \\
& \Omega = \ln \gamma + \gamma \Big(cov(\frac{v_i}{\bar{v}} \ln(\frac{v_i}{\bar{v}}), \frac{y_i^{-1}}{\bar{y}^{-1}}) + cov(\frac{y_i^{-1}}{\bar{y}^{-1}} \ln(\frac{y_i^{-1}}{\bar{y}^{-1}}), \frac{v_i}{\bar{v}}) \Big) \\
& \gamma = 1 - \frac{cov(v, y^{-1})}{\bar{r}}
\end{aligned}
$$

This first decomposition shows how the spatial economic disparities in the value-added-to-income ratio come from agglomeration economies on the one hand, and from the income sorting process on the other hand. In order to investigate whether the location process of firms and households takes place at the municipal level or at the cantonal one, we perform a second decomposition. In particular, we separate the Theil index into the within and between components for any subgroup:

$$
T_r = \underbrace{\sum_{j=1}^{M} s_j \cdot T_{r_j}}_{within} + \underbrace{\sum_{j=1}^{M} s_j \cdot \ln(\frac{\bar{r}_j}{\bar{r}})}_{between}
$$

$$
\text{where} \quad s_j = \frac{N_j \bar{r}_j}{N \bar{r}}
$$

49. We thank Jean-Marie Grether to have suggested that decomposition.

50. See the development in technical Appendix 3.8. r is equivalent to our ratio in the previous subsection, since the numerator and denominator are divided by the same factor. We use per capita terms in this decomposition to abstract from the scale effect.

where M is the number of cantons and N_j the number of observations in canton j. This second decomposition gives indications on the role of cantonal policies and the importance of distance-related costs in the location model. Notably, the within-cantons variation corresponds to the low commuting cost case described in section 3.2. The between inequality indicator will reflect to what extent cantons are also able to specialise. Inequality in the ratio may be particularly salient between small cantons (in surface area) like Geneva at one extreme (high ratio), and Schwyz at the other (low ratio).

These two decompositions can be run separately to investigate the drivers of spatial inequality in the value-added-to-income ratio. Does spatial inequality come from per capita value-added, average income or their interaction (first decomposition)? And is it rather due to intra- or inter-cantonal disparities (second decomposition)? The two decompositions can also be run jointly. This gives rise to the two-way matrix decomposition shown in table 3.1. [51]

Table 3.1: Two-way Matrix decomposition

	v	y	$Cov.$	Total
Within	$\sum_{j=1}^{M} s_j \gamma_j \frac{\bar{y}^{-1}}{\bar{y}_j^{-1}} T_{v_j}$	$\sum_{j=1}^{M} s_j \gamma_j \frac{\bar{v}}{\bar{v}_j} T_{y_j}$	$\sum_{j=1}^{M} s_j \Omega_j$	$\sum_{j=1}^{M} s_j T_{r_j}$
Between	$\sum_{j=1}^{M} s_j \frac{\gamma_j}{\gamma} \frac{\bar{y}^{-1}}{\bar{y}_j^{-1}} \ln\left(\frac{v_j}{\bar{v}_j}\right)$	$\sum_{j=1}^{M} s_j \frac{\gamma_j}{\gamma} \frac{\bar{v}}{\bar{v}_j} \ln\left(\frac{y_j^{-1}}{\bar{y}_j^{-1}}\right)$	$\sum_{j=1}^{M} s_j \Theta_j$	$\sum_{j=1}^{M} s_j \ln\left(\frac{\bar{r}_j}{\bar{r}}\right)$
Total	γT_v	γT_y	Ω	T_r

Where the within and between interaction terms, Ω_j and Θ_j, are given by

$$\Omega_j = \ln(\gamma_j) + cov\left(\frac{v_i}{\bar{v}} \ln\left(\frac{v_i}{\bar{v}}\right), \frac{y_i^{-1}}{\bar{y}^{-1}}\right) + cov\left(\frac{y_i^{-1}}{\bar{y}^{-1}} \ln\left(\frac{y_i^{-1}}{\bar{y}^{-1}}\right), \frac{v_i}{\bar{v}}\right) + \left[1 - \frac{\bar{y}^{-1}}{\bar{y}_j^{-1}}\right] \frac{\gamma_j}{\gamma} T_{v_j} + \left[1 - \frac{\bar{v}}{\bar{v}_j}\right] \frac{\gamma_j}{\gamma} T_{y_j}$$

$$\Theta_j = \ln\left(\frac{\gamma}{\gamma_j}\right) + \left[1 - \frac{\bar{y}^{-1}}{\bar{y}_j^{-1}}\right] \frac{\gamma_j}{\gamma} \ln\left(\frac{v_j}{\bar{v}_j}\right) + \left[1 - \frac{\bar{v}}{\bar{v}_j}\right] \frac{\gamma_j}{\gamma} \ln\left(\frac{y_j^{-1}}{\bar{y}_j^{-1}}\right)$$

Spatial analysis of the Production/Income ratio

We characterise the dispersion of the ratio across space using visualisation tools such as plots, maps and Moran scatter plots. In addition, we present the correlations between our ratio and the distribution of employment across industries. We should find higher ratios municipalities where agglomeration economies are important (see Tissot-Daguette (2019)) and where capital

51. See technical Appendix 3.8 for the detailed demonstration.

depreciation is high. On the contrary, the ratio should be lower in "residential" municipalities that attract rich people, and where local services are overrepresented (see chapter 1).

3.5 Results

Summary statistics

Table 3.2 presents the summary statistics of the main variables. Value-added is higher than net income on average, the mean of the ratio is between 1.29 and 1.33. This does not come as a surprise since the value-added is not entirely redistributed to shareholders. In particular, a part of the firm's earnings is dedicated to finance its own reserves and depreciation of capital. However, since data on "gross" value-added are available at a more disaggregated level, we conduct the first analyses without having removed depreciation of capital. Thus, we overestimate the value creation and, consequently, the ratio. According to Swiss National Accounts statistics, the capital depreciation, relative to "gross" value-added, was 20.8% on average over the years 2011-2015. Robustness tests are run in section 3.6, in which we use the rate of capital depreciation over value-added at the NOGA 2 level to obtain the net value-added, before calculating the ratio. Summary statistics of this alternative ratio is shown in table 3.12 of the Appendix.

The ratio of gross value-added to income has high variability across space, the minimum corresponds to the municipality Petit-Val (BE) and the maximum to Manno (TI). Variances are relatively large, with a standard deviation exceeding the mean for all variables. Moreover, the distribution of the production/income ratio is skewed to the left. [52]

Value-added and income seem quite stable over time. The same comment holds for the standard deviation, except for income (and consequently the ratio) in 2014. The increase in the variance is drawn by outliers, typically small municipalities. Anières, for instance, became ten times richer in 2014 compared to 2013 because one taxpayer earned a very high additional income (see Mabut (2015)).

52. The entire distribution of the ratio can be found in the Appendix, figure 3.12.

Table 3.2: Summary statistics

Variable	Year	Mean	Min	Median	Max	St. Dev.
Value-added*,†	2011	340 272.72	37.78	63 807.79	47 476 244.85	1 818 879.89
		(64.45)	(0.86)	(57.30)	(32364.37)	(988.54)
	2012	332 016.30	30.27	61 706.12	45 802 739.79	1 801 187.84
		(62.34)	(0.25)	(51.18)	(31786.18)	(957.75)
	2013	337 518.58	46.75	63 022.46	46 491 096.11	1 786 180.25
		(72.49)	(1.42)	(51.21)	(34804.81)	(1134.48)
	2014	346 478.18	4.09	65 366.33	47 531 862.49	1 814 765.50
		(58.87)	(0.00)	(50.05)	(27177.25)	(973.66)
	2015	347 554.75	89.63	63 091.95	49 037 735.71	1 908 821.42
		(53.99)	(1.54)	(59.49)	(30587.50)	(1033.16)
Income†	2011	176 731.76	3 868.80	76 061.78	16 911 838.19	582 799.33
	2012	176 515.85	3 596.67	76 646.63	16 986 498.63	581 377.24
	2013	181 161.47	3 301.33	79 046.47	17 310 091.68	595 074.27
	2014	185 285.87	3 572.27	80 346.07	17 772 252.31	612 262.09
	2015	188 478.43	3 757.13	82 155.08	18 694 467.63	632 915.94
Nb. of households	2011	2 736.45	62.00	1 216.00	249 257.00	8 835.13
	2012	2 750.76	64.00	1218.50	251 287.00	8 877.41
	2013	2 797.35	57.00	1245.00	254 108.00	9 002.98
	2014	2 819.83	61.00	1272.00	254 158.00	9 060.66
	2015	2 868.15	65.00	1286.50	263 358.00	9 302.38
Ratio*	2011	1.3381	0.0010	0.8351	20.8182	1.8316
		(0.0007)	(0.0000)	(0.0005)	(0.1315)	(0.0028)
	2012	1.2754	0.0005	0.8045	20.0877	1.7521
		(0.0005)	(0.0000)	(0.0005)	(0.1083)	(0.0022)
	2013	1.2977	0.0008	0.8148	21.3009	1.8303
		(0.0006)	(0.0000)	(0.0007)	(0.1180)	(0.0023)
	2014	1.3253	0.0003	0.8001	25.2474	1.9236
		(0.0005)	(0.0000)	(0.0006)	(0.2726)	(0.0041)
	2015	1.2890	0.0010	0.7828	24.5699	1.8660
		(0.0006)	(0.0000)	(0.0006)	(0.0912)	(0.0021)

Notes: * In parenthesis, the standard error of the given statistics average (i.e. statistics are computed on each of the 400 datasets and the mean and standard deviation of the mean of each statistics are reported).
† In thousands of CHF.

Theil decompositions

Table 3.3 shows the decomposition of the Theil index for the last year of our sample.

Comparing the table's rows, the first lesson is that within-cantonal inequality is much larger than the between disparities. 92% of the value-added per household disparities and 63% of the income inequality take place within cantons. The process of specialisation occurs mainly at the

Table 3.3: Two-way Matrix decomposition for 2015

	VPH*	IPH*	Cov.	Total
Within	0.58941	0.01310	-0.03913	0.56338
	(0.01000)	(0.00002)	(0.00272)	(0.00845)
Between	0.04461	0.00779	-0.01838	0.03402
	(0.00250)	(0.00001)	(0.00132)	(0.00192)
Total	0.63402	0.02089	-0.05752	0.59739
	(0.01073)	(0.00003)	(0.00353)	(0.00878)

Notes: * VPH is value-added per household and IPH inverse of income per household.
See Table 3.1 for the analytical expressions of the decomposition.

municipal level, which corresponds to the low commuting costs case of section 3.2.

Looking at the columns, we notice that value-added per household is far more uneven than average income. This holds within and between cantons. It appears that the forces behind firms concentration dominate the ones behind income sorting of households. This is consistent with one of the predictions of Borck et al. (2010): when commuting costs are non-prohibitive and trade costs are low, firms concentrate, while households spread. The covariance component is negative because of the positive relationship between the average income and value-added per household.[53] Large cities are for instance richer both in term of income and value-added. Similar figures appear for the other years (tables 3.8 – 3.11 in the Appendix). The observed facts seem stable over time.

Spatial analysis of the Production/Income ratio

The first set of graphs shows how do the distributions of value-added and income differ in general. The results for 2015 are shown in figure 3.1, whereas the same plots for 2011 are reported in the Appendix, figure 3.10. On the top graph, the points tend to align on the diagonal, showing that the richer municipalities in terms of income are also the richer in value-added terms. This reflects the scale effect. A radically different picture appears on the second graph, when we plot the variables in per capita terms. The richest households do not necessarily live in the municipalities with the highest average production. We also observe that the larger points lie in the upper-right part of each graph. This indicates that large urban centres attract both rich households (because of urban amenities) and productive firms (because of agglomeration

53. If the denominator tends to be large when the numerator is large, then the variability of the ratio is lower than if both variable are totally independent.

economies).

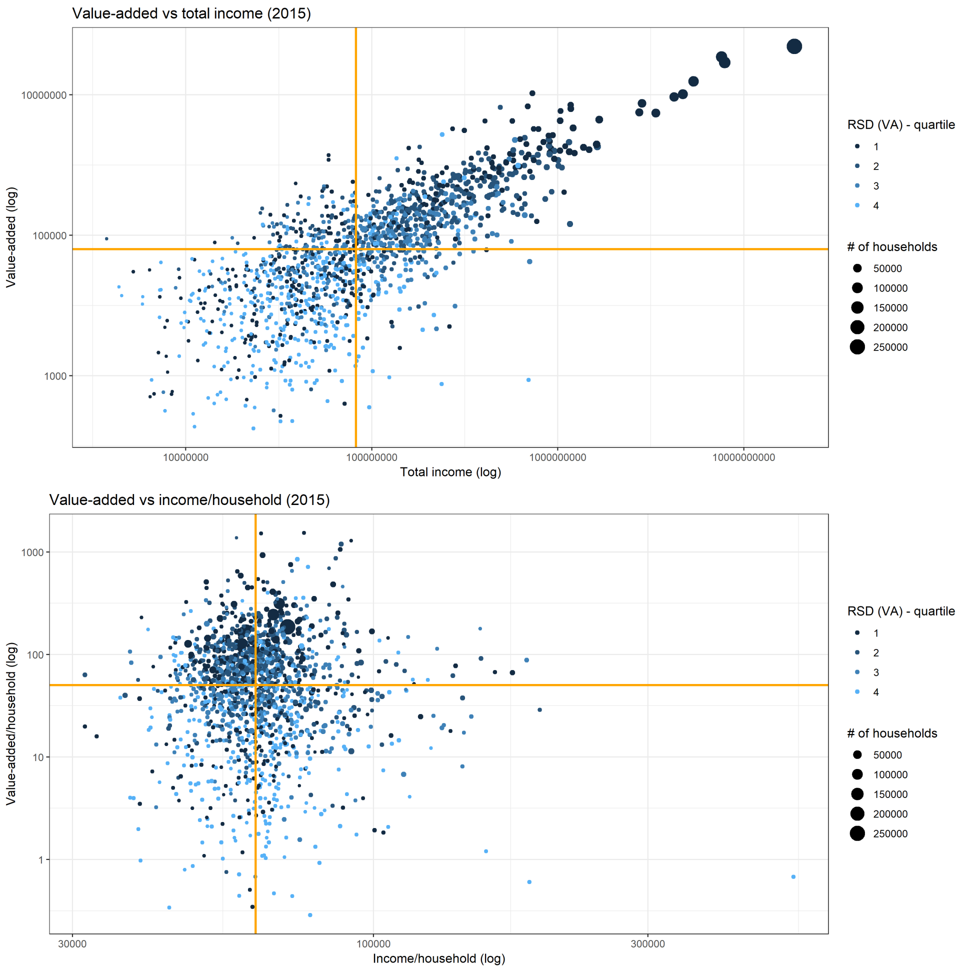


Figure 3.1: Income and value-added (total and per household) - 2015

Note: value-added (vertical axis) and income (horizontal axis) are represented on a logarithmic scale. Orange lines are drawn at the median. The colour represents the relative standard deviation of value-added across the 400 databases. The darker the point, the lower the relative standard deviation.

Figure 3.2 shows the spatial distribution of the ratio in 2015. The maps representing this ratio in the other years, as well as income and value-added are in the Appendix, figures 3.13, 3.14 and 3.15.

The highest ratios are found in either dense populated areas (e.g. Geneva (GE), Zurich (ZH), Basel (BS), Bern (BE)) or municipalities where sectors with high value-added are prominent, for instance the tobacco and gold production in Neuchâtel (NE) and the watch industry in la Vallée de Joux (VD). Using the typology of Rosenthal and Strange (2004), the cities profit from urbanisation economies, while the specialised municipalities are characterised by localisation

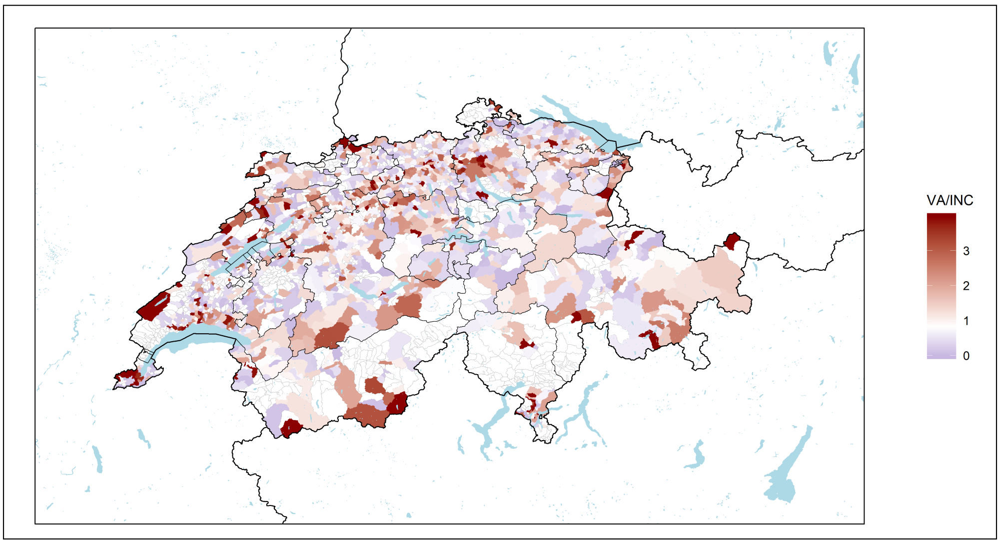


Figure 3.2: Spatial distribution of the ratio VA/INC

economies.[54] Towns next to the borders tend to have high ratios as well (e.g. Le Locle (NE), Boncourt(JU) and Sennwald (SG)). In those areas, firms benefit from the labour pool located in other countries. The access to the workforce is indeed one of the three types of Marshallian externalities giving a relative advantage to the border regions.

Oppositely, the low-ratio municipalities are typically located in a nice natural environment. For instance, the proximity to a lake is a prominent factor to attract rich households (e.g. the North of lakes Biel and Thun or the South of lakes Neuchâtel and Lucerne). This is in line with the observation made in chapter 1.

Turning to the geographical interdependence between high- and low-ratio municipalities, we observe a pattern of "surrounding belts", that is high-ratio municipalities whose most neighbours are low-ratio municipalities. This corresponds to the non-prohibitive commuting and low trade costs case in Borck et al. (2010): skilled workers are dispersed where the housing market is less congested and into municipalities that offer valuable amenities (among which accessibility, see chapter 1). This pattern can be observed around the largest cities such as Zurich (ZH), Basel (BS), Bern (BE), Lugano (TI) and Geneva (GE), but also around smaller productive centres such as Biel (BE), St-Imier (BE) or Monthey (VS).

The hypothesis of residential municipalities surrounding high-ratio units can be further explored with Moran scatterplots. Figure 3.3 represents the variable of interest in one municipality

54. By using data on employment, Tissot-Daguette (2019) has shown the importance of both effects in the Swiss economy. On average urbanisation economies seem to be prevalent in Switzerland, which might be due to the relative high urban density of the country.

(horizontal line) plotted against the same variable found on average in contiguous municipalities. The orange lines represent the median of each variable. We express income and value-added in per capita terms, in order to get rid of the scale effect. There is clear positive spatial autocorrelation in terms of income, with a slope of almost 1: a given municipality tends to be as rich as its neighbours. They indeed share many common characteristics (e.g. canton and natural environment).

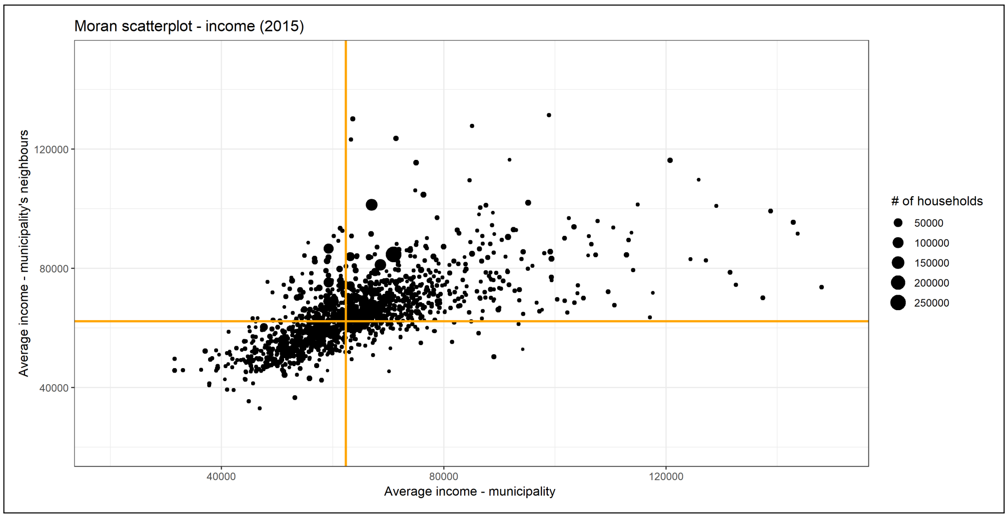


Figure 3.3: Moran Scatterplot - Income per taxpayer

However, we do not find such positive spatial spillovers in the value-added per household. The absence of spatial autocorrelation indicates that the process of firms concentration takes place mainly at the scope of the municipality. The large municipalities lie in the upper-right quadrant. Indeed, urbanisation economies often cross the municipal borders in these cases (e.g municipalities around Lausanne, Basel and Zurich). The picture remains the same as we add the sectoral dimension. The Moran scatterplots representing productivity by NACE sectors in 2015 are available in Appendix 3.8. We can not distinguish any positive relationship on these graphs.[55]

Positive autocorrelation seems also absent from the ratio, despite the spatial dependence of income. The residential municipalities lie on the left and spread out in the bottom and upper quadrant. This suggests that they are surrounded by both residential municipalities (low ratio) and productive centres (high ratio). The points on the right, which tend to be larger, represent productive centres. When situated in the upper quadrant, they suggest spatial spillovers across

55. The same conclusion holds considering NOGA 3 sectors (results available upon request).

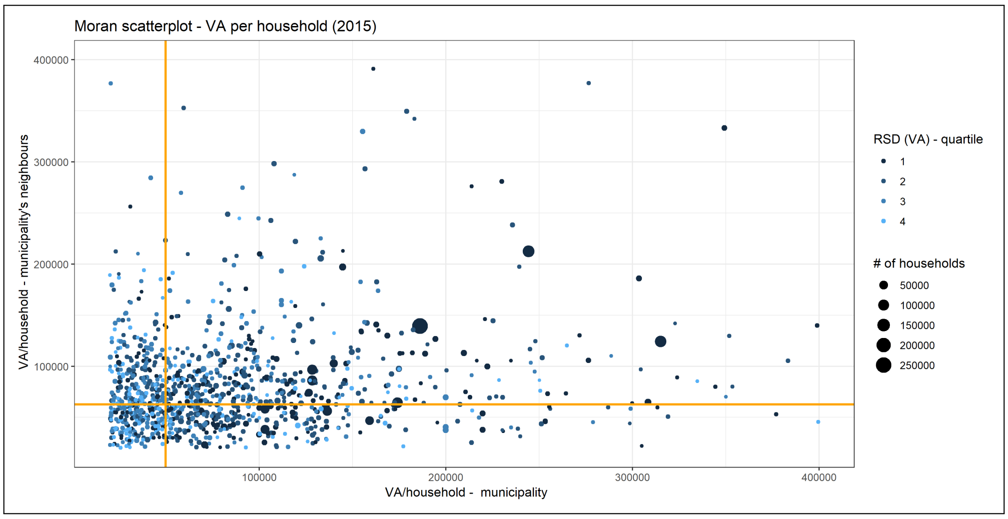


Figure 3.4: Moran Scatterplot - Value-added per household

municipal borders.

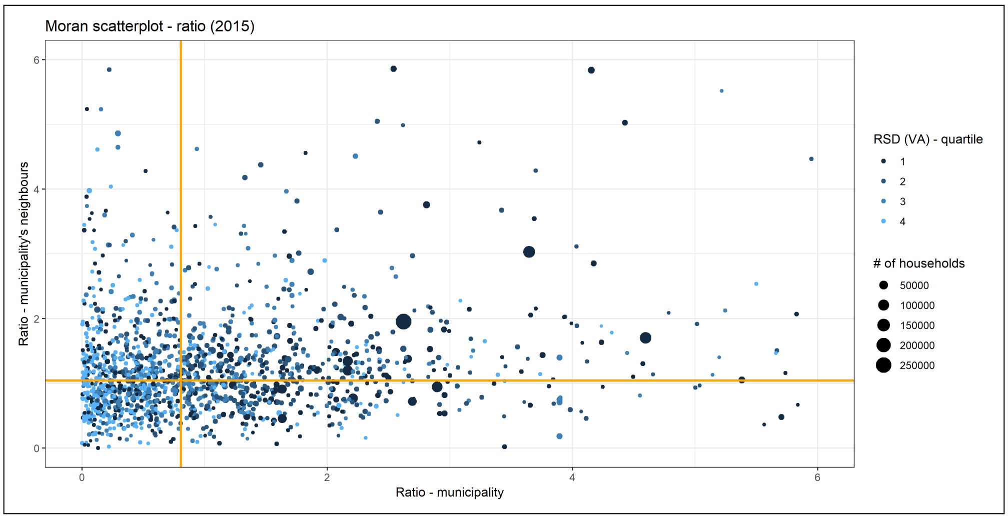


Figure 3.5: Moran Scatterplot - Ratio VA/INC

Moran scatterplots are in line with the "surrounding belt" hypothesis. As in the maps, large productive centres seem surrounded by residential municipalities. The former are likely to benefit from agglomeration economies, which depend on the industrial structure of the given municipality (see Tissot-Daguette (2019)). That is why, we explore the link between the ratio and the industry specialisation of a municipality in table 3.4. We compute the share of employment (Share_{is}) in each industrial section by municipality. [56] Then we calculate the correlation between our ratio and the share of employment in each industry. [57]

56. $\text{Share}_{is} = \text{FTE}_{is}/\text{FTE}_i$ where FTE_{is} is the number of full time-equivalents employed by sector s in municipality i and FTE_i is the total number of full time-equivalents in municipality i.

57. Since part of the value-added is imputed and leads to 400 different value-added samples, we again have 400

Table 3.4: Correlation between the share of employment and the VA/INC ratio, by NACE sections

Industrial classification (NACE)*		Year				
Code	Description	2011	2012	2013	2014	2015
2a	Manufacturing industries	0.10838	0.11929	0.11485	0.11064	0.11001
		(0.00754)	(0.00636)	(0.00567)	(0.00793)	(0.00597)
2	Extraction and other industries	-0.04667	-0.050790	-0.04591	-0.04903	-0.05064
		(0.00212)	(0.00212)	(0.00221)	(0.00270)	(0.00230)
3	Building sector	0.01580	0.00116	0.01300	0.03890	0.02643
		(0.00878)	(0.00614)	(0.00646)	(0.00938)	(0.00693)
4	Whole and retail sales, transport, hotels and restaurant, TIC	0.02190	0.02620	0.00798	-0.01130	0.00216
		(0.00748)	(0.00515)	(0.00593)	(0.00510)	(0.00676)
7	Real estate	-0.04923	-0.06047	-0.06960	-0.05287	-0.06297
		(0.01145)	(0.00358)	(0.00232)	(0.00352)	(0.00175)
8	Scientific and technical activities, administrative services	-0.034647	-0.02757	-0.00510	-0.00976	-0.01288
		(0.00321)	(0.00730)	(0.00428)	(0.00400)	(0.00415)
9	Public administration, defense, teaching, health and social activities	-0.12996	-0.12797	-0.12487	-0.12030	-0.11602
		(0.00364)	(0.00303)	(0.00317)	(0.00428)	(0.00251)
10	Other services	-0.07465	-0.07842	-0.07203	-0.07007	-0.07517
		(0.00243)	(0.00200)	(0.00221)	(0.00254)	(0.00206)

Notes: * Two NACE sectors are not covered by the value-added statistics: 1. Agriculture, forestry and fishing and 5.Financial services and insurances.

Each correlation is computed on the 400 datasets, the mean and standard deviation (in parenthesis) are reported.

At this disaggregation level we observe, in general, very low correlations, which are stable over time. Two exceptions are salient: the manufacturing and public sectors. The former is characterised by a relatively high and positive correlation, while the latter is especially negative. A region having an industrial structure biased (in terms of employment) towards manufacturing sector is associated with a high ratio. This municipality would therefore create more revenue than received. On the contrary, a region specialised in the public sector would be associated with a low ratio.

Making the link with the model of Borck et al. (2010), we notice that the manufacturing industry is characterised by a relatively high trade freeness and low transport costs, as evidenced by the relative importance of exports in this sector. Geographical concentration of value-added in this sector may be consequently particularly strong. In addition, the high ratio may reflect high share of capital depreciation, which enters into the measurement of the gross value-added, but is

candidate ratios per municipality. Therefore, the above described correlation is computed on each sample and results are averaged across samples.

not redistributed as income. On the small ratio side, public services are not easily transportable. Consider for instance health care. Family doctors, physiotherapists or residential care facilities are needed everywhere where households live. This sector is therefore overrepresented in the "residential" type of municipalities, in which the ratio is low.

The chosen level of industrial aggregation to calculate the correlation may appear too large. To investigate further, we perform the same exercise but at the NOGA 3 level. The results presented in figure 3.18 (Appendix) confirm the above comments.

3.6 Robustness

Capital depreciation

The wedge between income and value-added partly reflects capital depreciation. [58] We must therefore check that the identified spatial differentials of the ratio may not just reflect differences in non-distributed revenues. On the basis of the accounting results of Swiss firms (OFS, 2018a), we calculate the rate of depreciation over value-added by combining net benefits, turnover and depreciation at the NOGA 2 level. [59] We then deflate the gross value-added by the obtained rate at the pseudo-firm level. As could be expected, the corrected ratio is lower than the original one (see yearly statistics in table 3.12 in the Appendix). The spatial distribution of the ratio using net value-added is shown in figure 3.6 for 2015. The maps representing previous years are in the Appendix (figure 3.17). The patterns obtained when using net value-added are similar to the original ones reported in figure 3.2.

58. We thank Claudio Sfreddo for having suggested this robustness exercise.
59. Some NOGA 2 sectors are missing in the accounting results (i.e. 12, 14, 19, 39, 50, 51, 53, 60, 63 and 86). In these cases, we impute the average rate of depreciation in the corresponding NACE section.

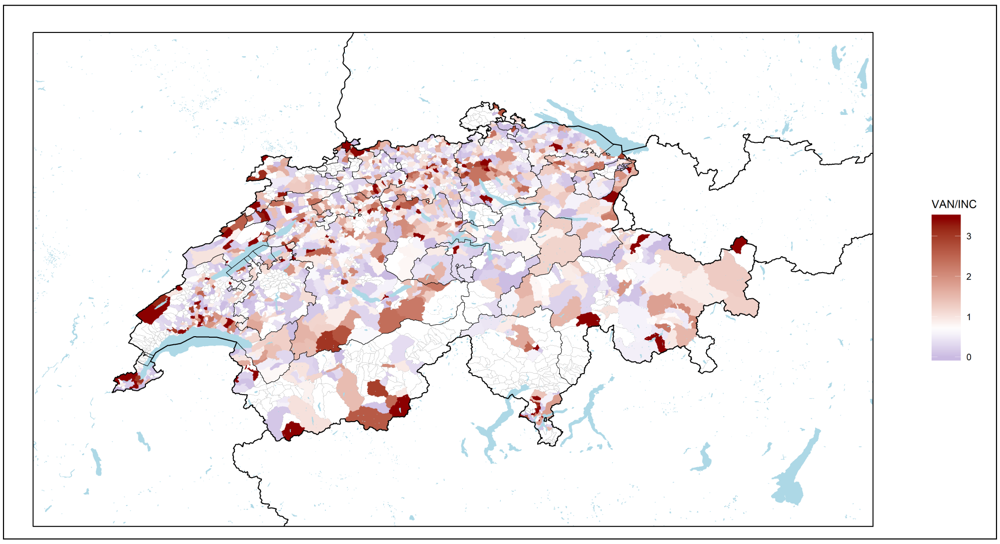


Figure 3.6: Spatial distribution of the ratio VAN/INC (2015) controlling for capital depreciation

Net benefits

The definition of municipality income as the sum of the residents' incomes can be discussed. One could argue that the net benefits from firms might also enter municipality income. To have a comprehensive view of spatial income distribution, we test this alternative measure of municipality revenue (i.e. adding firms' net benefits to households' incomes).[60] Summary statistics are provided in table 3.5.

60. The statistics are provided by the FTA, based on the Federal Direct Tax returns of legal entities. This measure overestimates total income, since dividends are counted twice.

Table 3.5: Summary statistics including net benefits

Variable	Year	Mean	Min	Median	Max	St. Dev.
Net benefit†	2011	139 075.89	0.00	4 565.65	48 354 630.30	1 601 382.13
	2012	120 923.96	0.00	4 871.95	27 935 717.20	1 164 915.41
	2013	129 812.60	0.00	5 262.50	29 845 486.80	1 263 686.38
	2014	144 235.16	0.00	5 598.15	36 717 571.90	1 398 765.85
	2015	190 525.22	0.00	5407.50	43 899 735.90	2 036 326.27
Ratio, including benefits*	2011	1.0574	0.0010	0.7384	15.1419	1.2622
		(0.0006)	(0.0000)	(0.0005)	(0.1158)	(0.0027)
	2012	0.9980	0.0005	0.7137	14.1241	1.1596
		(0.0004)	(0.0000)	(0.0005)	(0.0730)	(0.0017)
	2013	1.0026	0.0007	0.7169	14.7068	1.1844
		(0.0005)	(0.0000)	(0.0006)	(0.0651)	(0.0015)
	2014	1.0141	0.0003	0.7082	18.6790	1.2332
		(0.0005)	(0.0000)	(0.0006)	(0.2929)	(0.0047)
	2015	0.9928	0.0005	0.6905	21.9602	1.2038
		(0.0005)	(0.0000)	(0.0005)	(0.0847)	(0.0021)

Notes: * In parenthesis, the standard error of the given statistics mean (i.e. statistics are computed on each of the 400 datasets and the mean and standard deviation of the mean of each statistics are reported).
† In thousands of CHF.

Adding net firms benefit to residents income obviously reduces the ratio and its variance. Figures 3.7 and 3.8 confirm the trends already observed. Richer municipalities tend to produce more value-added, while no evident pattern emerges when getting rid of the scale effect (variables per household). This reinforces our previous findings: the most productive firms do not necessarily locate in the richest municipalities in terms of average income. [61]

Alternative weights

In our study, we rely on the set of weights ("weights II") recommended by (Tissot-Daguette and Grether, 2019) to construct a representative sample. As a final robustness exercise, we replace these weights by the alternative set ("weights I") presented by the authors (based on an ad hoc adjustment procedure). The identified trends are verified. In particular, we keep on observing a "surrounding belt" pattern around urban centres. [62]

61. The firms do not necessarily declare their benefits at the same place as they produce, either. See Figure 3.11 in Appendix. We also re-perform the Theil decompositions, with no substantial change in the results.
62. Detailed results available upon request.

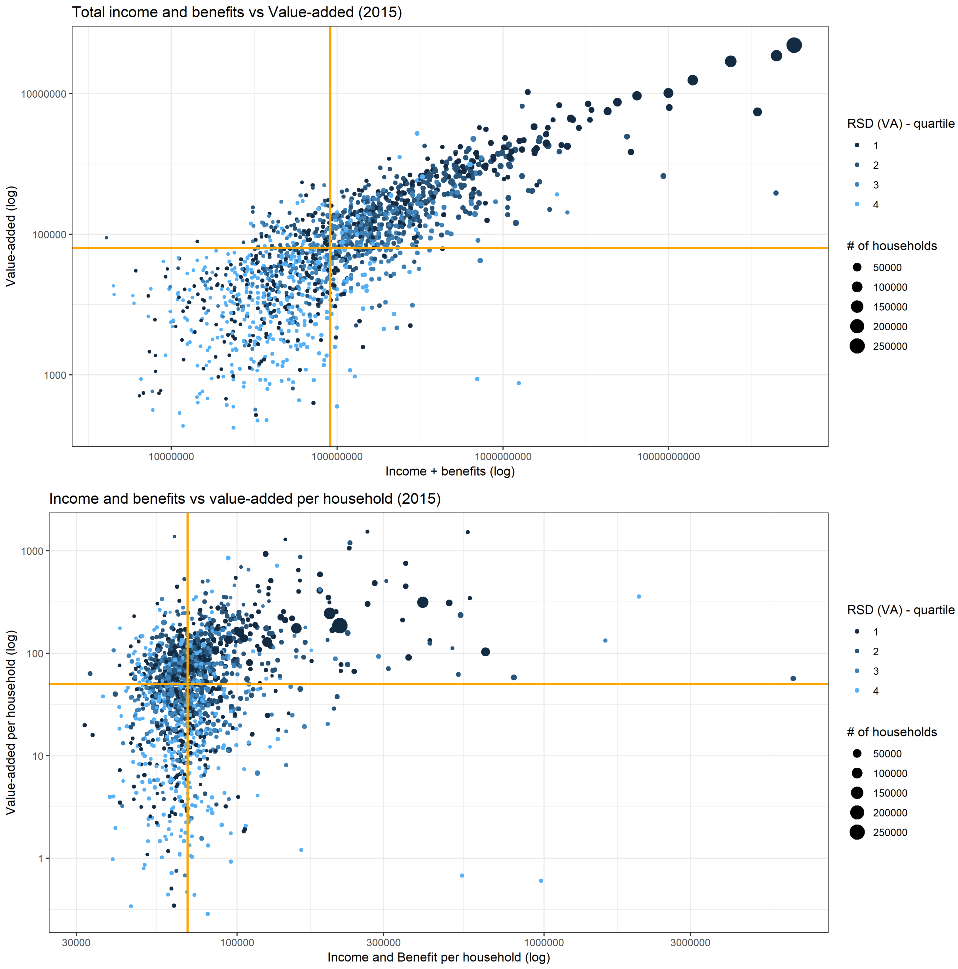


Figure 3.7: Total income (with net benefits) and value-added (total and per household) - 2015

Note: value-added (vertical axis) and total income (horizontal axis) are represented on a logarithmic scale. Orange lines are drawn at the median. The colour represents the relative standard deviation of value-added across the 400 databases. The darker the point, the lower the relative standard deviation.

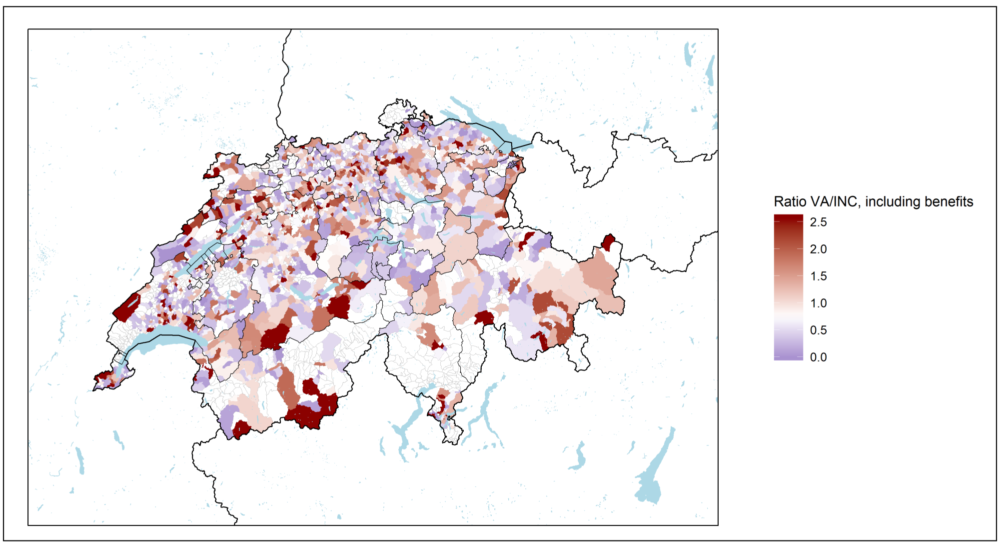


Figure 3.8: Spatial distribution of the ratio VA/INC, including benefits (2015)

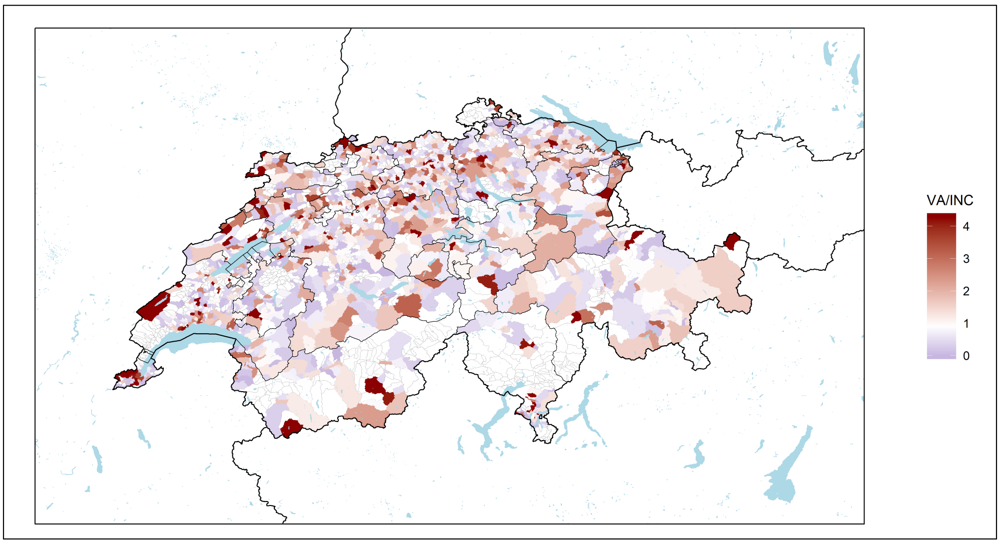


Figure 3.9: Spatial distribution of the ratio VA/INC, alternative weights definition (2015)

3.7 Conclusion

In this chapter, we explore the spatial distribution of income and production across Swiss municipalities. We find that the location process of value-added and income takes place at the municipal level, rather than at the cantonal level. Indeed, Theil decompositions show that spatial differentiation is substantial within each canton. We also note that value-added per household inequalities between municipalities are more pronounced than the spatial income disparities. The income sorting process, reinforced by the federalist structure of taxation and the adjustments on the housing market, is weaker than the concentration process of productive firms due to Marshallian externalities. Further analyses demonstrate a substantial spatial variability of the ratio between value-added and income. Municipalities attract productive firms or rich households to very different extents. The municipalities with the highest ratios are the ones in which there are agglomeration economies. On the other hand, the typical low-ratio municipality, where the share of local services is especially high, is situated near productive centres. Our results can be explained in the theoretical framework of Borck et al. (2010): when commuting costs are reasonably low, sectors with a high degree of trade freeness tend to agglomerate, and residences of skilled workers tend to spread out.

The empirical regularities uncovered by the present chapter are robust to alternative definitions of the main variables, and consistent with the interplay of forces identified by new economic geography models. As such, they constitute an adequate framework to discuss a number policy issues related to spatial economic development in the coming decades. A decrease in commuting costs (by investing in transport infrastructure for example) will probably reinforce the mismatch between productive and residential areas in the long run, increasing at the end the need for commuting and the cost of related externalities such as pollution or risks of accidents. The fiscal question comes naturally in mind as well. Tax decentralisation reinforces the income and value-added sorting processes (rich municipalities are able to set lower tax rates, which will attract rich households and/or more firms). Thus, in spite of fiscal equalisation schemes implemented between and within cantons, the fiscal system favours the spatial specialisation of municipalities, even if some fiscal equalisation schemes are put in place between and within cantons. Furthermore, our results show that a withholding tax (levied at the workplace rather than residence) would radically change the repartition of tax revenues among Swiss municipalities. Although such a system would not be necessarily desirable nor fairer (public spending depends on the number of residents, probably more than on the number of jobs), it illustrates how the location of production matters and should be carefully included into the negotiations about financial equalisation. The efforts to compensate the excess burden on urban centres go into that direction (see Ecoplan (2017)). Finally, as economic specialisation occurs mainly within cantons, a substantial part of spatial economic disparities could be addressed by public action at the cantonal level.

To conclude, it should be noted that the present analysis is only descriptive and therefore insufficient to address comprehensively the above-mentioned issues. More research is required to connect more precisely structural change and economic policies to spatial inequality measures. The latter will remain a fundamental assessment tool and deserve to be regularly monitored in the future.

3.8 Appendix

Empirical framework

The between-within Theil decomposition

Notation: there are M cantons and N_j municipalities i in each canton j. The total number of municipalities is therefore given by $N = \sum_{j=1}^{M} N_j$. The share of canton j in the national total with respect to variable a is given by $s_j = \frac{N_j \cdot \bar{a}_j}{N \cdot \bar{a}}$. Where $\bar{a}$ is the national mean and $\bar{a}_j$ the mean in canton j. We have therefore:

$$
\begin{aligned}
T_a &= \frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}\cdot\ln\left(\frac{a_i}{\bar{a}}\right)\\
&= \frac{1}{N}\sum_{j=1}^{M}N_j\cdot\frac{1}{N_j}\sum_{i=1}^{N_j}\frac{a_i\cdot\bar{a}_j}{\bar{a}_j\cdot\bar{a}}\left[\ln\left(\frac{a_i}{\bar{a}_j}\right)+\ln\left(\frac{\bar{a}_j}{\bar{a}}\right)\right]\\
&= \sum_{j=1}^{M}\underbrace{\frac{N_j\bar{a}_j}{N\bar{a}}}_{s_j}\cdot\underbrace{\frac{1}{N_j}\sum_{i=1}^{N_j}\frac{a_i}{\bar{a}_j}\cdot\ln\left(\frac{a_i}{\bar{a}_j}\right)}_{T_{a_j}}+\sum_{j=1}^{M}\underbrace{\frac{N_j\bar{a}_j}{N\bar{a}}}_{s_j}\cdot\frac{1}{N_j}\cdot\ln\left(\frac{\bar{a}_j}{\bar{a}}\right)\underbrace{\sum_{i=1}^{N_j}\frac{a_i}{\bar{a}_j}}_{N_j}\\
&= \underbrace{\sum_{j=1}^{M}s_j\cdot T_{a_j}}_{within}+\underbrace{\sum_{j=1}^{M}s_j\cdot\ln\left(\frac{\bar{a}_j}{\bar{a}}\right)}_{between}
\end{aligned}
$$

The product/ratio Theil decomposition

Assume the variable a is the product of b and c, i.e. $a_i = b_i \cdot c_i$. The Theil index for variable a is given by:

$$
\begin{aligned}
T_a &= \frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}\ln\left(\frac{a_i}{\bar{a}}\right)\quad\text{where}\quad\frac{a_i}{\bar{a}}=\frac{b_i}{\bar{b}}\cdot\frac{c_i}{\bar{c}}\cdot\frac{\bar{b}\bar{c}}{\bar{a}}=\frac{b_i}{\bar{b}}\cdot\frac{c_i}{\bar{c}}\cdot\gamma\text{ ,with }\gamma=1-\frac{cov(b,c)}{\bar{a}}\\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}\ln\left(\frac{b_i}{\bar{b}}\cdot\frac{c_i}{\bar{c}}\cdot\gamma\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}\ln\left(\frac{b_i}{\bar{b}}\right)+\frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}\ln\left(\frac{c_i}{\bar{c}}\right)+\ln\gamma\cdot\underbrace{\frac{1}{N}\sum_{i=1}^{N}\frac{a_i}{\bar{a}}}_{1}\\
&= \gamma\left[\frac{1}{N}\sum_{i=1}^{N}\frac{c_i}{\bar{c}}\cdot\frac{b_i}{\bar{b}}\ln\left(\frac{b_i}{\bar{b}}\right)+\frac{1}{N}\sum_{i=1}^{N}\frac{b_i}{\bar{b}}\cdot\frac{c_i}{\bar{c}}\ln\left(\frac{c_i}{\bar{c}}\right)\right]+\ln\gamma
\end{aligned}
$$

$$= \gamma\Bigg[\underbrace{\frac{1}{N}\sum_{i=1}^{N}\frac{b_i}{\bar{b}}\ln(\frac{b_i}{\bar{b}})}_{T_b}\underbrace{\frac{1}{N}\sum_{i=1}^{N}\frac{c_i}{\bar{c}}}_{1} + cov(\frac{b_i}{\bar{b}}\ln(\frac{b_i}{\bar{b}}), \frac{c_i}{\bar{c}}) +$$

$$\underbrace{\frac{1}{N}\sum_{i=1}^{N}\frac{c_i}{\bar{c}}\ln(\frac{c_i}{\bar{c}})}_{T_c}\underbrace{\frac{1}{N}\sum_{i=1}^{N}\frac{b_i}{\bar{b}}}_{1} + cov(\frac{c_i}{\bar{c}}\ln(\frac{c_i}{\bar{c}}), \frac{b_i}{\bar{b}})\Bigg] + \ln\gamma$$

$$= \gamma\Big[T_b + T_c\Big] + \underbrace{\Bigg[\ln\gamma + \gamma\Bigg(cov(\frac{b_i}{\bar{b}}\ln(\frac{b_i}{\bar{b}}), \frac{c_i}{\bar{c}}) + cov(\frac{c_i}{\bar{c}}\ln(\frac{c_i}{\bar{c}}), \frac{b_i}{\bar{b}})\Bigg)\Bigg]}_{\Omega}$$

$$= \gamma\Big[T_b + T_c\Big] + \Omega$$

Combining between-within and product/ratio Theil decompositions

We assume the variable a pertains to M groups, while being the product of b and c. In our application a is the ratio of interest (VA/INC), while b is value-added per households and c is the inverse of the income per households. In particular, we want to decompose the Theil index for variable a into within and between effects, but also determining the specific contributions of b, c and their interaction (named hereafter Cov). As a picture is worth thousands words, we aim to algebraically fill the matrix below. By using the properties of between-within and product/ratio decompositions, we can already define the borders of the matrix. The column "Total" is found using the between-within decomposition of the a Theil index, while the row "Total" is found using the product/ratio decomposition. We still need to derive algebraically the between and within effects specific to a, b and Cov.

Table 3.6: Two-way Matrix decomposition (before completion)

	b	c	$Cov.$	Total
Within	$(\text{within})_b$	$(\text{within})_c$	$(\text{within})_{Cov}$	$\sum_{j=1}^{M} s_j T_{a_j}$
Between	$(\text{between})_b$	$(\text{between})_c$	$(\text{between})_{Cov}$	$\sum_{j=1}^{M} s_j \ln(\frac{\bar{a}_j}{\bar{a}})$
Total	γT_b	γT_c	Ω	T_a

To fill table 3.6, we first need to notice that:

$$\bar{a}_j = \sum_{i=1}^{N_j} \frac{a_i}{N_j} = \frac{\bar{b}_j \cdot \bar{c}_j}{\gamma_j} \qquad \text{with} \quad \gamma_j = 1 - \frac{cov(b_j, c_j)}{\bar{a}_j}$$

$$\bar{a} = \sum_{i=1}^{N} \frac{a_i}{N} = \frac{\bar{b} \cdot \bar{c}}{\gamma} \qquad \text{with} \quad \gamma = 1 - \frac{cov(b, c)}{\bar{a}}$$

we have therefore that:

$$\frac{a_i}{a_j} = \frac{b_i c_i}{\bar{b}_j \bar{c}_j}\gamma_j \qquad \text{and} \qquad \frac{\bar{a}_j}{\bar{a}} = \frac{\bar{b}_j \bar{c}_j}{\bar{b}\ \bar{c}}\frac{\gamma}{\gamma_j}$$

Using these and the between-within decomposition shown above, we can write the Theil index of the ratio (T_a) as:

$$\begin{aligned}
T_a &= \sum_{j=1}^{M} s_j \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{b_i c_i}{\bar{b}_j \bar{c}_j}\gamma_j \Big[\ln\big(\frac{b_i c_i}{\bar{b}_j \bar{c}_j}\gamma_j\big) \Big] + \sum_{j=1}^{M} s_j \Big[\ln\big(\frac{b_j c_j}{\bar{b}\ \bar{c}}\frac{\gamma}{\gamma_j}\big) \Big] \\
&= \sum_{j=1}^{M} s_j \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{b_i c_i}{\bar{b}_j \bar{c}_j}\gamma_j \Big[\ln\big(\frac{b_i}{\bar{b}_j}\big) + \ln\big(\frac{c_i}{\bar{c}_j}\big) + \ln(\gamma_j) \Big] + \sum_{j=1}^{M} s_j \Big[\ln\big(\frac{b_j}{\bar{b}}\big) + \ln\big(\frac{c_j}{\bar{c}}\big) + \ln\big(\frac{\gamma}{\gamma_j}\big) \Big] \\
&= \sum_{j=1}^{M} s_j \gamma_j \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{c_i}{\bar{c}_j} \cdot \frac{b_i}{\bar{b}_j} \ln\big(\frac{b_i}{\bar{b}_j}\big) + \sum_{j=1}^{M} s_j \gamma_j \frac{1}{N_j} \sum_{i=1}^{N_j} \frac{b_i}{\bar{b}_j} \cdot \frac{c_i}{\bar{c}_j} \ln\big(\frac{c_i}{\bar{c}_j}\big) + \sum_{j=1}^{M} s_j \ln(\gamma_j) \underbrace{\frac{1}{N_j} \sum_{i=1}^{N_j} \frac{a_i}{\bar{a}_j}}_{1} + \\
&\quad \sum_{j=1}^{M} s_j \Big[\ln\big(\frac{b_j}{\bar{b}_j}\big) + \ln\big(\frac{c_j}{\bar{c}_j}\big) + \ln\big(\frac{\gamma}{\gamma_j}\big) \Big] \\
&= \sum_{j=1}^{M} s_j \gamma_j \Big[T_{b_j} + cov\big(\frac{b_i}{\bar{b}} \ln\big(\frac{b_i}{\bar{b}}\big), \frac{c_i}{\bar{c}}\big) \Big] + \sum_{j=1}^{M} s_j \gamma_j \Big[T_{c_j} + cov\big(\frac{c_i}{\bar{c}} \ln\big(\frac{c_i}{\bar{c}}\big), \frac{b_i}{\bar{b}}\big) \Big] + \sum_{j=1}^{M} s_j \ln(\gamma_j) + \\
&\quad \sum_{j=1}^{M} s_j \Big[\ln\big(\frac{b_j}{\bar{b}_j}\big) + \ln\big(\frac{c_j}{\bar{c}_j}\big) + \ln\big(\frac{\gamma}{\gamma_j}\big) \Big] \\
&= \sum_{j=1}^{M} s_j \Big\{ \gamma_j \Big[T_{b_j} + T_{c_j} \Big] + \Big[\ln(\gamma_j) + cov\big(\frac{b_i}{\bar{b}} \ln\big(\frac{b_i}{\bar{b}}\big), \frac{c_i}{\bar{c}}\big) + cov\big(\frac{c_i}{\bar{c}} \ln\big(\frac{c_i}{\bar{c}}\big), \frac{b_i}{\bar{b}}\big) \Big] \Big\} + \\
&\quad \sum_{j=1}^{M} s_j \Big\{ \ln\big(\frac{b_j}{\bar{b}_j}\big) + \ln\big(\frac{c_j}{\bar{c}_j}\big) + \ln\big(\frac{\gamma}{\gamma_j}\big) \Big\}
\end{aligned}$$

We want to decompose the respective b and c Theil indices also into within-between effects. To do so, we first need to remember, taking as example variable b, that:

$$T_b = \underbrace{\sum_{j=1}^{M} \frac{N_j \cdot \bar{b}_j}{N \cdot \bar{b}} \cdot T_{b_j}}_{within} + \underbrace{\sum_{j=1}^{M} \frac{N_j \cdot \bar{b}_j}{N \cdot \bar{b}} \cdot \ln\left(\frac{\bar{b}_j}{\bar{b}}\right)}_{between}$$

We notice that:

$$\frac{N_j \cdot \bar{b}_j}{N \cdot \bar{b}} = s_j \frac{\bar{a}}{\bar{a}_j} \frac{\bar{b}_j}{\bar{b}} = s_j \frac{\bar{b}\,\bar{c}}{\bar{b}_j \bar{c}_j} \frac{\bar{b}_j}{\bar{b}} \frac{\gamma_j}{\gamma} = s_j \frac{\bar{c}}{\bar{c}_j} \frac{\gamma_j}{\gamma}$$

Therefore:

$$T_b = \sum_{j=1}^{M} s_j \frac{\bar{c}}{\bar{c}_j} \frac{\gamma_j}{\gamma} \cdot T_{b_j} + \sum_{j=1}^{M} s_j \frac{\bar{c}}{\bar{c}_j} \frac{\gamma_j}{\gamma} \cdot \ln\left(\frac{\bar{b}_j}{\bar{b}}\right)$$

This implies that the within and between components of T_b (T_c) should be adjusted by a factor $\frac{\bar{c}}{\bar{c}_j}\frac{\gamma_j}{\gamma}$ (respectively $\frac{\bar{b}}{\bar{b}_j}\frac{\gamma_j}{\gamma}$). This also intervenes in the definition of the overall interaction term, Ω. Therefore, the decomposition rewrites:

$$\begin{aligned} T_a = &\sum_{j=1}^{M} s_j \left\{ \frac{\gamma_j}{\gamma} \left[\frac{\bar{c}}{\bar{c}_j} T_{b_j} + \frac{\bar{b}}{\bar{b}_j} T_{c_j} \right] + \left[\ln(\gamma_j) + cov\left(\frac{b_i}{\bar{b}} \ln\left(\frac{b_i}{\bar{b}}\right), \frac{c_i}{\bar{c}}\right) + cov\left(\frac{c_i}{\bar{c}} \ln\left(\frac{c_i}{\bar{c}}\right), \frac{b_i}{\bar{b}}\right) \right] \right\} + \\ &\sum_{j=1}^{M} s_j \left\{ \frac{\gamma_j}{\gamma} \frac{\bar{c}}{\bar{c}_j} \ln\left(\frac{b_j}{\bar{b}_j}\right) + \frac{\gamma_j}{\gamma} \frac{\bar{b}}{\bar{b}_j} \ln\left(\frac{c_j}{\bar{c}_j}\right) + \ln\left(\frac{\gamma}{\gamma_j}\right) \right\} + \\ &\sum_{j=1}^{M} s_j \left\{ \left[1 - \frac{\bar{c}}{\bar{c}_j}\right] \frac{\gamma_j}{\gamma} T_{b_j} + \left[1 - \frac{\bar{b}}{\bar{b}_j}\right] \frac{\gamma_j}{\gamma} T_{c_j} + \left[1 - \frac{\bar{c}}{\bar{c}_j}\right] \frac{\gamma_j}{\gamma} \ln\left(\frac{b_j}{\bar{b}_j}\right) + \left[1 - \frac{\bar{b}}{\bar{b}_j}\right] \frac{\gamma_j}{\gamma} \ln\left(\frac{c_j}{\bar{c}_j}\right) \right\} \end{aligned}$$

Hence, we can fill in table 3.7.

Table 3.7: Two-way Matrix decomposition (completed)

	b	c	$Cov.$	Total
Within	$\sum_{j=1}^{M} s_j \gamma_j \frac{\bar{c}}{\bar{c}_j} T_{b_j}$	$\sum_{j=1}^{M} s_j \gamma_j \frac{\bar{b}}{\bar{b}_j} T_{c_j}$	$\sum_{j=1}^{M} s_j \Omega_j$	$\sum_{j=1}^{M} s_j T_{a_j}$
Between	$\sum_{j=1}^{M} s_j \frac{\gamma_j}{\gamma} \frac{\bar{c}}{\bar{c}_j} \ln\left(\frac{b_j}{\bar{b}_j}\right)$	$\sum_{j=1}^{M} s_j \frac{\gamma_j}{\gamma} \frac{\bar{b}}{\bar{b}_j} \ln\left(\frac{c_j}{\bar{c}_j}\right)$	$\sum_{j=1}^{M} s_j \Theta_j$	$\sum_{j=1}^{M} s_j \ln\left(\frac{\bar{a}_j}{\bar{a}}\right)$
Total	γT_b	γT_c	Ω	T_a

Where the within and between interaction terms, Ω_j and Θ_j, are given by,

$$\Omega_j = \ln(\gamma_j) + cov\left(\frac{b_i}{\bar{b}} \ln\left(\frac{b_i}{\bar{b}}\right), \frac{c_i}{\bar{c}}\right) + cov\left(\frac{c_i}{\bar{c}} \ln\left(\frac{c_i}{\bar{c}}\right), \frac{b_i}{\bar{b}}\right) + \left[1 - \frac{\bar{c}}{\bar{c}_j}\right] \frac{\gamma_j}{\gamma} T_{b_j} + \left[1 - \frac{\bar{b}}{\bar{b}_j}\right] \frac{\gamma_j}{\gamma} T_{c_j}$$

$$\Theta_j = \ln\left(\frac{\gamma}{\gamma_j}\right) + \left[1 - \frac{\bar{c}}{\bar{c}_j}\right]\frac{\gamma_j}{\gamma}\ln\left(\frac{b_j}{\bar{b}_j}\right) + \left[1 - \frac{\bar{b}}{\bar{b}_j}\right]\frac{\gamma_j}{\gamma}\ln\left(\frac{c_j}{\bar{c}_j}\right)$$

Results

Plots

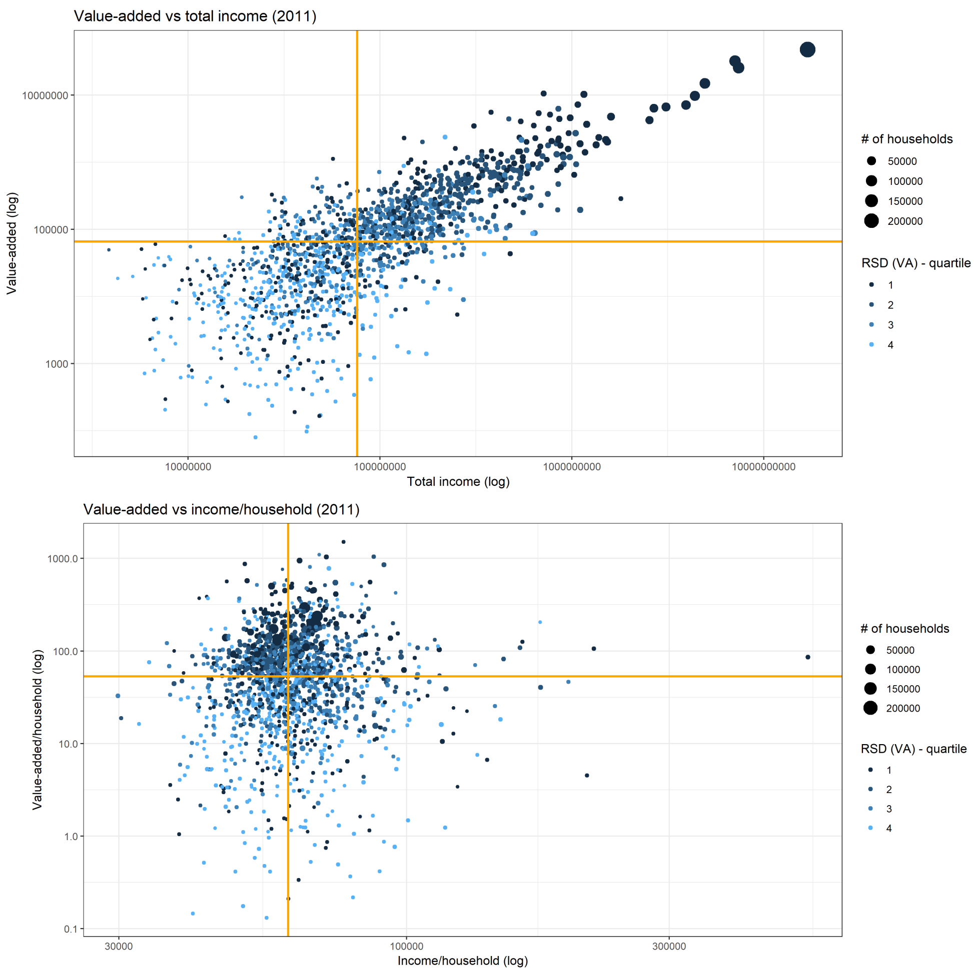


Figure 3.10: Income and value-added (total and per worker) - 2011

Note: value-added (vertical axis) and income (horizontal axis) are represented on a logarithmic scale. Orange lines are drawn at the median. The colour represents the relative standard deviation of value-added across the 400 databases. The darker the point, the lower the relative standard deviation.

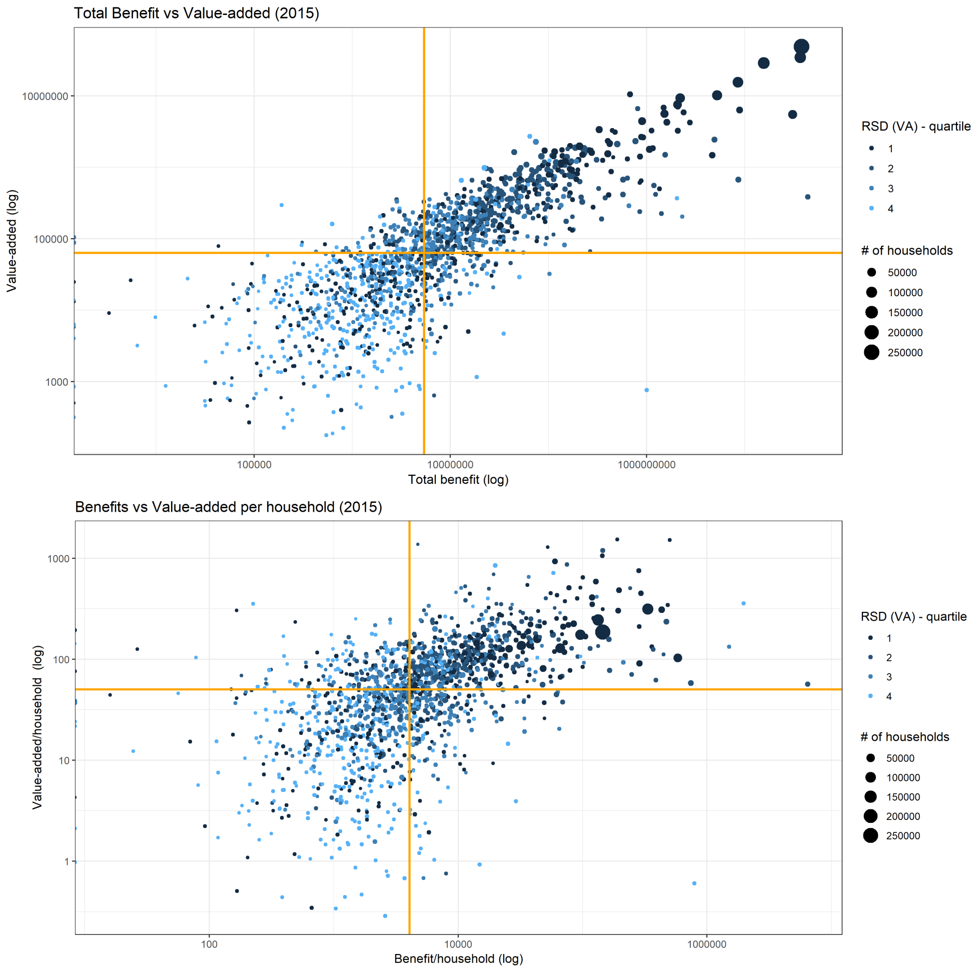


Figure 3.11: Total benefit and value-added (total and per worker) - 2011

Note: value-added (vertical axis) and benefit (horizontal axis) are represented on a logarithmic scale. Orange lines are drawn at the median. The colour represents the relative standard deviation of value-added across the 400 databases. The darker the point, the lower the relative standard deviation.

Ratio Density

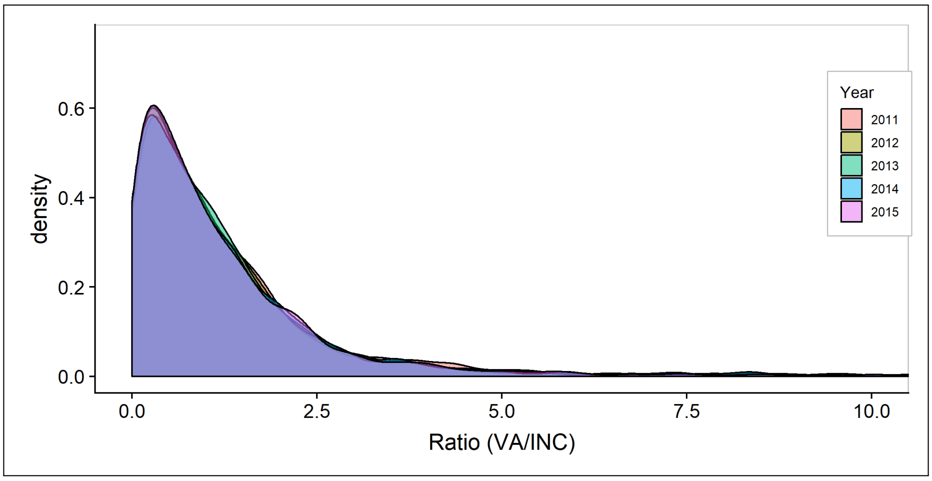


Figure 3.12: Distribution of the ratio of interest, value-added over income 2011-2015

Theil decompositions

Table 3.8: Two-way Matrix decomposition for 2011

	VPH*	IPH*	Cov.	Total
Within	0.56169	0.01285	-0.02845	0.54609
	(0.01111)	(0.00002)	(0.00285)	(0.01025)
Between	0.04269	0.00860	-0.02146	0.02983
	(0.00306)	(0.00001)	(0.00115)	(0.00272)
Total	0.60438	0.02145	-0.04991	0.57592
	(0.01230)	(0.00003)	(0.00345)	(0.01136)

Notes: * VPH is value-added per household and IPH inverse of income per household.
See table 3.1 for the analytical expressions of the decomposition.

Table 3.9: Two-way Matrix decomposition for 2012

	VPH*	IPH*	Cov.	Total
Within	0.55482	0.01249	-0.02801	0.53929
	(0.01090)	(0.00002)	(0.00297)	(0.00901)
Between	0.04783	0.00831	-0.02302	0.03313
	(0.00254)	(0.00001	(0.00124)	(0.00180)
Total	0.60265	0.02080	-0.05103	0.57242
	(0.01145)	(0.00003)	(0.00363)	(0.00922)

Notes: * VPH is value-added per household and IPH inverse of income per household.
See table 3.1 for the analytical expressions of the decomposition.

Table 3.10: Two-way Matrix decomposition for 2013

	VPH*	IPH*	Cov.	Total
Within	0.57907	0.01274	-0.03383	0.55797
	(0.01052)	(0.00002)	(0.00322)	(0.00873)
Between	0.04223	0.00812	-0.01904	0.03131
	(0.00287)	(0.00001)	(0.00143)	(0.00217)
Total	0.62129	0.02086	-0.05287	0.58928
	(0.01161)	(0.00003)	(0.00405)	(0.00929)

Notes: * VPH is value-added per household and IPH inverse of income per household.
See table 3.1 for the analytical expressions of the decomposition.

Table 3.11: Two-way Matrix decomposition for 2014

	VPH*	IPH*	Cov.	Total
Within	0.60502	0.01239	-0.04264	0.57476
	(0.01159)	(0.00002)	(0.00308)	(0.01189)
Between	0.04557	0.00821	-0.02153	0.03225
	(0.00278)	(0.00001)	(0.00123)	(0.00206)
Total	0.65059	0.02060	-0.06417	0.60702
	(0.01257)	(0.00003)	(0.00384)	(0.01276)

Notes: * VPH is value-added per household and IPH inverse of income per household.
See table 3.1 for the analytical expressions of the decomposition.

Spatial analysis of the Production/Income ratio

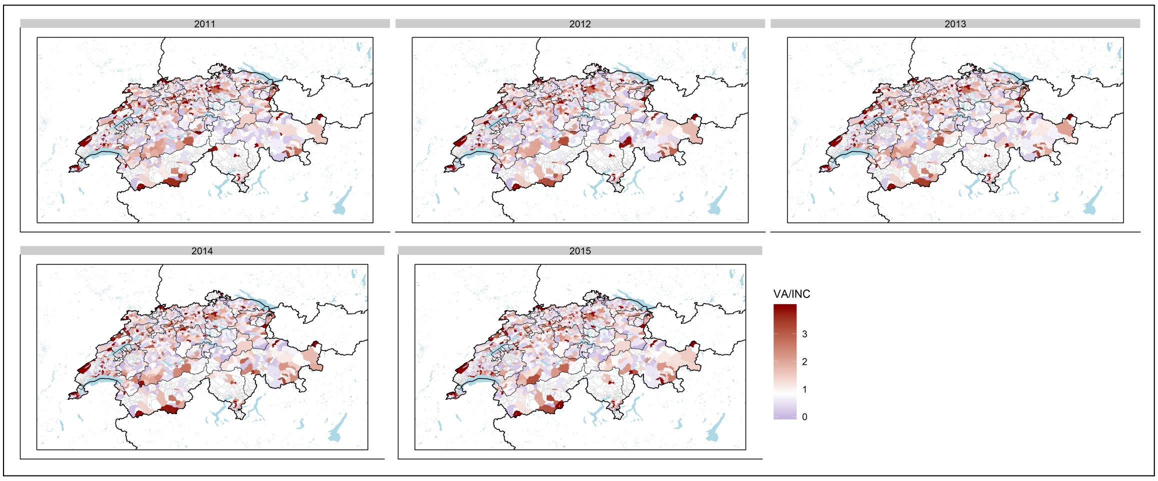


Figure 3.13: Spatial distribution of the ratio VA/INC (value-added over income)

The high correlation in the manufacturing sector is driven mainly by four sectors: the tobacco-based production (120), the pharmaceutical production (212), the manufacturing of

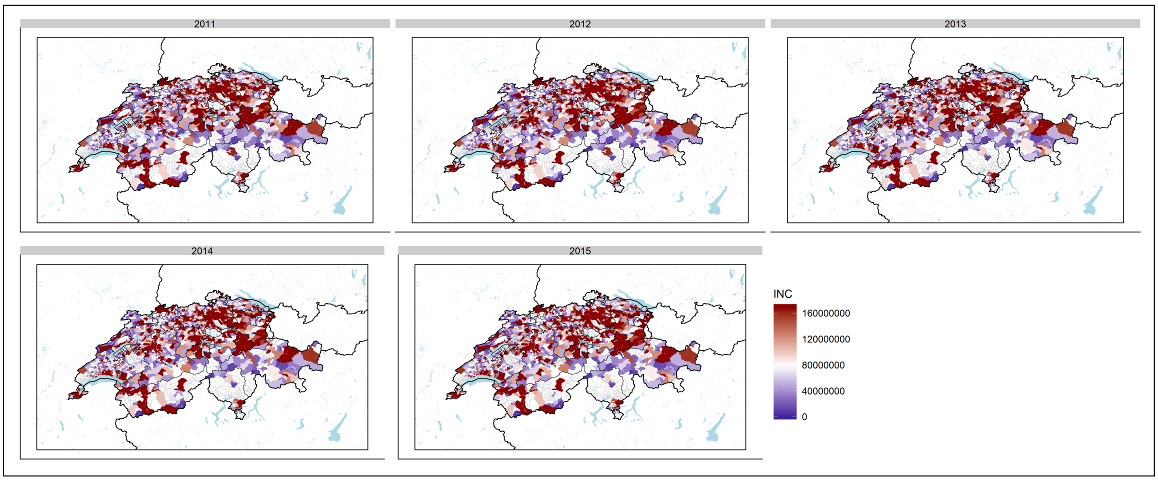


Figure 3.14: Spatial distribution of income (INC)

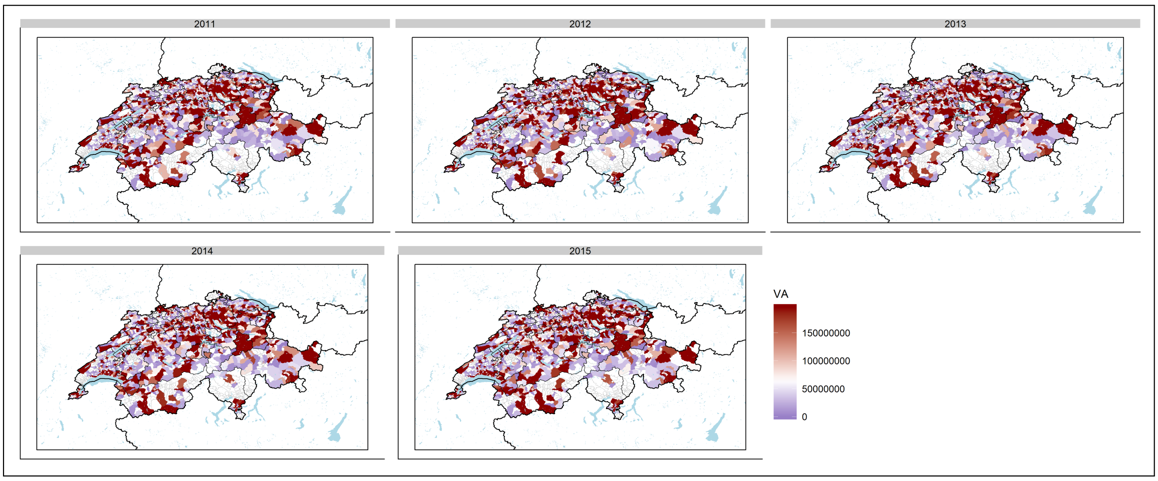


Figure 3.15: Spatial distribution of value-added (VA)

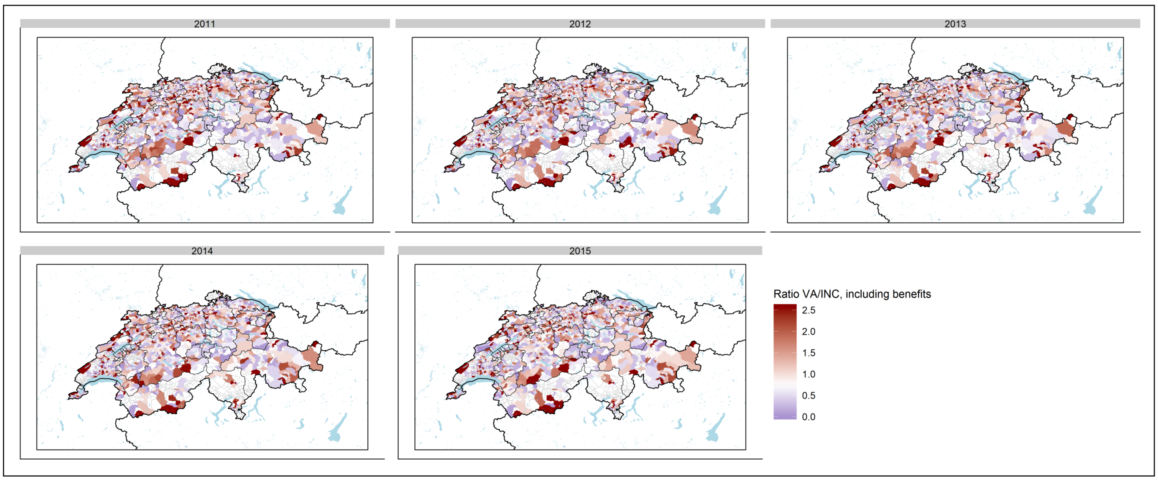


Figure 3.16: Spatial distribution of the ratio VA/INC, including benefits

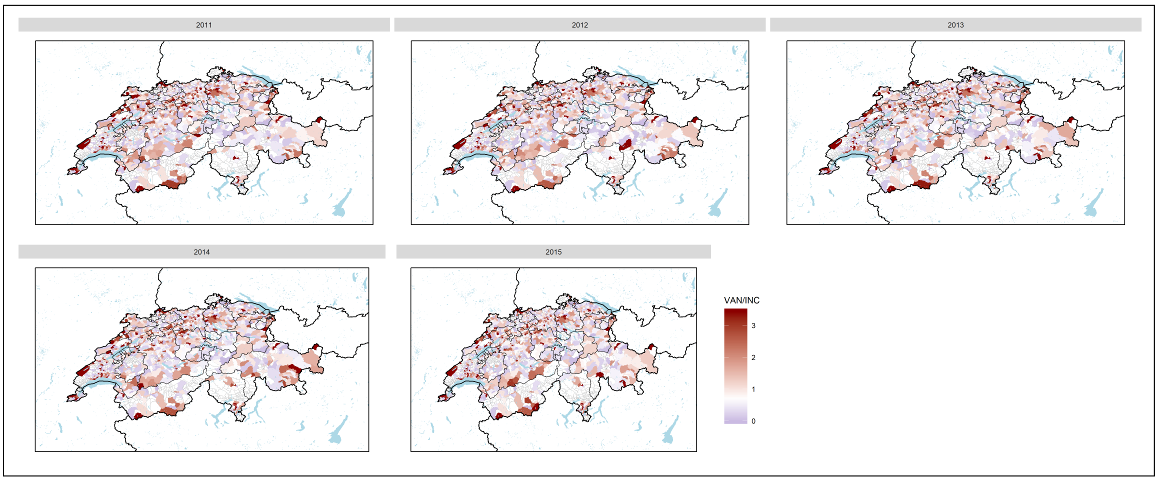


Figure 3.17: Spatial distribution of the ratio VAN/INC

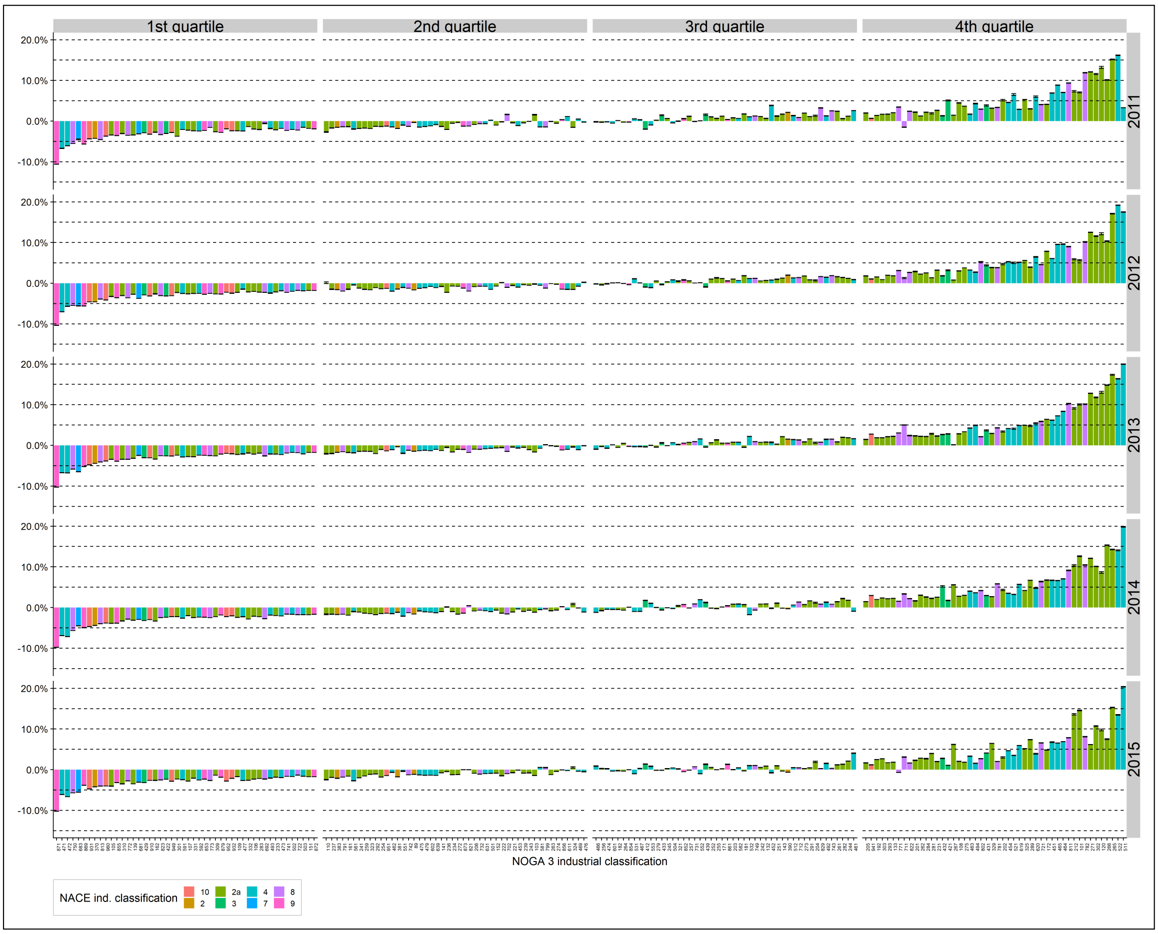


Figure 3.18: Correlation between the share of employment and the VA/INC ratio, by NOGA 3 sectors

Note: correlations are computed on each of the 400 datasets, averages are presented along with the means 95% confidence intervals. The correlations are ordered based on their yearly average value. Quartiles are computed on yearly averages as well.

medical irradiation and electromagnetic devices (266) and production of watches and precision measurement instruments (265). This explains the high ratios observed in la Vallée de Joux and le Locle, whose economy is based on the watch industry, in Neuchâtel, whose economy is fuelled by the watch and tobacco-based production and in Basel, turned towards the pharmaceutical industry. When looking at the other extreme, the most negative correlation was found in the health sector (NACE 9). This correlation is driven by the care-home residence sector (871). Such services should localise on the entire territory and not only in dense urban area or near related-industries. Therefore it does not benefit neither from urbanisation nor localisation economies. In the other industrial sections, we note other several interesting cases at a more disaggregated level. The wholesale sector (463 to 469) exhibits high correlation with the ratio, all being in the last distribution quartile, especially wholesale of domestic goods (464) (pharmaceutical products included). As shown by Tissot-Daguette (2019), this type of activities benefit mainly from urbanisation economies. Industries cluster together in dense urban areas which therefore have high value-added. Another important positive correlations can be found in the sector of aerial transportation (511) (and related services (522)), in placement agencies (781, 782) or house-cleaning service (811) for a similar reason. In other words, these industries locate in cities and, therefore, correlate with the large wealth created in urban center.

We also identify industries that are negatively correlated with the ratio of interest. The cases of retail business (471, 472) and restaurants (561) are similar to the one of home-care residence. They have a negative correlation because their consumption is locally determined and therefore spread over all the territory.

Figures 3.19 - 3.22 show the value-added per full-time equivalent specific to eight NACE sectors. The value for the municipality itself is represented on the horizontal axis, while the vertical dimension represents the average among its contiguous municipalities (that are available in the dataset, for this sector).

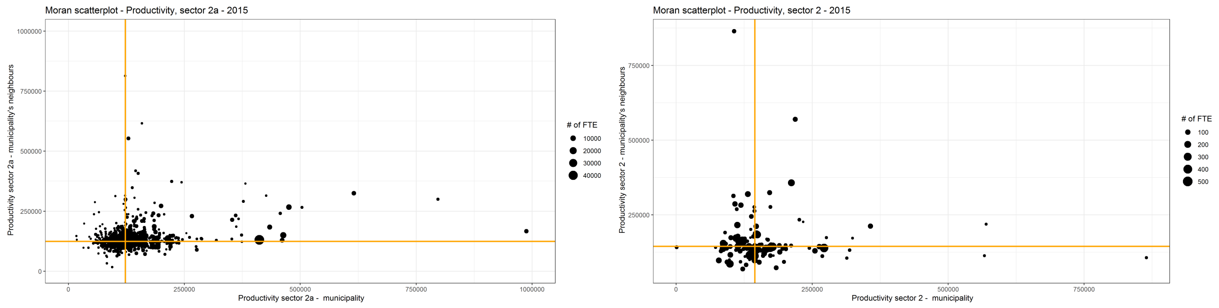


Figure 3.19: Moran scatterplot sectors 2a (manufacturing industries) and 2 (extraction and other industries) - 2015

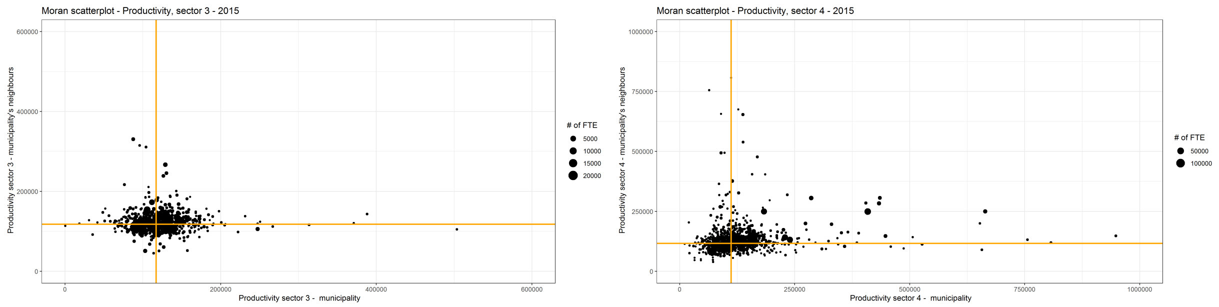


Figure 3.20: Moran scatterplots sectors 3 (building) and 4 (whole and retail sales, transports, hotels and restaurant, TIC) - 2015

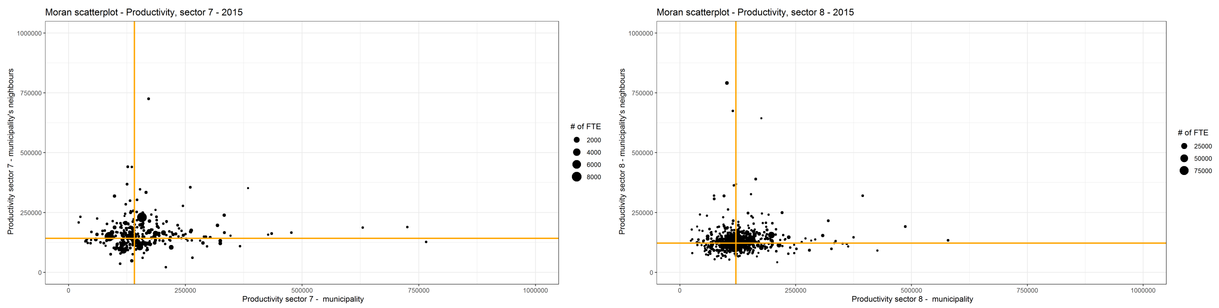


Figure 3.21: Moran scatterplot sectors 7 (real estate) and 8 (scientific and technical activities, administrative services) - 2015

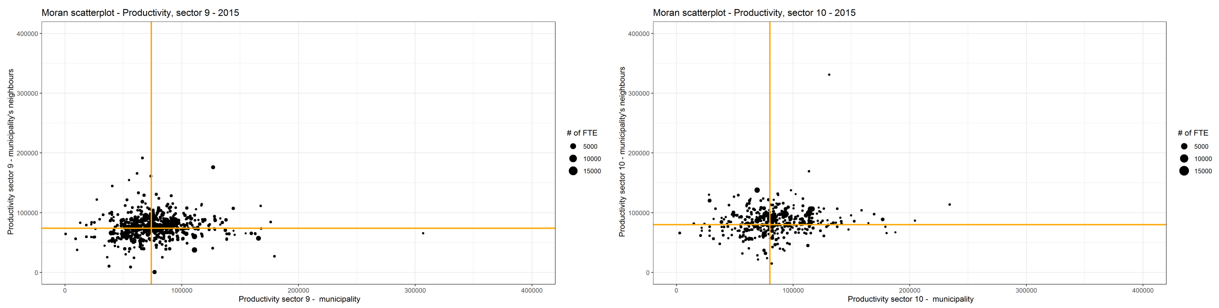


Figure 3.22: Moran scatterplot sectors 9 (public administration, defense, teaching, health and social activities) and 10 (other services) - 2015

Robustness

Table 3.12: Ratio of net value-added over income

Variable	Year	Mean	Min	Median	Max	St. Dev.
Ratio* †	2011	1.1577	0.0009	0.7271	17.4208	1.5763
		(0.0006)	(0.0000)	(0.0005)	(0.1155)	(0.0025)
	2012	1.1038	0.0004	0.6987	16.5976	1.5019
		(0.0004)	(0.0000)	(0.0004)	(0.0990)	(0.0018)
	2013	1.1241	0.0007	0.70570	17.7388	1.5765
		(0.0005)	(0.0000)	(0.0007)	(0.1013)	(0.0020)
	2014	1.1494	0.0003	0.6944	20.9693	1.6645
		(0.0005)	(0.0000)	(0.0005)	(0.2931)	(0.0041)
	2015	1.1152	0.0010	0.6836	20.5372	1.5925
		(0.0005)	(0.0000)	(0.0006)	(0.0819)	(0.0018)

Notes: * In parenthesis, the standard error of the given statistics mean (i.e. statistics are computed on each of the 400 datasets and the mean and standard deviation of the mean of each statistics are reported).

† Value-added, net of capital depreciation.

General Conclusion

The analysis of the spatial distribution of income can rely on diverse strands of the literature, among which income sorting models, convergence theories and New Economic Geography. The central question is however always the same: where is wealth located and why? The present thesis proposes answers along three dimensions.

The first dimension of the answer is presented in chapter 1: rich households live where natural amenities attract them - or at least do not repel them. Moreover, we find that within income inequalities are higher in municipalities located at the water's edge, in the mountains and/or far away from agglomerations. According to the general spatial equilibrium model of Fretz et al. (2017), this comes from trade and commuting costs. If this is true, then any decrease in distance-related costs should increase income sorting, i.e. decrease within- and increase between-municipalities income inequality.

The second dimension of the answer is the focus of chapter 2: income essentially stays at the same place over time. The spatial distribution of income crucially depends on its own past. Disparities are persistent because each municipality follows its own growth path. This is all the more true since the beginning of the 21st century, where the strength of spatial spillovers has decreased and initial conditions have gained in importance. According to our simulations, there is little hope for convergence among municipalities by 2040/41.

The last element of the response is provided by chapter 3: income does not originates where it is perceived. When we look at the spatial distribution of income and value-added together, we observe that they are naturally linked, but not completely tied. Some areas specialise into attracting residents whereas some other offer opportunities of agglomeration economies to productive firms. This specialisation takes place at the municipal level and gives rise to a "residential surrounding belt" pattern, in which rich households tend to settle in nice suburban environment around productive centres.

We innovate on several counts. First our set of data was never used. In particular, we combine several different sources (Federal Tax Administration, Federal Statistical Office, Federal Office for Spatial Development, Swisstopo) to conduct our analyses in the first chapter. We benefit from a new matching of the value-added statistics (WS - Wertschöpfungsstatistik) and the structural business statistics (STATENT - Statistique structurelle des entreprises) in the third. We therefore work at a highly disaggregated level and apply spatial econometrics tools to small geographical entities in a novel way. Finally, the issue of spatial distribution itself has been seldom explored in Switzerland. Many studies treat the role of taxation, but the outcome in terms of income sorting has not been the main concern.

In a nutshell, we find that the income distribution is tightly related to the natural environment, that disparities between Swiss municipalities are persistent over time and that the location of value-added and of income are driven by different forces, which lead to a spatial mismatch. These results have important implications in the public debate. In terms of redistribution policies, they mean that financial equalisation schemes will remain needed, all the more if tax differentials increase. In terms of regional development, we can expect growing concentration of economic activity in productive centres, which calls for urban densification policies. The persistence of the spatial mismatch between production and income also implies commuting. To reduce negative externalities related to mobility, the development of public transports appears to be a necessity.

In spite of these contributions, our thesis leaves several aspects of the spatial income distribution unexplored. It would be informative to work at an even more disaggregated geographical level, such as neighbourhoods. Intra-city income sorting is indeed an important phenomenon that deserves further attention. But given confidentiality issues, these analyses are still difficult to run. In addition, other variables would have been instructive: disposable or post-tax income, Theil index, share of income hold by the top rich, etc. They do not exist for now at the municipal level. This is a shame, insofar as the recent increase of income inequality in developed countries appears to come from the top of the distribution (see Dell et al. (2005) and Foellmi and Martínez (2017)). A study focusing on rich households at the municipal level remains to be done. This also raises the issue of tax evasion, which increases income inequality (see Alstadsæter et al. (2017)).

Second, we do not address the causality issue in all its complexity. Natural characteristics such as proximity to a lake and the altitude are exogenous to income distribution, but almost all other variables are not. Regarding relationships such as jobs vs residences choices, it is not clear how we can get out of the chicken or egg dilemma (Muth, 1971). Quantifying the impact of taxes, housing rents or value-added on municipal average income is left to other studies using more refined identification strategies (e.g. Basten et al. (2017)).

Finally, the role of distance-related costs calls for further empirical analysis. Mobility and trade costs are important frictions in any spatial equilibrium model, but empirically they are often identified by deduction, rather than observation. Data on daily mobility and commuting flows should also be investigated, if possible over a long period of time. Better measures of distance-related costs are warranted to improve our understanding of key factors affecting the spatial distribution of income, in particular neighbourhood effects and agglomeration economies.

Bibliography

AFC (2013). Administration Fédérale des Contributions: introduction aux chiffres-clés de l'impôt fédéral direct. *Département fédéral des finances DFF*.

AFC (2018a). Administration Fédérale des Contributions: charge fiscale en Suisse. `https://www.estv.admin.ch/estv/fr/home/allgemein/steuerstatistiken/fachinformationen/steuerbelastungen/steuerbelastung.html`.

AFC (2018b). Administration Fédérale des Contributions: impôt fédéral direct. `https://www.estv.admin.ch/estv/fr/home/allgemein/steuerstatistiken/fachinformationen/steuerstatistiken/direkte-bundessteuer.html`.

Albouy, D. (2016). What are cities worth? Land rents, local productivity, and the total value of amenities. *Review of Economics and Statistics*, 98(3):477–487.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *Quaterly Journal of Economics*.

Allison, P. D. (2009). *Fixed effects regression models*, volume 160. SAGE publications.

Alstadsæter, A., Johannesen, N., and Zucman, G. (2017). Tax evasion and inequality. Technical report, National Bureau of Economic Research.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2):93–115.

Anselin, L., Le Gallo, J., and Jayet, H. (2008). Spatial panel econometrics. In *The econometrics of panel data*, pages 625–660. Springer.

ARE (2013). Office Fédéral du Développement territorial - Durées de déplacement et niveaux de qualité de desserte par les transports publics : Mise à jour des données géographiques et des cartes. *DETEC*.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Armstrong, H. W. (1995). Convergence among regions of the European Union, 1950–1990. *Papers in Regional Science*, 74(2):143–152.

Badinger, H., Müller, W., and Tondl, G. (2004). Regional convergence in the European Union, 1985-1999: A spatial dynamic panel analysis. *Regional Studies*, 38(3):241–253.

Baldwin, R., Forslid, R., Martin, P., Ottaviano, G., and Robert-Nicoud, F. (2005). Economomic geography and public policy. *Princeton University Press*.

Baltagi, B. (2013). *Econometric analysis of panel data*. John Wiley & Sons, 5th edition.

Baltagi, B. H., Bresson, G., and Pirotte, A. (2012). Forecasting with spatial panel data. *Computational Statistics & Data Analysis*, 56(11):3381–3397.

Banzhaf, H. S. and Walsh, R. P. (2008). Do people vote with their feet? An empirical test of Tiebout's mechanism. *The American Economic Review*, pages 843–863.

Baranzini, A. and Schaerer, C. (2011). A sight for sore eyes: Assessing the value of view and land use in the housing market. *Journal of Housing Economics*, 20(3):191 – 199.

Barro, R. J. and Sala-i Martin, X. (2003). *Economic Growth, 2nd Edition*, volume 1 of *MIT Press Books*. The MIT Press.

Basten, C., Ehrlich, M., and Lassmann, A. (2017). Income taxes, sorting and the costs of housing: evidence from municipal boundaries in Switzerland. *The Economic Journal*, 127(601):653–687.

Bayer, P. and McMillan, R. (2012). Tiebout sorting and neighborhood stratification. *Journal of Public Economics*, 96(11–12):1129 – 1143. Fiscal Federalism.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive Cities: Sorting, Selection, and Agglomeration. *Journal of Political Economy*, 122(3):507–553.

Behrens, K. and Robert-Nicoud, F. (2014). Survival of the fittest in cities: Urbanisation and inequality. *The Economic Journal*, 124(581):1371–1400.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1):115–143.

Blundell, R. and Bond, S. (2000). GMM estimation with persistent panel data: an application to production functions. *Econometric reviews*, 19(3):321–340.

Borck, R., Pflüger, M., and Wrede, M. (2010). A simple theory of industry location and residence choice. *Journal of Economic Geography*, 10(6):913–940.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *The American Economic Review*, pages 531–553.

Boualam, B. (2014). Does culture affect local productivity and urban amenities? *Regional Science and Urban Economics*, 46(C):12–17.

Bouayad-Agha, S. and Védrine, L. (2010). Estimation strategies for a spatial dynamic panel using GMM. A new approach to the convergence issue of European regions. *Spatial Economic Analysis*, 5(2):205–227.

Bourguignon, F. (2018). World changes in inequality: An overview of facts, causes, consequences, and policies. *CESifo Economic Studies*, 64(3):345–370.

Brueckner, J. K. (2000). Urban Sprawl: Diagnosis and remedies. *International Regional Science Review*, 23(2):160–171.

Brueckner, J. K., Thisse, J.-F., and Zenou, Y. (1999). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review*, 43(1):91–107.

Brülhart, M., Bucovetsky, S., and Schmidheiny, K. (2015). Taxes in cities. *Handbook of regional and urban economics*, 5:1123–1196.

Brülhart, M. and Parchet, R. (2014). Alleged tax competition: The mysterious death of bequest taxes in Switzerland. *Journal of Public Economics*, 111:63–78.

Caliendo, M., Parro, F., and Rossi-Hansberg, E. (2015). The impact of regional and sectoral productivity changes on the u.s. economy. *working paper*.

Castells-Quintana, D. and Royuela, V. (2015). Are increasing urbanisation and inequalities symptoms of growth? *Applied Spatial Analysis and Policy*, 8(3):291–308.

Cheshire, P., Monastiriotis, V., and Sheppard, S. (2003). Income inequality and residential segregation: labour market sorting and the demand for positional goods. *Geographies of labour market inequality*, pages 83–109.

Crozet, M. (2004). Do migrants follow market potentials? An estimation of a New Economic Geography model. *Journal of Economic Geography*, 4(4):439–458.

Crozet, M. and Koenig, P. (2005). The cohesion vs growth tradeoff - Evidence from EU regions (1980-2000). ERSA conference papers, European Regional Science Association.

Cytron, N. et al. (2011). Ties that bind: income inequality and income segregation. *Community Investments*, 23(2). Federal Reserve Bank of San Francisco.

Debarsy, N. (2012). The Mundlak approach in the spatial Durbin panel data model. *Spatial Economic Analysis*, 7(1):109–131.

Deininger, K. and Squire, L. (1996). A new data set measuring income inequality. *The World Bank Economic Review*, 10(3):565–591.

Dell, F., Piketty, T., Saez, E., et al. (2005). Income and Wealth Concentration in Switzerland Over the 20th Century. Technical report, CEPR Discussion Papers.

Dewenter, R. and Westermann, M. (2005). Cinema Demand in Germany. *Journal of Cultural Economics*, 29(3):213–231.

Eaton, J. and Kortum, S. (2002). Technology, geography and trade. *Econometrica*.

Ecoplan (2017). Charges de centre des villes - rapport de synthèse. *Conférence des directrices et directeurs des finances des villes (CDFV)*.

Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):pp. 554–620.

Egger, P. and Pfaffermayr, M. (2006). Spatial convergence. *Papers in regional science*, 85(2):199–215.

Elhorst, J. P. (2014). Spatial panel data models. In *Spatial econometrics*, pages 37–93. Springer.

Ertur, C. and Koch, W. (2007). Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of applied econometrics*, 22(6):1033–1062.

Eugster, B. and Parchet, R. (2019). Culture and taxes. *Journal of Political Economy*, 127(1):296–337.

Fahrländer, S. and Lehner, M. (2014). Wohn- und Lebensqualitäts-Index - Methodenbericht. *Fahrländer Partner AG*.

Falck, O., Fritsch, M., Heblich, S., and Otto, A. (2015). Music in the air: Estimating the social return to cultural amenities. Technical report, CESifo Working Paper.

Feld, L. P., Frey, C., Schaltegger, C. A., Schmid, L. A., et al. (2018). Fiscal federalism and income inequality: An empirical analysis for Switzerland. Technical report, CESifo Group Munich.

Feld, L. P. and Kirchgässner, G. (2001). Income tax competition at the state and local level in Switzerland. *Regional Science and Urban Economics*, 31(2):181–213.

Fingleton, B. (1999). Estimates of time to economic convergence: an analysis of regions of the European Union. *International regional science review*, 22(1):5–34.

Fingleton, B. (2014). Forecasting with dynamic spatial panel data: practical implementation methods. *Economics and Business Letters*, 3(4):194–207.

Fingleton, B. and López-Bazo, E. (2006). Empirical growth models with spatial effects. *Papers in regional science*, 85(2):177–198.

Fischer, L. B. and Pfaffermayr, M. (2018). The more the merrier? migration and convergence among european regions. *Regional Science and Urban Economics*, 72:103–114.

Fischer, M. and Stumpner, P. (2008). Income distribution dynamics and cross-region convergence in Europe. *Journal of Geographical Systems*, 10(2):109–139.

Foellmi, R. and Martínez, I. Z. (2017). Volatile top income shares in Switzerland? Reassessing the evolution between 1981 and 2010. *Review of Economics and Statistics*, 99(5):793–809.

Fretz, S., Parchet, R., and Robert-Nicoud, F. (2017). Highways, market access, and spatial sorting. *CEPR Discussion Paper No. DP12437*.

Fujita, M., Krugman, P., and Venables, A. J. (1999). *The Spatial Economy.* The MIT Press.

Fujita, M. and Thisse, J.-F. (2002). *Economics of agglomeration : cities, industrial location and regional growth.* Cambridge University Press.

Gallo, J. and Ertur, C. (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980–1995. *Papers in Regional Science*, 82(2):175–201.

Gallo, J. L. (2004). Space-time analysis of GDP disparities among European regions: a Markov chains approach. *International Regional Science Review*, 27(2):138–163.

Geary, F. and Stark, T. (2016). What happened to regional inequality in britain in the twentieth century? *The Economic History Review*, 69(1):215–228.

Getis, A. and Griffith, D. A. (2002). Comparative spatial filtering in regression analysis. *Geographical analysis*, 34(2):130–140.

Getis, A. and Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical analysis*, 24(3):189–206.

Helpman, E. (1998). The size of regions. *Topics in public economics: Theoretical and applied analysis*, pages 33–54.

Islam, N. (1995). Growth empirics: a panel data approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

Kakamu, K. and Fukushige, M. (2005). Divergence or convergence?: Income inequality between cities, towns and villages in Japan. *Japan and the World Economy*, 17(4):407–416.

Kim, J. H., Pagliara, F., and Preston, J. (2005). The intention to move and residential location choice behaviour. *Urban Studies*, 42(9):1621–1636.

Koster, H. R., van Ommeren, J. N., and Rietveld, P. (2014). Historic amenities, income and sorting of households. *Journal of Economic Geography*, 16(1):203–236.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499.

Kübler, D. and Rochat, P. E. (2017). "Tamed Tieboutianism" and Spatial Inequalities in Swiss Metropolitan Areas. In *Inequality and Governance in the Metropolis*, pages 107–124. Springer.

Kukenova, M. and Monteiro, J.-A. (2009). Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation. Mpra paper, University Library of Munich, Germany.

Lazarsfeld, P. F., Merton, R. K., et al. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18–66.

Lee, S. and Lin, J. (2017). Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *The Review of Economic Studies*, 85(1):663–694.

Levy, A. (1997). Median income versus per capita income: Implications for assessing growth and convergence. Technical report, School of Economics, University of Wollongong, NSW, Australia.

Liebig, T. and Sousa-Poza, A. (2006). The influence of taxes on migration: evidence from Switzerland. *Cambridge Journal of Economics*, 30(2):235–252.

Lin, J. et al. (2015). The puzzling persistence of place. *Federal Reserve Bank of Philadelphia Business Review Q*, 2:1–8.

López-Bazo, E., Vayá, E., and Artis, M. (2004). Regional externalities and growth: evidence from European regions. *Journal of regional science*, 44(1):43–73.

López-Bazo, E., Vayá, E., Mora, A. J., and Suriñach, J. (1999). Regional economic dynamics and convergence in the European Union. *The Annals of Regional Science*, 33(3):343–370.

Mabut, J.-F. (2015). Anières doit son pactole fiscal à un grand horloger, Tribune de Genève. https://www.tdg.ch/geneve/actu-genevoise/anieres-doit-pactole-fiscal-grand-horloger/story/24943561.

Marcouiller, D. W., Kim, K.-K., and Deller, S. C. (2004). Natural amenities, tourism and income distribution. *Annals of tourism research*, 31(4):1031–1050.

Marshall, A. (1920). *Principles of Economics*. Macmillan, London.

Martin, P. (2009). The geography of inequalities in Europe. *Spatial Disparities and Development Policy, Washington, DC: The World Bank*, pages 239–256.

Martínez, I. (2016). Beggar-thy-neighbour tax cuts: Mobility after a local income and wealth tax reform in Switzerland. Technical report, Verein für Socialpolitik/German Economic Association.

Martínez, I. (2017). Les hauts revenus en Suisse depuis 1980: répartition et mobilité suisse. *Social Change in Switzerland*, 11. FORS.

Mazek, W. F. and Chang, J. (1972). The chicken or egg fowl-up in migration: comment. *Southern Economic Journal*, pages 133–139.

Milanovic, B. (1999). *True World Income Distribution, 1988 and 1993: First Calculations, Based on Household Surveys Alone*. The World Bank.

Mincer, J. (1978). Family migration decisions. *Journal of Political Economy*, 86(5):749–773.

Moreno, R. and Trehan, B. (1997). Location and the growth of nations. *Journal of Economic Growth*, 2(4):399–418.

Morger, M. (2017). Heterogeneity in Income Tax Capitalization: Evidence from the Swiss Housing Market. *Swiss Journal of Economics and Statistics*, 153(3):227–259.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society*, pages 69–85.

Müri Leupp, R., Rappl, I., Bröhl, A., Walker, U., Fischer, F., and Ingold, K. (2011). Ruhe bitte! Wie Lage und Umweltqualität die Schweizer Mieten bestimmen. *Zürcher Kantonalbank*.

Muth, R. F. (1971). Migration: chicken or egg? *Southern Economic Journal*, pages 295–306.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, pages 1417–1426.

Oates, W. E. (1969). The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the Tiebout hypothesis. *Journal of political economy*, 77(6):957–971.

OFS (2018a). Les résultats comptables des entreprises suisses - années fiscales 2015-2016. *Office fédérale de la statistique - Section Statistique monétaire des entreprises.*

OFS (2018b). Office fédérale de la statistique - Application Suisse des communes. `https://www.agvchapp.bfs.admin.ch/fr/mutated-communes/query`.

OFS (2019). Office fédérale de la statistique - Consumer Prices. `https://www.bfs.admin.ch/bfs/en/home/statistics/prices/consumer-price-index.html`.

Openshaw, S. (1984). The modifiable areal unit problem. *Concepts and techniques in modern geography.*

Piketty, T. (2013). *Le Capital au XXIe siècle.* Les Livres du nouveau monde. Le Seuil.

Piketty, T. and Saez, E. (2006). The evolution of top incomes: a historical and international perspective. *American Economic Review*, 96(2):200–205.

Portnov, B. A., Axhausen, K. W., Tschopp, M., and Schwartz, M. (2011). Diminishing effects of location? Some evidence from Swiss municipalities, 1950–2000. *Journal of Transport Geography*, 19(6):1368–1378.

Quah, D. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review*, 37(2):426–434.

Rey, S. (2016). Space-time patterns of rank concordance: Local indicators of mobility association with application to spatial income inequality dynamics. MPRA Paper 69480, University Library of Munich, Germany.

Rey, S. J. (2004a). Spatial analysis of regional income inequality. *Spatially integrated social science*, 1:280–299.

Rey, S. J. (2004b). Spatial dependence in the evolution of regional income distributions. *Spatial econometrics and spatial statistics*, pages 194–213.

Rey, S. J. and Montouri, B. D. (1999). US regional income convergence: a spatial econometric perspective. *Regional studies*, 33(2):143–156.

Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of Political Economy*, pages 1257–1278.

Roller, M. and Schmidheiny, K. (2016). Effective tax rates and effective progressivity in a fiscally decentralized country. Technical report, CEPR Discussion Papers.

Rosen, K. T. and Smith, L. B. (1983). The price-adjustment process for rental housing and the natural vacancy rate. *The American Economic Review*, 73(4):779–786.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2119–2171. Elsevier.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers*, 3(2):135–146.

Schaeffer, Y., Cremer-Schulte, D., Tartiu, C., and Tivadar, M. (2016). Natural amenity-driven segregation: Evidence from location choices in French metropolitan areas. *Ecological Economics*, 130:37–52.

Schaltegger, C. A., Somogyi, F., and Sturm, J.-E. (2011). Tax competition and income sorting: Evidence from the Zurich metropolitan area. *European Journal of Political Economy*, 27(3):455 – 470.

Schirmer, P., Van Eggermond, M., and Axhausen, K. (2013). Measuring location in residential location choice: an empirical study in the Canton of Zurich. In *13th International Conference on Computers in Urban Planning and Urban Management, Utrecht, the Netherlands*, pages 2–5.

Schmidheiny, K. (2006). Income segregation and local progressive taxation: Empirical evidence from Switzerland. *Journal of Public Economics*, 90(3):429 – 458. Special issue published in cooperation with the National Bureau of Economic Research: Proceedings of the Trans-Atlantic Public Economics Seminar on Fiscal Federalism 20–22 May 2004.

Schmidheiny, K. (2017). Emerging lessons from half a century of fiscal federalism in Switzerland. *Swiss Journal of Economics and Statistics*, 153(2):73–101.

Schunck, R. (2013). Within and between estimates in random-effects models: Advantages and drawbacks of correlated random effects and hybrid models. *Stata Journal*, 13(1):65–76(12).

Segessemann, A. and Crevoisier, O. (2016). Beyond economic base theory: The role of the residential economy in attracting income to Swiss regions. *Regional Studies*, 50(8):1388–1403.

Shorrocks, A. and Wan, G. (2005). Spatial decomposition of inequality. *Journal of Economic Geography*, 5(1):59–81.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48(3):613–625.

Sinha, P. and Cropper, M. L. (2015). Household location decisions and the value of climate amenities. Working Paper 21826, National Bureau of Economic Research.

Sjaastad, L. A. (1962). The costs and returns of human migration. *The Journal of Political Economy*, pages 80–93.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1):65–94.

Stadelmann, D. and Billon, S. (2012). Capitalisation of fiscal variables and land scarcity. *Urban Studies*, 49(7):1571–1594.

Stiglitz, J., Sen, A. K., and Fitoussi, J.-P. (2007). Report by the Commission on the Measurement of Economic Performance and Social Progress.

Stiglitz, J. E. (2012). *The price of inequality: How today's divided society endangers our future.* WW Norton & Company.

Throsby, D. (1994). The production and consumption of the arts: a view of cultural economics. *Journal of Economic Literature*, pages 1–29.

Tiebout, C. M. (1956). A pure theory of local expenditures. *The Journal of Political Economy*, pages 416–424.

Tissot-Daguette, B. (2019). Measuring agglomeration economies in Switzerland. *IRENE Working Papers*, (tbd).

Tissot-Daguette, B. and Grether, J.-M. (2019). Multiple imputation techniques: An application to swiss value-added data. *IRENE Working Papers*, (tbd).

van Duijn, M. and Rouwendal, J. (2015). Sorting based on Urban Heritage and Income: Evidence from the Amsterdam Metropolitan Area. Technical report, Tinbergen Institute.

Waltert, F., Schulz, T., and Schläpfer, F. (2011). The role of landscape amenities in regional development: Evidence from Swiss municipality data. *Land Use Policy*, 28(4):748–761.

WID (2019). World Inequality Database - Switzerland. `https://wid.world/country/switzerland/`.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Young, A. T., Higgins, M. J., and Levy, D. (2008). Sigma convergence versus beta convergence: Evidence from US county-level data. *Journal of Money, Credit and Banking*, 40(5):1083–1093.

Yu, J. and Lee, L.-f. (2012). Convergence: A spatial dynamic panel data approach. *Global Journal of Economics*, 1(01):1250006.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 57(298):348–368.

Zondag, B. and Pieters, M. (2005). Influence of accessibility on residential location choice. *Transportation Research Record: Journal of the Transportation Research Board*, (1902):63–70.