UNIVERSITY OF LJUBLJANA
SCHOOL OF ECONOMICS AND BUSINESS

MASTER'S THESIS

AN EMPIRICAL ANALYSIS OF DAILY MIGRATION IN SLOVENIA

Ljubljana, December 2022

MARINA TOMIĆ

AUTHORSHIP STATEMENT

The undersigned Marina Tomić, a student at the University of Ljubljana, School of Economics and Business, (hereafter: SEB LU), author of this written final work of studies with the title An Empirical Analysis of Daily Migration in Slovenia, prepared under supervision of prof. Sašo Polanec, PhD.

D E C L A R E

1. this written final work of studies to be based on the results of my own research;
2. the printed form of this written final work of studies to be identical to its electronic form;
3. the text of this written final work of studies to be language-edited and technically in adherence with the SEB LU's Technical Guidelines for Written Works, which means that I cited and / or quoted works and opinions of other authors in this written final work of studies in accordance with the SEB LU's Technical Guidelines for Written Works;
4. to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Criminal Code of the Republic of Slovenia;
5. to be aware of the consequences a proven plagiarism charge based on the this written final work could have for my status at the SEB LU in accordance with the relevant SEB LU Rules;
6. to have obtained all the necessary permits to use the data and works of other authors which are (in written or graphical form) referred to in this written final work of studies and to have clearly marked them;
7. to have acted in accordance with ethical principles during the preparation of this written final work of studies and to have, where necessary, obtained permission of the Ethics Committee;
8. my consent to use the electronic form of this written final work of studies for the detection of content similarity with other written works, using similarity detection software that is connected with the SEB LU Study Information System;
9. to transfer to the University of Ljubljana free of charge, non-exclusively, geographically and time-wise unlimited the right of saving this written final work of studies in the electronic form, the right of its reproduction, as well as the right of making this written final work of studies available to the public on the World Wide Web via the Repository of the University of Ljubljana;
10. my consent to publication of my personal data that are included in this written final work of studies and in this declaration, when this written final work of studies is published.

Ljubljana, ________________________ Author's signature: _________________________

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

EU – European Union

SARS-COV-2 – severe acute respiratory syndrome coronavirus 2

AMM – Alonso-Muth-Mills

SURS – Statistical Office of the Republic of Slovenia

OLS – Ordinary Least Squares

PPML - Poisson Pseudo-Maximum-Likelihood

INTRODUCTION

The journey to work has long been recognized as an important factor in the daily lives of individuals. This is especially relevant since the majority of individuals travel to and from work daily. According to the IZA Institute of Labor Economics commuting to and from work is one of the most important trips in a worker's day (Giménez-Nadal, Molina Chueca & Velilla, 2020). According to the study: "one out of five workers in Europe spend more than 90 minutes commuting each day, equivalent to about 29 km distance (Giménez-Nadal, Molina Chueca & Velilla, 2020, p. 1). By 2050, more than 82% of EU citizens, are estimated to be working in an urban area that is different from their place of residence (ibid.). There is a body of literature dedicated to studying the social, economic and financial effects of daily commuting, most of which is essential for the understanding of human mobility patterns and behaviors, as well as having wider implications for urban planning and policy design (Giménez-Nadal, Molina Chueca & Velilla, 2020; Lyons & Chatterjee, 2008). Given the lack of research on daily migration in Slovenia, it is the scope of this thesis to fill the gap in the literature and provide some new insights based on empirical analysis.

The purpose of this thesis is to examine the commuting patterns of workers in Slovenia and characterize the geographical structure of Slovenia's labor market. Firstly, we will provide the characteristics of commuting flows in Slovenia based on data at the level of county of residence and work for all employed persons in Slovenia, generally, by gender and educational attainment. Secondly, we will examine the structure of Slovenia's labor market by generating flow-based clusters dependent on the identification of regional employment centers, using a clustering algorithm applied to data on bilateral commuting flows. We perform this analysis by imposing several restrictions on the model and applying it to data on general commuting flows as well as commuting flows by education level. Lastly, we will discuss our estimates of distance elasticities for the labor market in Slovenia, using the Poisson Pseudo-Maximum-Likelihood estimator. It is our primary goal to assess the effect of commuting times on the willingness of workers to commute and we expand our analysis to include the effects of education, gender, and wages on commuting patterns in Slovenia.

We aim to provide an estimate of distance elasticities for employment for the labor market in Slovenia and measure how these elasticities change in time, by gender as well as by level of education. We also hope to identify the possible effect of wages on commuting flows as well as provide a general analysis of the geographical structure of the labor market in Slovenia. The results of this thesis could have possible implications for labor policy, specifically regarding job referral strategies the Employment Service of Slovenia employs when matching job opportunities to potential candidates. We believe our empirical analysis could shed light on the commuting patterns of Slovenian workers and show just how far different groups of employees are willing to commute for work.

1 LITERATURE OVERVIEW

1.1 The Daily Commute

The word “commute” has become a well-known term describing the daily journey from home to work and vice versa. In contrast to the term “travelling”, commuting refers to the repeated nature of the journey, usually associated with work, as opposed to “travelling”, which mostly implies leisurely activities (Lyons & Chatterjee, 2008). In economics “commuting” is an action characterized by the spatial interaction between housing and labor markets. It generally connects an individual’s place of work and residence, and provides information about home location relative to the place of employment (Jacob, Munford, Rice & Roberts, 2019; O’Kelly, Niedzielski & Gleeson, 2012). As Horton and Wittick explain: “The journey-to-work has long been recognized as an important factor in the residential-location decision process. When selecting a new home people will usually choose a location which limits their journey-to-work trip to a reasonable length […]” (Horton & Wittick, 1969, p. 223). Since the article was published, the field has seen increasing interest by researchers to understand the connections between households, location, housing choice and journey-to-work travel distance, as we will see in the forthcoming chapters (O’Kelly, Niedzielski & Gleeson, 2012).

Spatial labor markets are characterized by demand for workers with a certain skillset that generally differs from the supply of labor with such capabilities. Thus, commuting presents a necessary mechanism for reaching spatial equilibrium in the labor market (Evers, 1989; Persyn & Torfs, 2016). As Persyn and Torfs explain: “in standard closed-economy labor market models, commuting reduces disparities in regional labor market outcomes such as unemployment rates and wages, and brings aggregate welfare gains” (Persyn & Torfs, 2016, p. 155). However, we must not neglect the fact that commuting has its cost. There are obvious costs that are directly related to commuting, such as travel expenses and time-based opportunity costs, and more indirect, nuanced costs, such as linguistic barriers, informational deficiencies and cultural barriers (Persyn & Torfs, 2016). As Lyons and Chatterjee point out, there are larger economic, health and social effects of commuting that are largely neglected in the literature, that have an undeniable effect on the daily lives of individuals and their work-life balance (Lyons & Chatterjee, 2008). Nevertheless, commuting has become an integral part of today’s society. In an economy where both people and jobs are becoming increasingly heterogeneous – looking at skills, qualifications, and location preferences – an equilibrium is becoming harder to achieve. The most obvious solution to this problem is of course labor mobility (Evers, 1989).

The majority of working individuals travel to and from work daily. According to the IZA Institute of Labor Economics commuting to and from work is one of the most important trips in a worker’s day (Giménez-Nadal, Molina Chueca & Velilla, 2020). According to the study: “one out of five workers in Europe spend more than 90 minutes commuting each day, equivalent to about 29 km distance (Giménez-Nadal, Molina Chueca & Velilla, 2020, p. 1). By 2050, more than 82% of EU citizens, are estimated to be working in an urban area that is different from their place of residence (ibid.). The trends are similar in the U.S, where employees devote, on average, 38 minutes per day to commuting, with the number increasing each year (Giménez-Nadal, Molina & Velilla, 2019). Since then, the SARS-CoV-2 pandemic has disrupted the normal travel patterns of working individuals, with no indication of when these trends will converge back to pre-Covid levels. Nevertheless, these numbers prove to be worrisome since studies have shown, that commuting is one of the least enjoyable activities in an individual's daily life (Kahneman, Krueger, Schkade, Schwarz & Stone, 2004). Commuting can also cause high levels of stress and can have a negative effect on an individual's social and family life as well as general welfare (Novaco & Gonzalez, 2009; Stone & Schneider, 2016).

Given its importance, the journey to work is essential in daily mobility planning. As such, research devoted to understanding commuting trends and behaviors is crucial in mobility planning and policy design. As the research suggests, there are social, economic, and financial benefits to improving the commuting experience and lowering the average commuting time (Giménez-Nadal, Molina Chueca & Velilla, 2020; Lyons & Chatterjee, 2008). It is also the purpose of this thesis, to shed light on the commuting experience in Slovenia, that could have further implications, not only for economic research but policy change as well. In the next chapter, we will look at previous research and empirical analyses done on the topic of commuting, within different theoretical and methodological frameworks.

1.2 Previous Research

Understanding the basic laws of human mobility is of fundamental importance as it relates to problems such as traffic forecasting, congestion alleviation, social stability and even disease control (González, Hidalgo & Barabási, 2008; Yan, Wang, Gao & Lai, 2017). Numerous studies find that human mobility patterns mimic those found in the animal kingdom, specifically foraging and hunting patterns of animals (Brockmann, Hufnagel & Geisel, 2006; González, Hidalgo & Barabási, 2008; Reynolds & Rhodes, 2009; Yan, Wang, Gao & Lai, 2017). For many years, the presiding conceptual model describing animal movement was based on the Brownian motion or random walk. The theory states that a movement through space is made up of distinct, random movements drawn from a Gaussian distribution (Kareiva & Shigesada, 1983). However, new research has shown that animal trajectories show more similarities to a Lévy flight (Klafter, Shlesinger & Zumofen, 1996; Mantegna & Stanley, 1994). A Lévy flight is a random walk for which the step length distribution obeys a power law and for which the probability distribution is long-tailed. As Reynolds and Rhodes explain, a Lévy flight is a "frequently occurring but relatively short straight-line movement randomly alternate with more occasionally occurring longer movements, and so on with this pattern repeated on all scales" (Reynolds

& Rhodes, 2009, p. 879). Such a single-step distribution has been shown to maximize foraging and hunting in uncertain environments and has been documented in movements of albatrosses (Viswanathan et al., 1996), bees (Reynolds, 2008) and other animals (Reynolds & Rhodes, 2009), as well as humans (Brockmann, Hufnagel & Geisel, 2006).

In their 2006 article titled: "The Scaling Laws of Human Travel" Brockmann et al. used the circulation of more than half a million U.S. banknotes as a proxy for human movement, the trajectories of which were tracked over a semi-long time period. The study showed the existence of two scaling laws: (1) the distribution of the traveled distance of banknotes revealed a power-law decay, which is best described as a Lévy flight, and (2) the probability of staying in a smaller region mimics a long-tailed distribution with an exponential cutoff, which is consistent with superdiffusive behavior. As such, an individual following a Lévy flight has a significant probability of travelling long distances in a single step, which is in line with human travel patterns (Brockmann, Hufnagel & Geisel, 2006; Yan, Wang, Gao & Lai, 2017). For the most part, humans travel over relatively short distances – for instance commuting from home to work – while sometimes we take longer trips. As such, these trajectories are best modelled by a continuous-time random walk with incorporated scale-free jumps, as well as waiting time distributions (Brockmann, Hufnagel & Geisel, 2006; González, Hidalgo & Barabási, 2008).

Other studies have shown, that mobile phones appear to be a better proxy for capturing human movement patterns (Hidalgo & Rodriguez-Sickert, 2008; Onnela et al., 2007; Palla, Barabási & Vicsek, 2007). In their 2008 study, González et al. studied the trajectory of one hundred thousand anonymous mobile phone users over a six-month period. Measuring the distribution of displacements, the authors find that compared to the Lévy flight and random walk models, human movements show a significant level of temporal and spatial regularity. This would suggest that human movement patterns are not completely random and unpredictable, since individuals have a tendency to return to frequently visited locations, such as home or work (González, Hidalgo & Barabási, 2008). This regularity was not detected in banknotes since the banknotes followed the path of their current owner and not the path of a single person. Thus, dollar bills diffuse, while humans do not (ibid.). Han et al. (2011) took a step further. They introduced a geographical model that simulates the hierarchy of a real-life traffic system, upon which they were able to generate a random walk scenario that was able to reproduce the power-law displacement distribution as discussed in Brockmann et al. (2006), showing that the Lévy flight theory can be applied to more complex traffic systems (Han, Hao, Wang & Zhou, 2011).

Other authors disagree. Song et al. (2010) argue that continuous-time random walk models are in direct conflict with empirical results. The authors show, using data from 3 million mobile phone users collected over a one-year period, that human trajectories follow various scaling laws, that cannot be explained with Lévy flight models or directly contradict them. These scaling laws show that (1) the tendency of humans to travel to new locations decreases with time and (2) there is a high probability that they will return to locations they frequently visited before, such as home or work, suggesting that human trajectories are all but random. Instead, the authors focus on two mechanisms unique to human mobility – exploration and preferential return. The first challenges the assumption that the next step in a random walk scenario is independent of the previously visited location and the latter incorporates the tendency of humans to return to previously visited locations. By doing so, the authors are able to generate an individual mobility model that is able to account for these scaling anomalies, while providing the necessary framework for capturing the basic properties of human mobility (Song, Koren, Wang & Barabási, 2010).

However, by far the most popular modeling framework in the field of human mobility has been the gravity model (Choukroun, 1975; Wilson, 1998; Zipf, 1946). Derived from Newton's law of gravity, the gravity model is one of the principal contributions in the field of spatial interaction models. Its simple formulation and general applicability have made it the predominant framework in the field of international trade, traffic flow and congestion alleviation, as well as commuting and migration (O'Kelly, 2009; Rodrigue, 2020). The basic premise of the model is the expectation that there is a positive association between the size of a city (or population) and the volume (or flow) of commuters and migrants to that location, mediated by distance, cost and travel time (Fotheringham, 2001). In other words: "the gravity law assumes that the number of individuals T_{ij} that move between locations i and j per unit of time is proportional to some power of the population of the source (m_i) and destination (n_j) locations, and decays with the distance r_{ij} between them" (Simini, González, Maritan & Barabási, 2012, p. 96). The formula is thus:

$$T_{ij} = \frac{m_i^{\alpha} n_j^{\beta}}{f(r_{ij})} \tag{1}$$

where α and β are adjustable exponents. Despite its extensive use, the gravity model has some significant limitations. Firstly, the model has an inadequate theoretical framework, thus diminishing it's predictive capacity (Fotheringham, 2001; Ramos, 2016). Secondly, as pointed out by Simini et al. (2012), the model suffers from several inconsistencies, such as (1) lack of derivation for the principal equation, (2) the need for extensive trajectory data for proper mobility prediction and (3) several analytical inconsistencies, such as allowing the number of commuters to exceed the source population m_j etc. As a solution to these issues, several authors recognized radiation models as the better option in predicting the patterns of human mobility (Lenormand, Huet, Gargiulo & Deffuant, 2012; Ren, Ercsey-Ravasz, Wang, González & Toroczkai, 2014; Simini, González, Maritan & Barabási, 2012). Simini et al. (2012) proposed a stochastic radiation model that derives mobility fluxes using only information on population distribution. The model captures local mobility decisions based on job opportunities. The main assumption is that individuals will choose the closest job opportunity that provides the most benefit. The proposed model is able to overcome several of the limitations found in the gravity model and provides a consistent estimate of human mobility and migration patterns, that can be applied to several fields such as urban geography, resource flow economics, epidemiology etc. (Simini, González, Maritan & Barabási, 2012).

Similar to Simini et al., Yan et al. (2017) diverge from the classic gravity equation and propose a universal model of human mobility using memory effect and population-induced competition to predict human mobility patterns, relying on population distribution only. The resulting model is able to capture human mobility at different levels and spatial scales while accounting for the above-mentioned scaling laws (Yan, Wang, Gao & Lai, 2017). Lenormand et al. (2012) follow a similar method, using a doubly-constrained gravity model along with an individual-based approach to construct a universal model that is more flexible and less data demanding. The authors show that the model is able to follow a simple universal law – dependent only on the average surface of geographic units – and is able to derive a matrix of flows with a very good confidence, thus outperforming the universal model proposed by Simini et al. (Lenormand, Huet, Gargiulo & Deffuant, 2012). Ren et al. (2014) also build on the radiation model, expanding the model to commuter flows, by adding a cost-minimizing algorithm to solve the flux distribution problem. The resulting model is able to capture the log-normal distribution of traffic and accurately predict traffic and roadway networks (Ren, Ercsey-Ravasz, Wang, González & Toroczkai, 2014).

As we have seen random walk models and spatial interaction models have proven to be a necessary benchmark for understanding the general laws and principles of human movements. However, human mobility is not an isolated incident. How people move is directly related to where they choose to live, work, spend their free time and how they interact with other people (So, Orazem & Otto, 2001). In order to fully understand commuting choices and patterns, we must consider the effects of variables such as job opportunities, rent, wages and location preferences, when modeling human mobility. Firstly, let’s look at location and property prices. An array of literature relates property prices to spatial structures. The basis of which are the classic works of Alonso (1964), Mills (1967), and Muth (1969), also known as the Alonso-Muth-Mills model. The authors develop a monocentric city model to prove the existence of rent gradients and show that property prices increase the closer we are to the central business district, where all the job opportunities are located. Similarly, cost-of travel is increasing the further we move from the center, indicating a trade-off between affordable housing and lengthy commutes, for workers solving the residence choice problem (Alonso, 1964; Mills, 1967; Muth, 1969).

The basic implications of the AMM model are, that if property prices increase disproportionately – affecting the central business district more so than the periphery – then it is very likely that people will relocate to the suburbs to avoid higher rents, settling for longer commutes (Ahrens & Lyons, 2021). There are several studies[1] that find empirical evidence to support the finding of the AMM model, however, none of them

[1] (Ahlfeldt, 2011; Boarnet, 2005; Carlino & Mills, 1987; Freedman & Kern, 1997; Renkow & Hoover, 2000)

consider the full scope of parameters, such as commuting costs, housing prices, wages, job opportunities and residence location (So, Orazem & Otto, 2001). So et al. try to close this gap by modelling the decision of where to live and where to work, using a restricted multinomial logit approach. The authors show that the probability of choosing a certain residence is negatively affected by housing price, but positively affected by wage levels, which is in line with the AMM model. People are more likely to commute, if the wages in that market are higher, meaning that on average commuters have higher wages than non-commuters. Since commuting is modelled as a cost, the probability of commuting is negatively affected by commuting time/distance. As such the probability of commuting decreases with the distance of the commute, with a one hour commute being the maximum (So, Orazem & Otto, 2001).

A study done by O'Kelly et al. (2012) also supports the findings of Alonso (1964). The authors attempt to model the direct benefits of accessible housing, using a spatial interaction model applied to commuting data for Ireland. The approach uses shadow prices as a proxy for measuring rent and wage gradients as found in Alonso, showing that housing located closer to high-income jobs will command a rent premium over other locations (O'Kelly, Niedzielski & Gleeson, 2012). This suggests that people with higher wages can afford housing with better amenities and tend to live closer to the central business district, pushing people with low wages to less desirable locations (Alonso, 1964; O'Kelly, Niedzielski & Gleeson, 2012). Ahrens and Lyons (2021) use the same data for Ireland to model the effects of rental price change on commuting times, using a gravity model of bilateral commuting flows. The study finds that a 10% rise in rents in central business districts or "employment centers" correlates with a 0.6 minute longer commute, nationally, which represents 2.2% of the average commuting time (Ahrens & Lyons, 2021). These findings have several implications for urban policy and design – a disproportional increase in rents in the center can lead to urban sprawl (Travisi, Camagni & Nijkamp, 2010) or result in excess commuting (Ma & Banister, 2006), both of which negatively impact urban planning and traffic alleviation efforts.

Other authors use different methods to study the commuting-housing nexus. Manning and Petrongolo (2017) study the relationship between housing and commuting through unemployment and job vacancies. They propose a job search model to measure how local are labor markets. The study finds that distance has its cost - the probability of a person taking a job that is 5 kilometers away, as opposed to one nearby, is only 19%. Furthermore, the study finds that workers are less likely to apply to job positions with stronger competition, implying that larger markets do not necessarily offer a better matching of workers and jobs (Manning & Petrongolo, 2017). Monte et al. (2018) develop a general equilibrium model to measure the determinants of local distance elasticities. The authors show that commuting flows display a gravity relationship with a higher distance elasticity in comparison to trade flows, implying that it is more costly to move people than goods. The study also finds that countries with more open labor markets, exhibit a higher increase in employment when exposed to a positive labor market shock (Monte, Redding & Rossi-Hansberg, 2018). Other significant contributions include the use of employment potential as a proxy for measuring the impact of employment on land prices (Ahlfeldt & Wendland, 2016) and the use of commuting flows measured by location "attractiveness" to predict the spatial distribution of wages (Kreindler & Miyauchi, 2021).

There are several other relevant perspectives in the field of human mobility that are outside the scope of this thesis, but we will cover them briefly. A significant array of literature is devoted to understanding commuting behavior based on modal choice (Commins & Nolan, 2011; Tiwari & Kawakami, 2001; Vega & Reynolds-Feighan, 2008). In recent years the car has become the dominant mode of transport in many countries, allowing workers to live further away from work. A consequence of this have been increased traffic volumes and more frequent congestions in urban areas (Gutiérrez-i-Puigarnau & van Ommeren, 2015; McArthur, Kleppe, Thorsen & Ubøe, 2011; Ren, Ercsey-Ravasz, Wang, González & Toroczkai, 2014). Another relevant, yet seldom researched area in commuting is the gender gap. Research shows that there is still a visible gap between how men and women commute, with men commuting, on average, longer than women (Crane, 2007). This gap increases even further in households with children (Vega & Reynolds-Feighan, 2008). Other studies research the effects of trade shocks (Artuç, Chaudhuri & McLaren, 2010), wage inequality (Artuç & McLaren, 2015) and other social, economic and financial effects of commuting (Lyons & Chatterjee, 2008). In the next chapter, we will look at the urban characteristics and regional development of Slovenia, as well as introduce the previous research conducted on daily migration in Slovenia.

1.3 Daily Commuting in Slovenia

One of the most noticeable and impactful phenomena of the 21st century, has without a doubt, been urbanization. The fast and exponential growth of urban settlements has been documented globally, with researchers estimating that by 2040, more than 70% of all people will be living in cities (Vodeb et al., 2016). This is not the case for Slovenia. Unlike fast-growing countries, Slovenia is characterized by a low urbanization rate and a dispersed spatial structure indicative of suburbanization[2]. According to recent data, the share of people in Slovenia who live in cities is less than 50%, with the United Nations predicting that urbanization will increase to only about 61% by 2050 (United Nations, 2014). Furthermore, more than 80% of municipalities are in sparsely populated areas, with only Ljubljana and Maribor being considered densely populated. As a consequence, Slovenian towns are distinctively small. As Vodeb et al. (2016) explain "95% of settlements have less than 1000 inhabitants, 47% of settlements have less than 100

[2] Suburbanization can be a potentially harmful phenomena causing a spatial misallocation of labor across the country. As Hsieh and Moretti (2019) point out, suburbanization can be a direct result of improper zoning restrictions, poor land use regulation or even a shortage of housing supply (Hsieh & Moretti, 2019).

inhabitants and 59 settlements have no inhabitants" (Vodeb et al., 2016, p. 7). The unique makeup of Slovenia's urban structure has led to several statistical inconsistencies. For example, while the Statistical Office of the Republic of Slovenia (SURS) recognizes 156 urban settlements, Eurostat recognizes only two – Ljubljana and Maribor (ibid.).

Slovenia is a fascinating case even in terms of regional development. In their 2003 article, Rovan and Sambt analyzed the socio-economic differences between Slovenian municipalities, dividing them into four distinct groups. The authors considered demographic variables, such as ageing and population growth, social variables, like unemployment and number of students, as well as several economic variables. Based on these socio-economic indicators, the authors were able to determine four groups of municipalities ranging from most developed to least developed. The most developed municipalities[3] consisted of the capital Ljubljana with adjourning municipalities, including the Primorska region. These municipalities scored the highest in population growth, standard of living and had the lowest unemployment rate of all four groups. The second group[4] consisted predominantly of eastern municipalities, including Maribor, Jesenice, Krško and Ptuj. Based on the socio-economic indicators, these municipalities are still considered developed, scoring just below the first group. The third group[5] consisted mainly of less attractive areas in the eastern part of Slovenia, around Kozjansko, Haloze and Slovenske Gorice. These municipalities were deemed less developed, due to a low number of students and commuters, as well as a high agricultural population. The last group[6] consisted mostly of smaller municipalities near the borders as well as 11 municipalities

[3] Ajdovščina, Bled, Bohinj, Borovnica, Brda, Brezovica, Cerklje na Gorenjskem, Cerknica, Cerkno, Divača, Dobrova-Polhov Gradec, Dol pri Ljubljani, Dolenjske Toplice, Domžale, Gorenja Vas-Poljane, Grosuplje, Horjul, Hrpelje-Kozina, Idrija, Ig, Ivančna Gorica, Izola, Kamnik, Kanal, Komen, Komenda, Koper, Kranj, Kranjska Gora, Ljubljana, Logatec, Loška dolina, Lukovica, Medvode, Mengeš, Miren-Kostanjevica, Mislinja, Moravče, Naklo, Nova Gorica, Novo mesto, Piran, Postojna, Preddvor, Sežana, Šempeter-Vrtojba, Šenčur, Škofja Loka, Škofljica, Šmartno ob Paki, Trzin, Velike Lašče, Vipava, Vodice, Vrhnika, Železniki, Žiri, Žirovnica.

[4] Bovec, Brežice, Celje, Črna na Koroškem, Črnomelj, Dobrepolje, Dravograd, Gornja Radgona, Hoče-Slivnica, Hrastnik, Ilirska Bistrica, Jesenice, Jezersko, Kidričevo, Kobarid, Kočevje, Krško, Laško, Lenart, Lendava, Litija, Ljutomer, Maribor, Metlika, Mežica, Miklavž na Drav. Polju, Mozirje, Murska Sobota, Muta, Nazarje, Pivka, Polzela, Prebold, Prevalje, Ptuj, Rače-Fram, Radeče, Radenci, Radlje ob Drave, Radovljica, ravne na Koroškem, Ribnica, Rogaška Slatina, Ruše, Semič, Sevnica, Slovenj Gradec, Slovenska Bistrica, Slovenske Konjice, Šentilj, Šentjernej, Šentjur pri Celju, Šoštanj, Štore, Tolmin, Trbovlje, Trebnje, Tržič, Velenje, Vojnik, Vuzenica, Zagorje ob Savi, zreče, Žalec.

[5] Beltinci, Benedikt, Bloke, Braslovče, Cerkvenjak, Črenšovci, Destrnik, Dobje, Dobrna, Dornava, Duplek, Gorišnica, Gornji Grad, Hajdina, Juršinci, Kobilje, Kozje, Križevci, Kungota, Ljubno, Loški Potok, Lovrenc na Pohorju, Luče, Majšperk, Markovci, Mirna Peč, Odranci, Oplotnica, Ormož, Pesnica, Podčetrtek, Podlehnik, Podvelka, Razkrižje, Ribnica na Pohorju, Rogatec, Selnica ob Dravi, Sodražica, Solčava, Starše, Sveta Ana, Sveti Andraž v Slov. goricah, Sveti Jurij, Škocjan, Šmarje pri Jelšah, Tabor, Tišina, Trnovska vas, Turnišče, Veržej, Videm, Vitanje, Vransko, Zavrč, Žetale, Žužemberk.

[6] Bistrica ob Sotli, Cankova, Dobrovnik, Gornji Petrovci, Grad, Hodoš, Kostel, Kuzma, Moravske Toplice, Osilnica, Puconci, Rogašovci, Šalovci, Velika Polana.

from the north-eastern part of Slovenia, Goričko. The most concerning aspect surrounding this group of municipalities is their level of (under)development. The group scored the lowest in income per capita, standard of living as well as unemployment rate in the country. These findings shed light on the developmental gap between the well-developed west and the less-developed east (Rovan & Sambt, 2003).

These results were confirmed by Bole (2004), showing that the Municipality of Ljubljana has by far, the widest employment attraction, offering the highest diversity and number of employment places relative to its population. The study showed that there were only twenty-four municipalities with less than 1% of commuters travelling to Ljubljana. The extent of the city's employment power was attributed to its prime location, higher wages, and job diversity. The study also identified several other homogeneous employment centres, such as Murska Sobota, Koper, Nova Gorica, Novo mesto and Slovenj Gradec, as well as some competitive centres, like Celje and Velenje, Kranj and Ljubljana, Jesenice and Kranj, Maribor and Ptuj, etc. (Bole & Gabrovec, 2012). Not long after, the construction of new highways and the abolition of the toll system thoroughly altered the mobility flows, as well as the regional structure of Slovenia. Bole (2011) showed that the new traffic network increased the attractiveness of all the major employment centres, not only by increasing the spatial range of individual centres but also by increasing the flow between them. Interestingly, the study also found an increase in spatial range even in areas where no road connections were improved. A more worrying result of the study is the fact that Ljubljana was the only regional centre that recorded a considerable expansion in its spatial range compared to other employment centres. This again points to the disproportionate development of the eastern and western regions (Bole, 2011).

In the last couple of decades, Slovenia has witnessed significant changes in commuting dynamics as well as mode of transportation. The most notable change being the gradual decrease in the use of public transport. As Bole explains: "the number of registered cars has nearly doubled from 1985 to 2005", while "the percent of workers who use public transport for their daily commuting has decreased from over 64% in 1981 to just 10% in 2001" (Bole & Gabrovec, 2012, p. 172). This shift can be largely attributed to the lack of sufficient public transport networks as well as slow commuting times, making the choice to drive more favorable. With the development of new highway infrastructure and the uncompetitive nature of the public transport system, the gap between commuting times by car vs. by public transport has been growing. As a result, more than 78% of commuters travel to work by car every day, as opposed to only 10% who use public transport. This type of car-dominant traffic network is not only harmful to the environment, but also leads to the development of dispersed settlements with sparse population density, leading to urban sprawl (Bole & Gabrovec, 2012).

In the past decade, there have been several authors, who have tried to study and understand the commuting patterns of Slovenians, through different methodological approaches. In their 2008 study, Drobne et al. examined the influence of accessibility[7] on regional commuting flows in Slovenia. The authors used an extended inter-regional gravity model with GIS-tools that allowed introduction of accessibility into the model. The study showed that the average gross earnings per person from a certain region directly correlate to the accessibility of motorway connections in the region (Drobne, Bogataj & Lisec, 2008). A similar approach was used by Drobne et al. (2011, 2012) in their analysis of the municipalities' stickiness and attractiveness levels for commuters in Slovenia. They examined the commuting data for the period 2000-2009 and applied an extended gravity model to measure the stickiness and attractiveness of the municipalities, based on parameters, such as travel time by car, employment rate and average gross earnings per municipality. The authors concluded that both the stickiness of the population in the municipality of origin and attractiveness of the population in the municipality of destination were increasing. Also increasing was the influence of distance on commuting flows, implying that the accessibility of (labor-intensive) municipalities should increase in order to maintain the current commuting flows (Drobne, Bogataj, Zupan & Lisec, 2011; Drobne, Bogataj & Lisec, 2012).

Another study by Drobne et al. (2013) looked at the dynamics of migration and how it relates to commuting. The study aimed to measure the effect of the attractiveness of individual employment centers, municipalities of origin and the effect of the distance between them. Using a standard spatial interaction model, the authors analyzed the data divided into two time periods - before the economic crisis (2000-2007) and during the crisis (2008-2011). The study showed that before the economic crisis, migration and commuting flows were positively correlated. After the crisis, that dynamic changed. The authors argue, that after the year 2008 the migration and commuting flows became negatively correlated, though they were unable to confirm this hypothesis due to insufficient data at the time. The study also found that the number of migrations and commutes increased prior to the economic crisis, and dropped abruptly, after the financial crash (Drobne, Rajar & Lisec, 2013). Finally, in their 2014 study, Drobne et al. studied the effect of distance on the intensity of commuting flows. The authors analyzed several distance-decay functions using commuting data for Slovenia in the period 2010-2011. Their results show that a normalized power-exponential function presents the best fit to the data, applying both to commuting and migration flows (Drobne & Lakner, 2014).

In 2020, commuting and migration flows were critically disrupted by the outbreak of the SARS-CoV-2 pandemic, that started at the beginning of February. Since then, countries all around the world have taken extreme precautionary measures in order to contain the spread of the virus (Shibayama, Sandholzer, Laa & Brezina, 2021). The closure of schools, universities and workplaces, the restriction of commercial activities, as well as the complete shut-down of the public transport system, were only some of the measures taken

[7] Accessibility in Drobne et al. (2008) is defined in two parts. Firstly, as distance (in time) to different regional centers such as the capital Ljubljana, 12 regional and administrative centers as well as municipal centers. Secondly, as distance (in time) to motorway connections.

by the Slovenian government at the time. Due to the strict nature of these restrictions, the size of commuting flows decreased drastically in 2020 and 2021, as people transitioned to either not working (e.g. furloughed workers) or working remotely from home (Bartolj, Murovec & Polanec, 2022). Moreover, due to the complete closure of the public transport system, the car became the dominant mode of transport for people still commuting to and from work, as well as for shopping and other daily activities (Brezina, Tiran, Ogrin & Laa, 2021). This, in and of itself, poses a methodological issue since the shift to work-from-home cannot be properly measured due to the methodological framework in which the commuting data is collected. In the next chapter, we move on to the empirical part of this thesis.

2 METHODOLOGY

As we have shown in the previous chapter, Slovenia is uniquely interesting in terms of its regional development. It is a small country, nestled in the central part of Europe, not only geographically but also economically centered at the crossroads between the eastern and western part of the continent. The aforementioned studies characterize Slovenia as a sub-urbanized country, with relatively small towns and few large urban settlements, a distinct developmental and labor gap dividing the eastern and western part of the region. This poses an interesting question as to how these distinct urban features affect the geographical structure of its labor market, how they shape the formation of employment centers and the commuting flows that connect them.

The purpose of this thesis is to examine the commuting patterns of workers in Slovenia and define the geographical structure of Slovenia's labor market. Firstly, we will discuss the characteristics of commuting flows in Slovenia based on data at the level of county of residence and county of employment for all employed persons in Slovenia, generally, by gender and education. Secondly, we will examine the structure of Slovenia's labor market by generating flow-based clusters dependent on the identification of regional employment centers, using a clustering algorithm applied to data on bilateral commuting flows. We perform this analysis by imposing several restrictions on the model and applying it to data on general commuting flows as well as commuting flows by education level. Lastly, we will discuss our estimates of distance elasticities for the labor market in Slovenia, using the Poisson Pseudo-Maximum-Likelihood estimator. It is our primary goal to assess the effect of commuting times on the willingness of workers to commute and we expand our analysis to include the effects of education, gender, and wages on commuting patterns in Slovenia.

We aim to provide an estimate of distance elasticities for the labor market in Slovenia and measure how these elasticities change in time, by gender as well as by level of education. We also hope to identify the possible effect of wages on commuting flows as well as provide a general analysis of the geographical structure of the labor market in Slovenia. The results of this thesis could have possible implications for labor policy, specifically regarding job referral strategies the Employment Service of Slovenia uses when matching job opportunities to potential candidates. We believe our empirical analysis could shed light on the commuting patterns of Slovenian workers and show just how far employees are willing to commute for work, based on different factors such as education, gender etc.

2.1 Estimation Method

As the literature suggests there are several ways to approach studying human mobility. From continuous-time random walk models to spatial interaction models, the research offers several feasible estimation methods for measuring commuting flows. For the purpose of this thesis, we will focus on the gravity model for commuting flows, developed by Ahrens and Lyons in their 2021 article.

2.1.1. The Augmented Gravity Model

For the basis of their model, Ahrens and Lyons (2021) employ the cross-section commuting model proposed by Ahlfeldt and Wendland (2016). The basic equation is:

$$\pi_{ij} = \tau t_{ij} + o_i + d_j + u_{ij} \tag{2}$$

where the dependent variable π_{ij} is equal to the logarithm of the probability of commuting from residence *i* to place of work *j*, where C_{ij} is the number of commuters travelling from *i* to *j* and P_i is the number of residents.

$$\pi_{ij} = \ln\left(\frac{C_{ij}}{P_i}\right) \tag{3}$$

Variable t_{ij} measures either the geographic distance or travel time between *i* and *j* and parameters o_i and d_j are defined as the push and pull factors, which measure the attractiveness of place of residence (*i*) and place of work (*j*), and u_{ij} captures the error term. This approach is able to measure the spatial decay of commuting probabilities quite well, although it suffers from a few drawbacks. Firstly, the fixed effects o_i and d_j capture only the common push and pull factors and do not account for the fact that place of work *j* has a different attractiveness factor for different places of residence *i*. Secondly, the model is not able to differentiate the effects of different push and pull factors (e.g., rental prices) from other local characteristics. The authors address both weaknesses using a two-period panel framework.

Ahrens and Lyons (2021) consider a linear first-difference gravity model, as an extension to the basic specification given in equation (3). The augmented equation is then:

$$\pi_{ij,t} = f\left(r_{i,t}, r_{j,t}\right) + \boldsymbol{x'}_{\boldsymbol{i,t}}\boldsymbol{\theta} + \boldsymbol{x'}_{\boldsymbol{j,t}}\boldsymbol{\delta} + \mu_{ij} + \varepsilon_{ij,t} \tag{4}$$

where the dependent variable $\pi_{ij,t}$ now has an additional time index *t* and the variables $\boldsymbol{x'}_{\boldsymbol{i,t}}$ and $\boldsymbol{x'}_{\boldsymbol{j,t}}$ are now vectors describing the time-varying properties of locations *i* and *j*. The vectors contain information on the number of jobs and residents, and other socio-economic characteristics (e.g., information on properties, demographic factors, etc.). The fixed effects μ_{ij} measure the time-invariant attractiveness of commuting from location *i* to location *j*. The variable μ_{ij} depends on distance and other factors which are more difficult to measure and are not as easily observed. Finally, $\varepsilon_{ij,t}$ denotes the error terms. In order to eliminate these unknown factors, the authors compute the first difference of the model:

$$\Delta\pi_{ij,t} = \Delta f\left(r_{i,t}, r_{j,t}\right) + \Delta\boldsymbol{x'}_{\boldsymbol{i,t}}\boldsymbol{\theta} + \Delta\boldsymbol{x'}_{\boldsymbol{j,t}}\boldsymbol{\delta} + \Delta\varepsilon_{ij,t} \tag{5}$$

The main crux of the model lies within variables r_i and r_j, which represent the (logarithm of) rent prices at place of residence *i* and place of work *j*. The authors specify two options concerning function $f\left(r_{i,t}, r_{j,t}\right)$:

$$f\left(r_{it}, r_{jt}\right) = \alpha(r_{jt} - r_{it}) \tag{6}$$

$$f\left(r_{it}, r_{jt}\right) = \beta_1 r_{it} + \beta_2 r_{jt} \tag{7}$$

Equation (6) represents the difference between r_{jt} and r_{it}, where the goal is to estimate the effect α. Alternatively, equation (7) includes both rents in an additive way, the goal of which is to find estimates for β_1 and β_2. The reasoning behind the consideration for the rental difference is that the location decision, or in our case the decision to commute, might be determined, not by the absolute values of rents at the location of residence and location of work, but in fact by the relative prices between them. A reasonable hypothesis then might be, that an increase in the rental price difference between locations *i* and *j* may lead to higher commuting flows. On the other hand, the additive model differentiates the push and pull factors and provides insight into whether location (or commuting) decisions are primarily driven by push factors (e.g., higher rent prices in the center) or by pull factors (e.g., lower rent prices in the periphery). The additive model is expected to produce $\beta_1 < 0$ and $\beta_2 > 0$ (Ahrens & Lyons, 2021).

2.1.2. Ordinary Least Squares (OLS)

To estimate this model empirically, one might consider the Ordinary least squares method. However, several studies[8] have pointed out, that applying this estimation method to log-linearized equations has significant drawbacks, which must be considered. The crux of the issue relates to Jensen's inequality, a consequence of which is when exposed to heteroscedasticity, the OLS estimator produces biased results. As Santos-Silva and Tenreyro (2006) explain, the concept of Jensen's inequality has been known for a while,

[8] Santos-Silva & Tenreyro, 2006; Yotov, 2012; Yotov, Piermartini, Monteiro & Larch, 2017.

however, it has been overlooked in many econometric applications, including in the use of log-linearized equations.

The inequality states that the expected value of the natural logarithm of a (random) variable is not equal to the natural logarithm of its expected value. Or in other words:

$$E(\ln y) \neq E(y) \tag{8}$$

An important implication of this inequality is, as Santos-Silva and Tenreyro explain, that “the standard practice of interpreting the parameters of log-linearized models estimated by ordinary least squares (OLS) as elasticities can be highly misleading in the presence of heteroskedasticity” (Santos-Silva & Tenreyro, 2006, p. 641).

To illustrate the issue, let’s look at an example using the gravity equation for trade, as defined by Tinbergen (1962):

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \tag{9}$$

The gravity equation states that the trade flow T_{ij} from country i to country j is proportional to the product of the two GDPs of the countries denoted by Y_i and Y_j and inversely proportional to the distance between them (D_{ij}). The analogy fails in practice since there is no set of parameters for which the gravity equation would hold exactly for all given observations. To account for these empirical inconsistencies, we apply the stochastic version of the gravity equation:

$$T_{ij} = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3}\ \eta_{ij} \tag{10}$$

where η_{ij} is the error factor and $\mathrm{E}(\eta_{ij} | Y_i, Y_j, D_{ij}) = 1$. This leads to the assumption that η_{ij} is statistically independent of our regressors, which implies:

$$E(T_{ij} | Y_i, Y_j, D_{ij}) = \alpha_0 Y_i^{\alpha_1} Y_j^{\alpha_2} D_{ij}^{\alpha_3} \tag{11}$$

Finally, the log-linearized equation is:

$$\ln T_{ij} = \ln \alpha_0 + \alpha_1 \ln Y_i + \alpha_2 \ln Y_j + \alpha_3 \ln D_{ij} + \ln \eta_{ij} \tag{12}$$

The validity of estimating this function with the OLS estimator depends fundamentally on the assumption that η_{ij} (as well as $\ln \eta_{ij}$) are statistically independent of the regressors. We can see that the expected value of the logarithm of the random variable depends on the mean as well as the higher-order moments of the distribution, such as variance. Santos-Silva and Tenreyro (2006) explain that if the variance of η_{ij} in equation (10) depends on regressors Y_i, Y_j or D_{ij}, then so will the expected value of $\ln \eta_{ij}$, thus violating the condition for consistency. When testing the OLS model against other estimators, the authors find overwhelming evidence that even when controlling for fixed effects, the OLS estimator produces considerably biased estimates when exposed to heteroscedasticity (Santos-Silva & Tenreyro, 2006).

Another issue with the application of OLS is the presence of zeros. If we look at the Newtonian theory of gravity more closely, we can see that the gravitational force can theoretically be very small, even approach zero, but can never reach it. In trade this is not the case. There are several instances where there is no recorded trade between countries. For instance, countries separated by large geographical distances or countries with high variable trade costs are less likely to trade, or not at all (ibid.). As a result, establishing a model that includes observations with zero flows is of paramount importance.

Such observations pose no issue for the gravity model in its multiplicative form but can create problems when using the log-linearized version. In these instances, the standard procedure of OLS is to eliminate the observations for which $T_{ij} = 0$. The issue here is twofold. Firstly, as we have shown, countries do not trade for several observable reasons. All of which contain important information which is pertinent to producing accurate estimation results. Secondly, deleting zero-value observations could eliminate a considerable amount of observations from the dataset and could potentially lead to inconsistent and biased estimation results (Santos-Silva & Tenreyro, 2006). This point is especially relevant in our empirical analysis since the dataset used in our estimation has less than 34% of observations which are different from zero[9].

As we have shown, there is overwhelming evidence that supports the fact that when using OLS to estimate the log-linearized forms of the gravity equation, the error terms produced are heteroscedastic and thus violate the assumption that the term $\ln \eta_{ij}$ is statistically independent of the regressors. This leads to the conclusion that OLS produces inconsistent estimates of the elasticities in question and leads to biased results (ibid.). These findings are further supported by Yotov (2012) and Yotov et al. (2017), further cementing the fact that OLS is not the appropriate estimation method to be used in our analysis.

2.1.3. Poisson Pseudo-Maximum-Likelihood (PPML) Estimator

Instead, Santos-Silva and Tenreyro (2006) suggest the application of the Poisson pseudo-maximum likelihood estimator, which can be expressed as:

$$\sum_{i=1}^{n}[y_i - \exp(x_i\tilde{\beta})]x_i = 0 \tag{13}$$

[9] The data on commuting flows containing the information on the county of residence and county of employment for all persons in employment (excluding farmers) in Slovenia for the years 2000 – 2020 contains 253.956 zero value observations out of a total of 380.556 observations. This equals 67% of all observations.

where y_i and x_i denote the dependent and explanatory variables, respectively, and β is a set of estimation parameters. The PPML estimator is considered to be a better fit, since it gives the same weight to all observations, instead of emphasizing those for which $\exp(x_i\beta)$ is large. This occurs under the assumption that $E[y_i|x]$ is proportional to $V[y_i|x]$, or $E[y_i|x] \propto V[y_i|x]$. If this holds, all observations have the same information on the parameters as the additional information on the curvature of the conditional mean (especially resulting from observations with large $\exp(x_i\beta)$), is offset by their larger variance. Additionally, the PPML estimator is structured in a way that the only necessary condition for the estimates to be consistent is the specification of the conditional mean, which is $E[y_i|x] = \exp(x_i\beta)$. Therefore, the data need not be Poisson at all for the estimator to be consistent and y_i does not have to be an integer as required for Poisson model. In their empirical analysis, Santos-Silva and Tenreyro find, that the PPML estimator is robust to different patterns of heteroscedasticity and provides a natural way of dealing with zero values in the data. As such the PPML estimator provides a reasonable alternative to the OLS estimator and is the one we chose to apply in our empirical analysis (Santos-Silva & Tenreyro, 2006).

3 DATA

For the purpose of this thesis, all the data used was obtained from two sources: (i) the Statistical Office of the Republic of Slovenia and (ii) PNZ svetovanje projektiranje d.o.o. From SURS we obtained data on commuting flows, both aggregate and disaggregated by gender and level of education, as well as gross and net average wages. From PNZ d.o.o. we obtained information on commuting times.[10]

The publicly available data on commuting flows from SURS, attainable at the Si.Stat web page, contains the information on the county of residence and the county of employment for all employed persons (excluding farmers) in Slovenia for the years 2000 – 2020. The data were obtained separately for males and females, as well as all persons in employment. The structure of the data has changed throughout the years of collection, with 20 new counties added between the year 2002 – 2010 and several changes to the structure of the counties registered in that time frame. In order to eliminate the effects of changes in geographical boundaries of municipalities on measured commuting flows, we created time-invariant county definitions mainly coinciding with boundaries before creation of new counties[11.]

[10] We would like to thank the Institute for Economic Research, who acquired the PNZ data on commuting flows for development of the Microsimulation model on labor supply choices, for allowing us to use these data for the purpose of this thesis.

[11] The registered changes to the number and structure of counties implied that county codes were time varying, which would prevent us from properly measuring the change in commuting flows through time. To mitigate this issue, we constructed time-invariant codes, which merged the time-variant counties into distinct

The data on commuting flows by education level is comprised of data at the level of the county of residence and the county of employment for all persons in employment (excluding farmers), based on their level of education for the years 2010-2020.[12] The data is categorized by klasius (1,2,3), klasius 1 representing all workers with elementary school education or less, klasius 2 representing all workers with high school education and klasius 3 representing all workers with higher education. Since the data spans only 11 years, most of the structural changes noted in the first dataset were omitted, with some exceptions. These inconsistencies were mitigated using time-invariant counties. Due to a small number of commuters in certain counties, some of the observations in the dataset were statistically protected due to privacy concerns. These observations were given a value of zero.

The data on commuting times contains information on the commuting time (in minutes) from the county of residence to the county of work and back for the years 1997, 2007 and 2017. The dataset also recognizes distinct subcategories within counties, which imply different commuting times depending on the location of residence and location of work within specific counties. Based on the benchmark data, we were able to generate the approximate commuting time trends for the years 1997-2020, that were used in the empirical analysis. There were no inconsistencies in the data regarding the changes in the number of counties, however, the data was converted to a time-invariant coding system for the purpose of consistency using unweighted averages of commuting times.

The publicly available data on wages is comprised of both gross and net average monthly wages per county for the years 2005-2020. Due to similar methodological issues, time-invariant counties were used.

---

units in order to contain the changes in question into single entities, thus eliminating any possible biases. Another methodological issue we faced was related to information gathering. Since the data are created by relating the registered domicile to the registered place of work, the method is not able to capture information on persons not registered in their domicile or their place of work, rendering the data incomplete. Similarly, the method is also unable to provide an accurate estimation of the number of persons working from home, which is especially relevant for the year 2020 due to the implementation of Covid-19 restrictions.

[12] We are grateful to Nuška Brnot from SURS for kindly providing us with data for commuting flows disaggregated by the level of education.

4 DESCRIPTIVE STATISTICS

4.1 Commuting Flows

Commuting flows were analysed using data on the county of residence and county of employment for all persons in employment (excluding farmers) in Slovenia for the years 2000-2020. In Table 1 we can see the distributional moments for the share of people working outside the county of their residence. It is evident that the share of people working outside the county of residence has been increasing in the last 20 years. The weighted share in the year 2000 amounted to 39.9% of the population and increased to 53.1% in the year 2020, which is more than half of the working population. From 2000-2010 the share of people working outside the county of residence increased by almost 22%, whereas from 2010-2020 the share increased at a steadier pace, by 9%. The weighted average was calculated using the share of working persons within a county relative to all working persons in Slovenia.

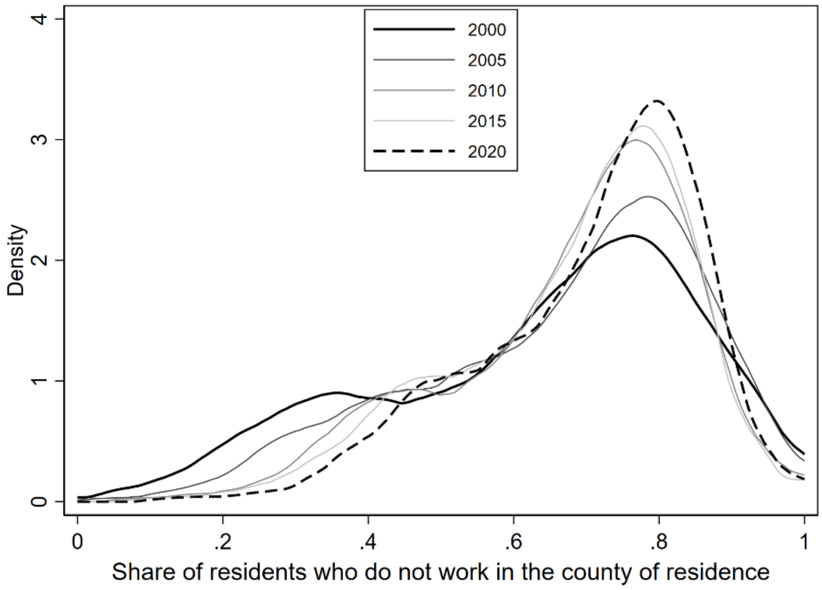
Table 1: Share of persons working outside county of residence in Slovenia, 2000-2020

	Mean			
Year	Weighted	Unweighted	Median	Std. Dev.
2000	0.399	0.622	0.609	0.213
2001	0.408	0.666	0.615	0.209
2002	0.417	0.667	0.620	0.205
2003	0.428	0.672	0.627	0.200
2004	0.439	0.677	0.636	0.196
2005	0.450	0.693	0.645	0.192
2006	0.459	0.699	0.652	0.188
2007	0.468	0.707	0.658	0.183
2008	0.474	0.714	0.661	0.178
2009	0.478	0.707	0.657	0.172
2010	0.486	0.704	0.662	0.168
2011	0.492	0.707	0.666	0.167
2012	0.496	0.710	0.666	0.165
2013	0.497	0.709	0.665	0.164
2014	0.500	0.705	0.655	0.162
2015	0.503	0.711	0.669	0.162
2016	0.507	0.707	0.671	0.160
2017	0.512	0.716	0.677	0.159
2018	0.517	0.713	0.680	0.158
2019	0.520	0.724	0.683	0.159
2020	0.531	0.733	0.692	0.153

Source: Own work

Figure 1 and Figure 2 illustrate the unweighted and weighted kernel density functions, representing the share of persons working outside their county of residence. The kernel density estimation represents a non-parametric way of estimating the probability density function and helps us visualise the distribution of the share of residents working outside their county of residence. As we can see from the figures below, there has been a steady increase in the share of people commuting to work outside their county of residence. This trend is also evident in the weighted kernel density estimation.

Figure 1: Kernel density distribution for the unweighted share of persons working outside county of residence in Slovenia, 2000-2020



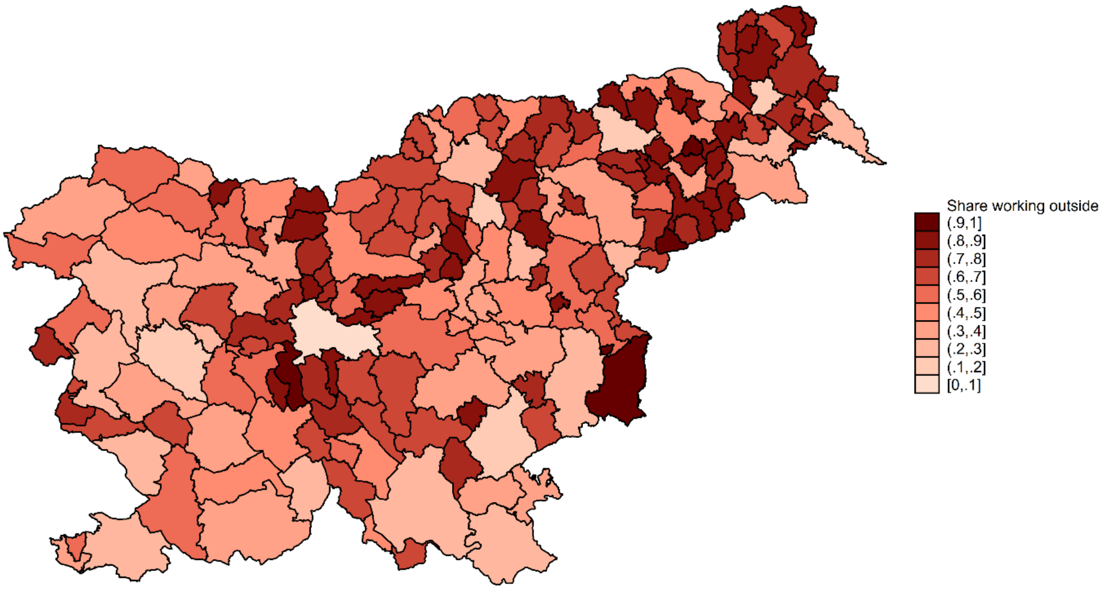
Source: Own work

Figure 2: Kernel density distribution for the weighted share of persons working outside county of residence in Slovenia, 2000-2020

Source: Own work

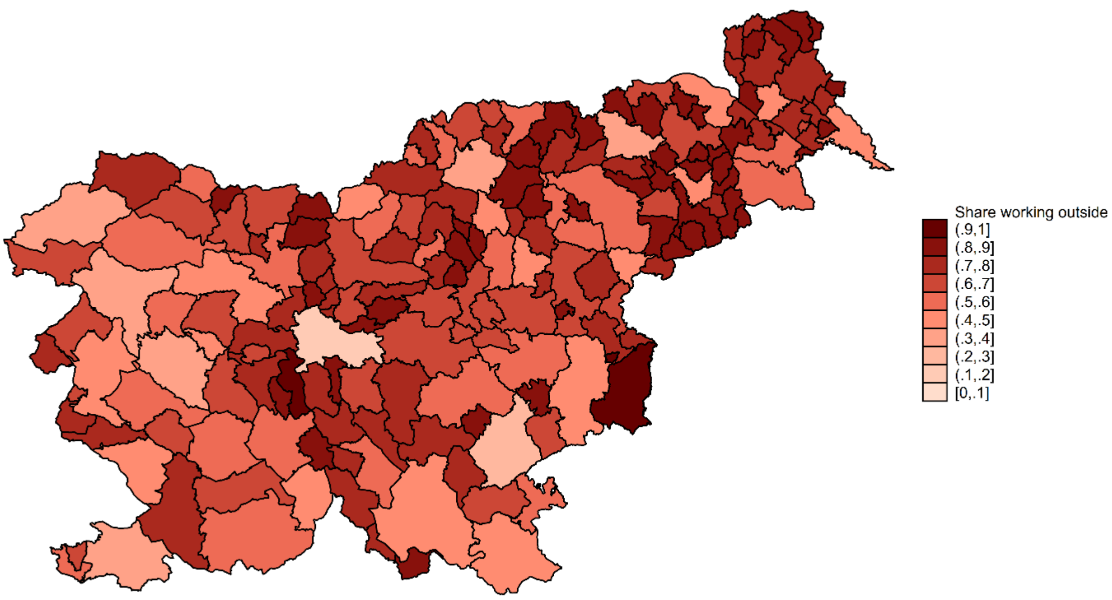
We can see a similar trend in the next two figures. Figures 3 and 4 represent the share of persons working outside their county of residence in the years 2000 and 2020. This share has increased throughout the whole country in the past 20 years, but most noticeably in the Osrednjeslovenska region. It is evident that, today, more people are commuting outiside their county of residence to the main employment centers, such as Ljubljana, Maribor, Murska Sobota, Velenje and Zreče, and are willing to travel longer distances than in 2000.

Figure 3: Share of persons working outside county of residence in Slovenia, 2000



Source: Own work

Figure 4: Share of persons working outside county of residence in Slovenia, 2020



Source: Own work

4.1.1. Cluster Analysis

Using a flow-based clustering algorithm by Meekes & Hassink (2018), we were able to identify employment center clusters, based on relational data of flows. The data used for generating employment clusters were bilateral commuting flows for the years 2000-2020. Due to several methodological changes in the data occurring through the years 2002-2010, we shifted our focus to the changes in the 2010-2020 time frame. The choice of stopping criteria used was 12 clusters, which directly correspond to the 12 statistical regions[13] in Slovenia. The purpose of generating employment clusters was to see whether the flow-based clusters would coincide with the corresponding statistical regions.

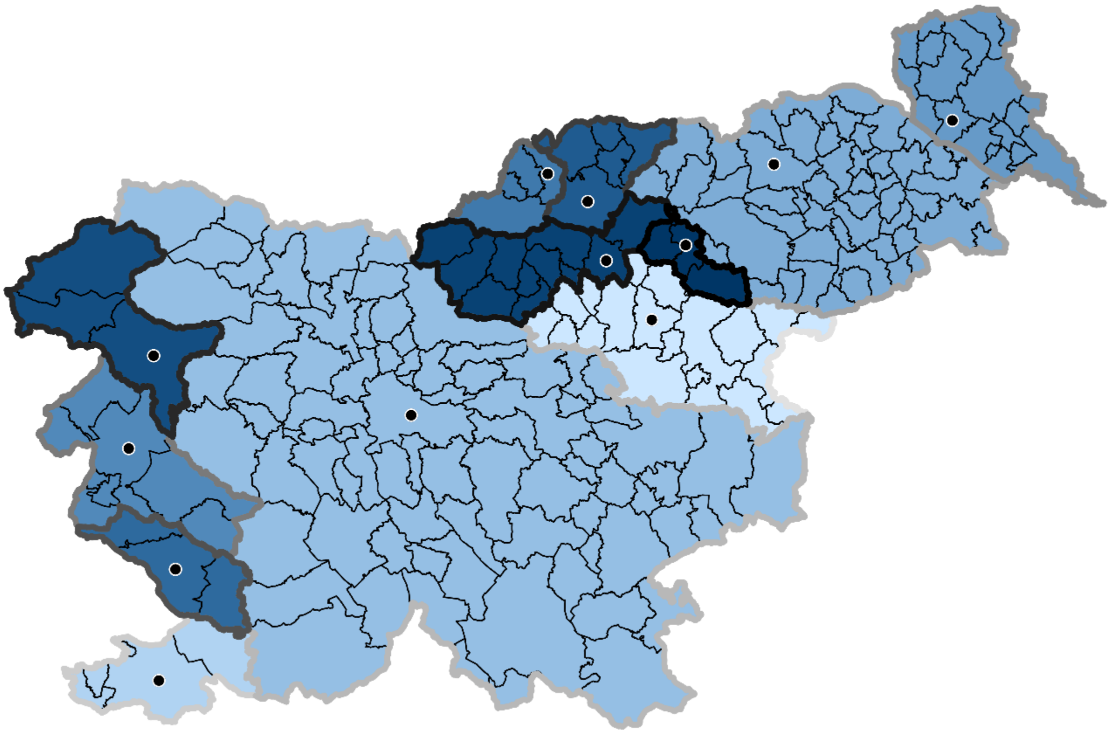
Figure 5 shows the 12 identified employment clusters, based on 2010 data. The 12 biggest employment centers identified in 2010 were Celje, Koper, Ljubljana, Maribor, Murska Sobota, Nova Gorica, Ravne na Koroškem, Sežana, Slovenj Gradec, Tolmin, Velenje and Zreče. The point markers represent the individual employment centers, while the area surrounding them represents the geographical reach of each corresponding employment center. As we can see Ljubljana, Maribor and Murska Sobota represent dominant employment centers in their regions, with Maribor and Murska Sobota coinciding with their respective statistical regions. The Koroška and Savinjska Region appear to be more fragmented, with Savinjska having three strong employment centers (Celje, Velenje and Zreče) and Koroška having two (Ravne na Koroškem and Slovenj Gradec). Also fragmented are Goriška and Obalno-kraška regions, with the former having two employment centers (Sežana and Tolmin) and the latter also two (Koper and Nova Gorica). The biggest employment center in the region is Ljubljana, with its geographical reach spreading across six statistical regions.

Figure 6 represents the 12 identified employment clusters based on 2020 data. The 12 biggest employment centers identified in 2020 were Bovec, Celje, Ljubljana, Maribor, Murska Sobota, Nazarje, Nova Gorica, Slovenj Gradec, Slovenjske Konjice, Tolmin, Velenje and Zreče. The most notable change since 2010 has been the geographical reach of Ljubljana, which has expanded to include the entire statistical region of Obalno-kraška, meaning that the employment reach of Ljubljana expands over more than 50% of the entire geographical area of Slovenia. We can also see the formation of several new employment centers predominantly in the Goriška, Koroška and Savinjska regions. In the Goriška region, we see the formation of a new employment cluster, Bovec in the northern part of the region. In the Koroška region, Ravne na Koroškem are no longer an employment center, with the majority of flows directed towards Slovenj Gradec. The most noticeable change is in the Savinjska region, with much more fragmented employment

[13] Statistical regions in Slovenia are Gorenjska, Goriška, Jugovzhodna Slovenija, Koroška, Obalno-kraška, Osrednjeslovenska, Podravska, Pomurska, Posavska, Primorsko-notranjska, Savinjska and Zasavska region.

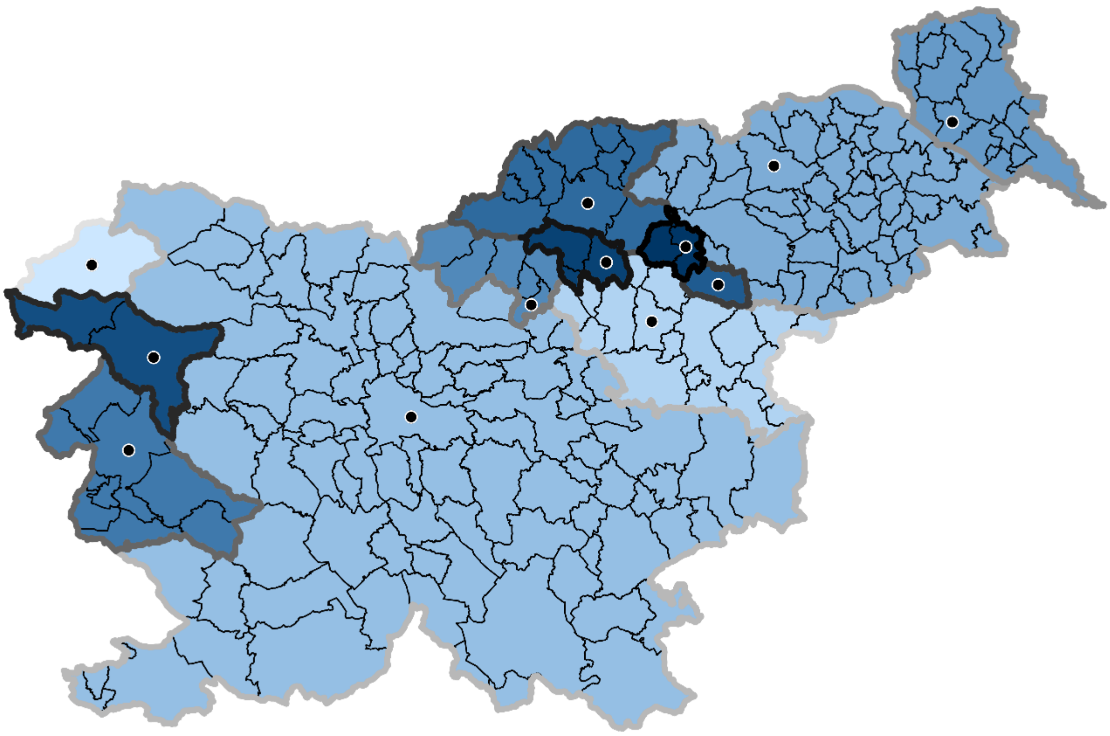
clusters than in the year 2010 and employment centers Nazarje and Slovenjske Konjice.

Figure 5: 12 Clusters based on commuting flows in Slovenia, 2010



Source: Own work

Figure 6: 12 Clusters based on commuting flows in Slovenia, 2020

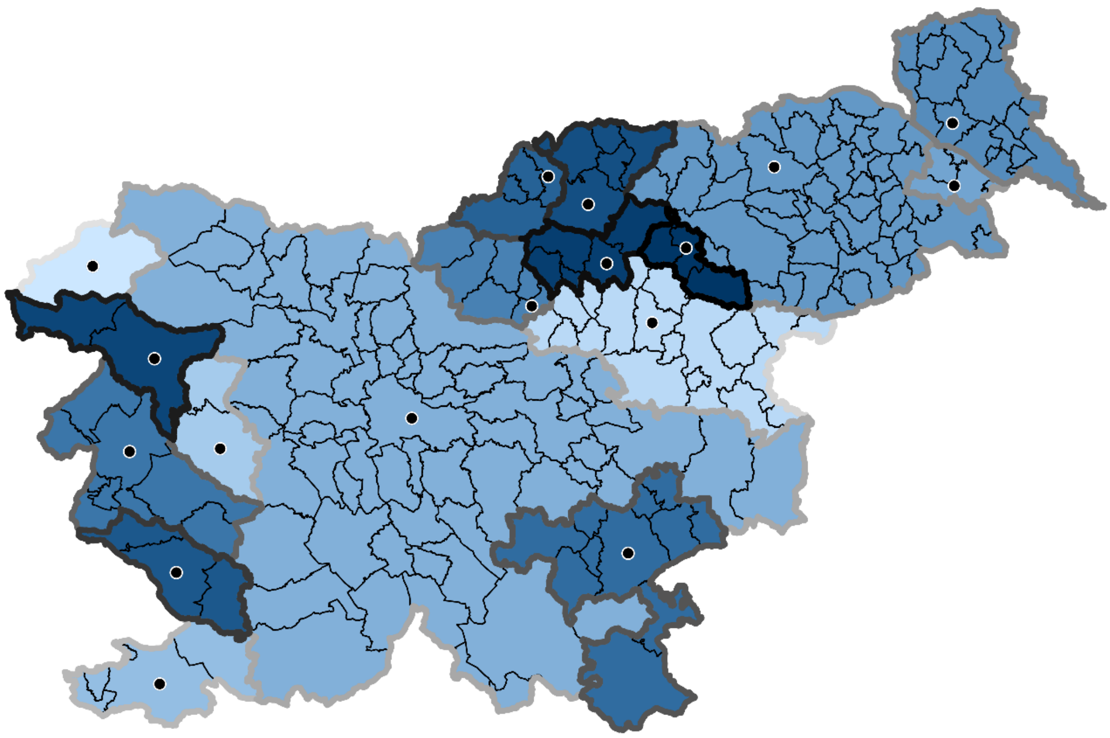


Source: Own work

The flow-based clustering algorithm allows us to impose different stopping criteria for generating clusters. One of the stopping criteria is the minimum level of internal relative flows (denoted L_m in Meekes & Hassink, 2018). The algorithm defines clusters based on the set criteria of L_m and continues to cluster counties until the stopping criteria are met.

Figure 7 shows the results of imposing a 50% minimum level of internal relative flows on data for the year 2010. As we can see the algorithm identified 17 different employment clusters[14], meaning that within each cluster at least 50% of the residents commute to work within the defined cluster. Figure 8 shows the results of the same analysis, used on data for 2020. As we can see the algorithm identified 11 different employment clusters that satisfy the minimum level of internal relative flows of 50%. The employment centers are Bovec, Celje, Ljubljana, Maribor, Murska Sobota, Nazarje, Nova Gorica, Slovenj Gradec, Tolmin, Velenje and Zreče. The results of this analysis are similar to the results in Figure 6, where we imposed a set number of clusters to be generated. This further shows, not only that the share of people working outside their county of residence has increased in the past 10 years, but also that the number of strong employment centers is diminishing, and people are commuting further to work.

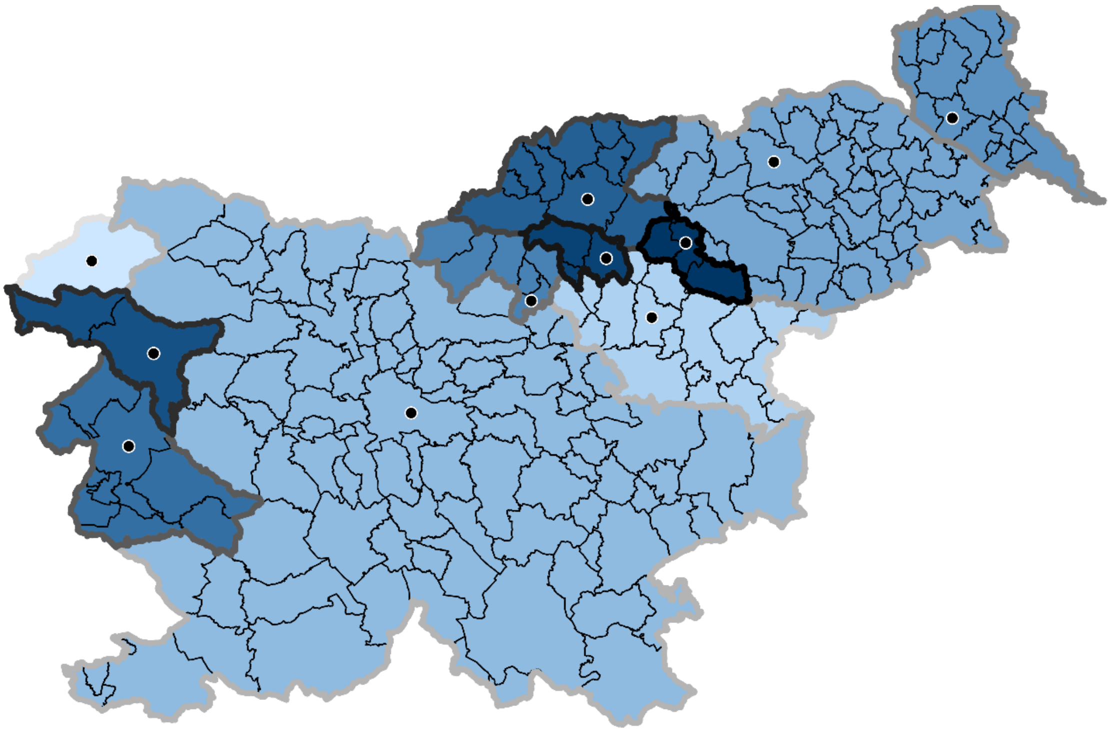
Figure 7: Clusters based on a minimum of internal relative commuting flows of 50% in Slovenia, 2010



Source: Own work

[14] Bovec, Celje, Idrija, Koper, Ljubljana, Ljutomer, Maribor, Murska Sobota, Nazarje, Nova Gorica, Novo mesto, Ravne na Koroškem, Sežana, Slovenj Gradec, Tolmin, Velenje and Zreče.

Figure 8: Clusters based on a minimum of internal relative commuting flows of 50% in Slovenia, 2020



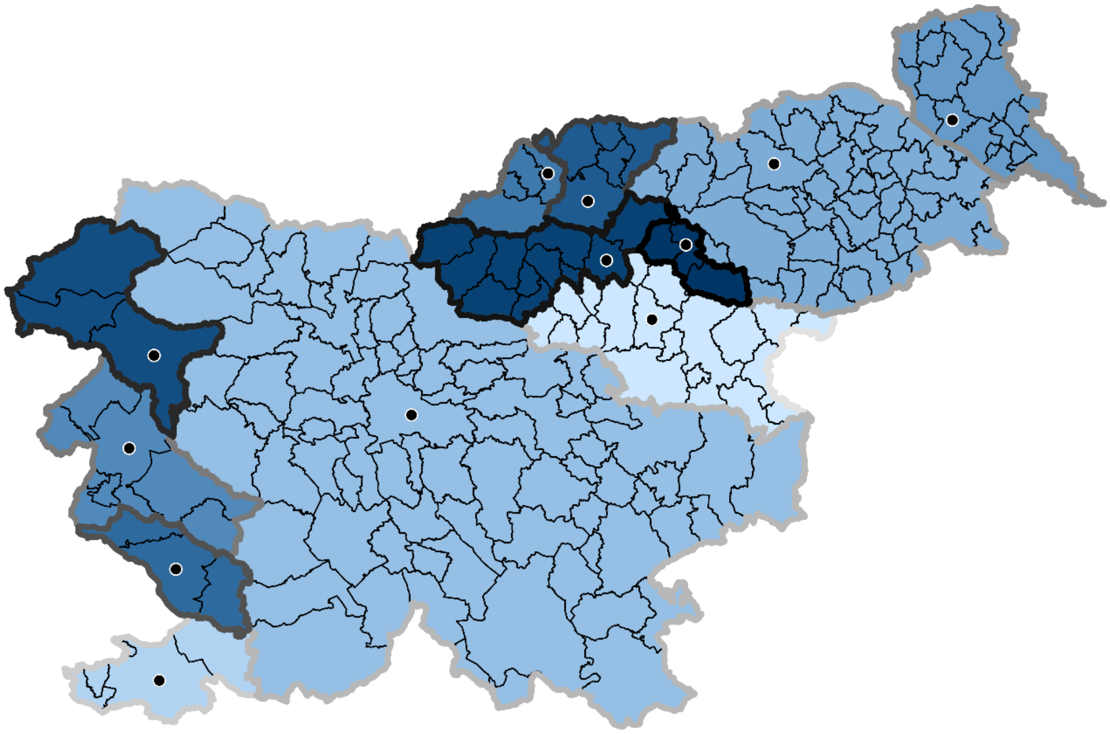
Source: Own work

Another stopping criterion the program allows is a minimum level of interaction with a source unit, defined as Q in Meekes & Hassink (2018). The algorithm defines commuting clusters based on the imposed minimum cross-county flow threshold. Figure 9 shows the results of imposing a 15% minimum external flow threshold, a minimum level of interaction at which a source unit is aggregated to a destination unit, on the data for 2010.[15] The algorithm identified 12 different employment centers that satisfied the 15% minimum external flow limit. The analysis produced the same results as in Figure 6, identifying the same 12 employment centers.

Figure 10 shows the results of imposing a 15% minimum external flow threshold on the data for 2020. Interestingly, applying the same criteria as in 2010, the algorithm identified 4 major employment clusters that satisfy the imposed threshold. The most noticeable of which is the cluster surrounding Ljubljana, which seems to dominate the employment market. There is a considerable number of counties that satisfy the 15% external flow minimum, which amounts to almost 78% of the geographical area of Slovenia. The other employment clusters are in the Koroška, Podravska and Pomurska regions, with employment centers in Slovenj Gradec, Maribor and Murska Sobota. These clusters are condensed in the eastern part of Slovenia, where the distance to the capital is the largest and commuting times would be the longest.

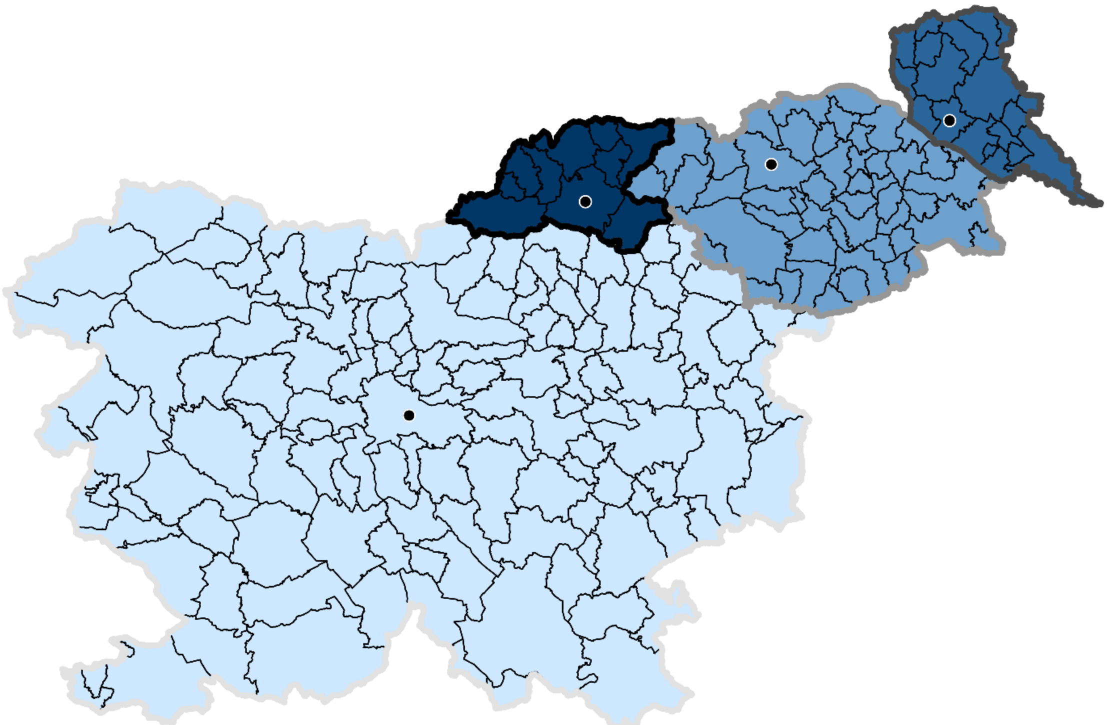
[15] The 15% minimum external flow threshold was chosen because the number of clusters generated using the data for 2010 was very similar to the number of clusters generated in Figures 5 and 6. Using a similar benchmark case with a different stopping criterion, we were able to measure how the external flows changed from 2010 to 2020.

Figure 9: Clusters based on a minimum of external commuting flows of 15% in Slovenia, 2010



Source: Own work

Figure 10: Clusters based on a minimum of external commuting flows of 15% in Slovenia, 2020



Source: Own work

4.2 Commuting Times

With the data on commuting times (in minutes) from the county of residence to the county of work (and back) for the years 1997, 2007 and 2017, we were able to generate estimated commuting time trends for the 1997-2020 time frame. The first entry in Table 2 shows us the unweighted mean time of all inter-county flows for each year. As we can see the average commuting time has been decreasing over the years, dropping from 194.81 to 179.35 between 1997 and 2007 and from 179.35 to 171.15 between 2007 and 2017 (with commuting time representing time travelling to and from work). The second entry in Table 2 is the weighted average mean, using aggregate county flows.[16] We can see that the actual commuting times differ immensely from the unweighted commuting times, with the aggregate times ranging from 35-40 minutes. These results are expected since workers prefer a shorter commute, to a longer one. Interestingly, while the unweighted mean time is decreasing through time, the aggregate mean is increasing slightly, each year. Most of the statistics suggest that commuting times have increased over the years, with a slight decrease recorded between 2006-2007 (due to the building of new road infrastructure). The improvements recorded in 2007 reduced the weighted average commuting time by 3.3. minutes, with the unweighted mean and median decreasing much less.

Table 2: Dynamics of commuting times (in minutes) within Slovenia, 1997-2020

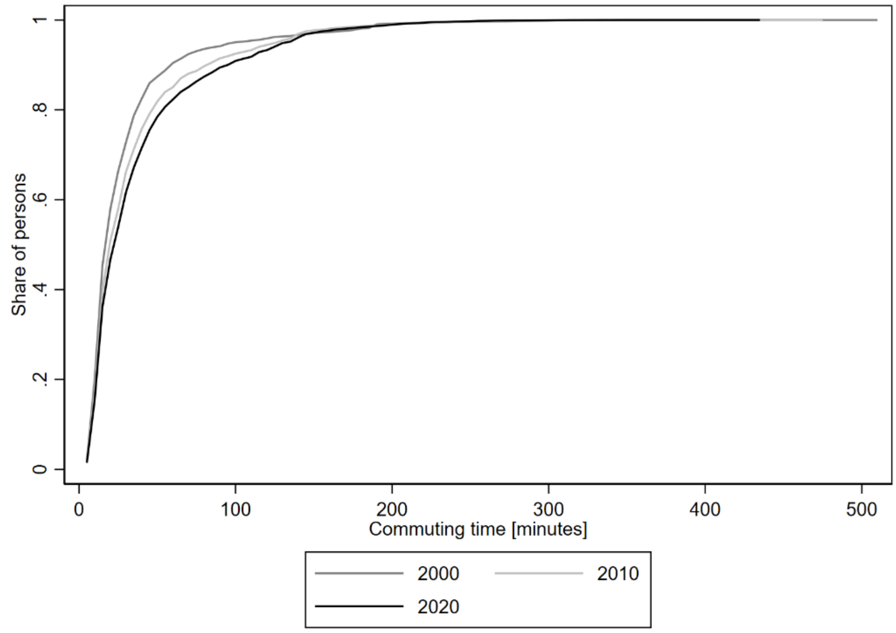
	Unweighted	Aggregate	County-Level		
Year	Mean Time	Mean	Mean	Median	Std. Dev.
1997	194.81				
1998	194.81				
1999	194.81				
2000	194.81	34.52	30.52	34.21	20.08
2001	194.81	35.14	30.73	34.90	20.30
2002	194.81	35.72	31.48	35.67	20.63
2003	194.81	36.66	32.22	37.22	20.86
2004	194.81	37.71	33.29	38.44	21.43
2005	194.81	38.73	34.44	39.54	22.17
2006	194.81	39.80	35.27	40.88	22.49
2007	179.35	36.53	34.13	39.15	21.37
2008	179.35	37.08	34.80	39.75	22.01
2009	179.35	37.03	34.54	39.62	21.62
2010	179.35	37.62	35.06	39.96	22.03
2011	179.35	38.43	35.80	40.64	22.48
2012	179.35	38.52	35.74	41.17	22.29
2013	179.35	38.66	35.92	41.14	22.41
2014	179.35	39.00	36.08	41.44	22.51
2015	179.35	39.20	36.81	42.24	23.09
2016	179.35	39.50	37.03	42.48	23.16
2017	171.15	39.11	36.33	41.66	22.50
2018	171.15	39.43	36.75	42.32	22.77
2019	171.15	39.60	37.00	42.73	22.84
2020	171.15	40.52	38.06	43.77	23.49

Source: Own work

[16] The aggregate mean was calculated as a weighted average mean using yearly flows, while the county level statistics were weighted using yearly county flows.

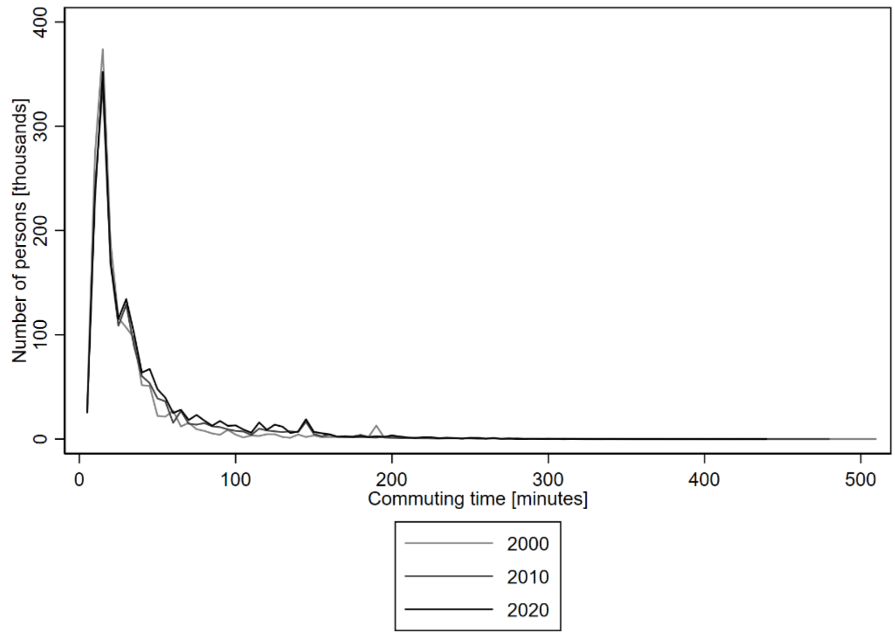
Figures 11 and 12 represent the cumulative distribution of commuting times from 2000-2020 and the commuting time frequency from 2000-2020. Workers on average spend 30-40 minutes commuting to and from work, with a small percentage reporting longer commutes. Nevertheless, commutes longer than 100 minutes are rare (likelihood is less than 10 percent). The results further show that commuting times have been increasing for the past 20 years.

Figure 11: Cumulative distribution of commuting times in Slovenia, 2000-2020



Source: Own work

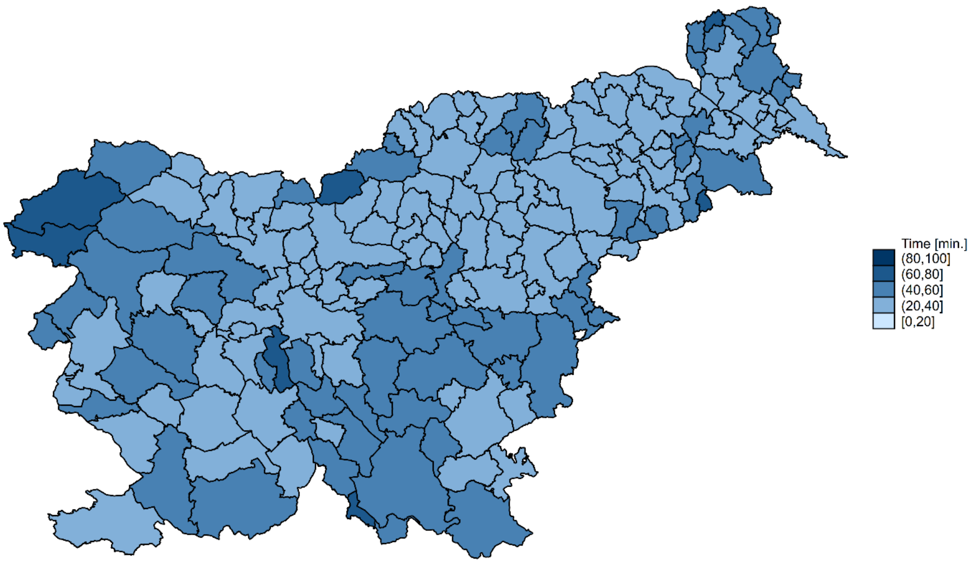
Figure 12: Commuting time frequency in Slovenia, 2000-2020



Source: Own work

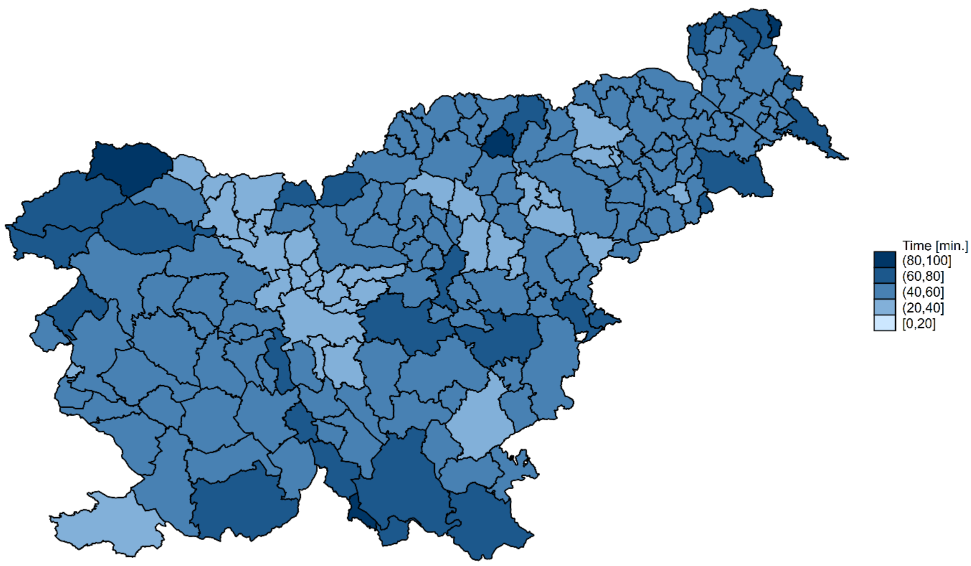
Figure 13 and Figure 14 show the county-level average commuting times in Slovenia for the years 2000 and 2020. These two figures further visualise the increase in commuting times in the past 20 years. We can see that commuting times have increased all over the country, but most notably around the main employment centers such as Ljubljana, Maribor and Murska Sobota, as well as in counties on the outskirts, meaning that people are driving to these employment centers from further away.

Figure 13: County level average commuting times (minutes) in Slovenia, 2000



Source: Own work

Figure 14: County level average commuting times (minutes) in Slovenia, 2020



Source: Own work

4.3 Education

Commuting flows by education were analysed using data on the county of residence and county of employment for all persons in employment (excluding farmers) in Slovenia based on their level of education for the years 2010-2020. The data is categorized using klasius (1,2,3), klasius 1 representing all workers with elementary school education or less, klasius 2 representing all workers with high school education and klasius 3 representing all workers with higher education.

In Table 3 we can see the weighted and unweighted share of people working outside the county of residence for persons with elementary school education or less, in the last 10 years. The weighted share in the year 2010 amounted to 41.5% of the population and increased to 45.3% in 2020, which corresponds to an increase of 9.2%. The weighted average was calculated using the share of working persons within a county, among all working persons in Slovenia.

Table 3: Share of persons working outside county of residence in Slovenia (klasius 1), 2010-2020

Klasius 1				
	Mean			
Year	Weighted	Unweighted	Median	Std. Dev.
2010	0.415	0.594	0.628	0.190
2011	0.423	0.597	0.633	0.190
2012	0.424	0.592	0.623	0.190
2013	0.424	0.594	0.617	0.192
2014	0.428	0.595	0.607	0.194
2015	0.431	0.600	0.613	0.198
2016	0.436	0.601	0.601	0.200
2017	0.443	0.609	0.618	0.199
2018	0.446	0.608	0.628	0.196
2019	0.451	0.612	0.626	0.198
2020	0.453	0.617	0.637	0.194

Source: Own work

Table 4 represents the weighted and unweighted shares of people working outside the county of residence for persons with high school education, in the last 10 years. The weighted share in the year 2010 amounted to 47.8% of the population and increased to 51.3% in 2020, which corresponds to an increase of 7.3%. Evidently, the share of people working outside their county of residence is slightly higher for high school graduates than for elementary school graduates, with both shares increasing gradually throughout the years. Furthermore, as the unweighted mean and the median change less than the weighted means, it seems that the county population shares changed as well – in favour of those commuting outside the county.

Table 4: Share of persons working outside county of residence in Slovenia (klasius 2), 2010-2020

Klasius 2				
	Mean			
Year	Weighted	Unweighted	Median	Std. Dev.
2010	0.478	0.644	0.686	0.171
2011	0.484	0.648	0.685	0.170
2012	0.487	0.646	0.680	0.167
2013	0.487	0.645	0.684	0.167
2014	0.491	0.645	0.681	0.167
2015	0.493	0.549	0.683	0.168
2016	0.497	0.651	0.683	0.167
2017	0.501	0.657	0.691	0.168
2018	0.505	0.656	0.692	0.163
2019	0.506	0.654	0.686	0.162
2020	0.513	0.663	0.695	0.161

Source: Own work

In Table 5 we can see the weighted and unweighted share of people working outside the county of residence for persons with higher education, in the last 10 years. The weighted share in the year 2010 amounted to 49.7% of the population and increased to 53.9% in 2020, which corresponds to an increase of 8.4%. The data shows that the share of people working outside their county of residence is the highest for people with higher education, meaning that highly educated people are more likely to commute outside their county of residence for work.

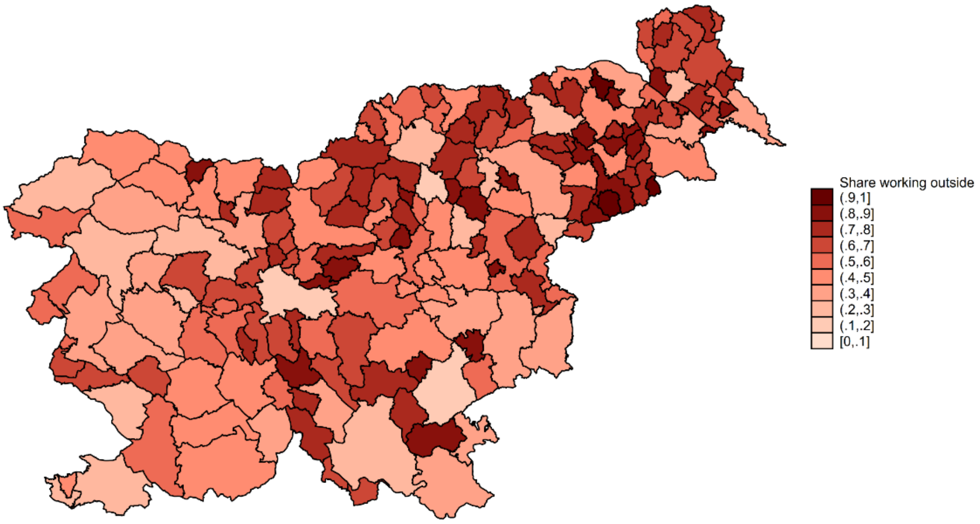
Table 5: Share of persons working outside county of residence in Slovenia (klasius 3), 2010-2020

Klasius 3				
	Mean			
Year	Weighted	Unweighted	Median	Std. Dev.
2010	0.497	0.721	0.764	0.155
2011	0.502	0.720	0.774	0.153
2012	0.505	0.719	0.765	0.153
2013	0.506	0.715	0.763	0.151
2014	0.508	0.716	0.760	0.151
2015	0.510	0.716	0.764	0.150
2016	0.514	0.719	0.766	0.148
2017	0.518	0.721	0.765	0.147
2018	0.523	0.725	0.765	0.144
2019	0.525	0.726	0.769	0.145
2020	0.539	0.738	0.778	0.141

Source: Own work

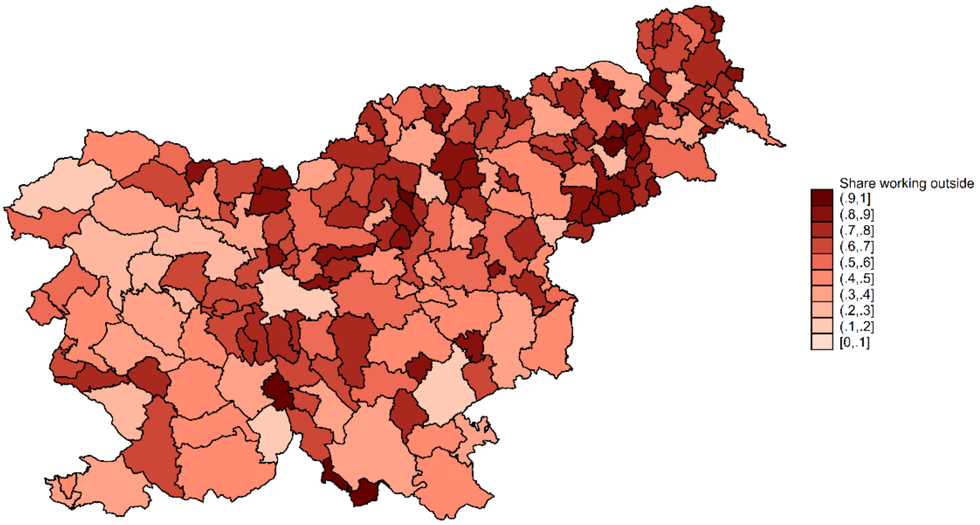
Figures 15-20 contain spatial representation of the share of people working outside their county of residence by their level of education for the years 2010 and 2020. They reveal interesting insights into the travelling patterns of workers, based on their level of education. Figures 15 and 16 represent the share of workers with an elementary school education or less, working outside their county of residence for the years 2010 and 2020. We can see that the counties with the highest share are close to major employment centers such as Ljubljana, Maribor, Murska Sobota and others in the Koroška region. We can see that the shares have increased in the year 2020, but the distribution remains the same.

Figure 15: Share of persons working outside county of residence in Slovenia (klasius 1), 2010



Source: Own work

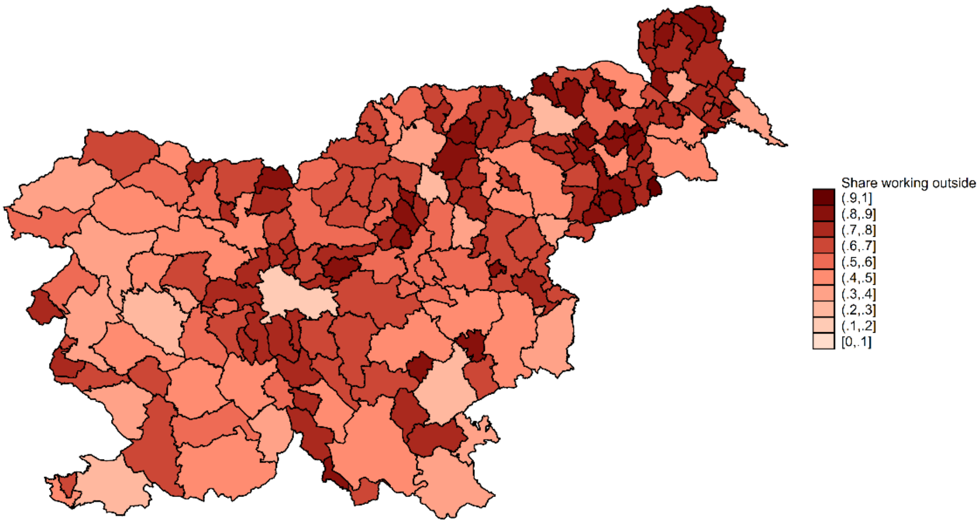
Figure 16: Share of persons working outside county of residence in Slovenia (klasius 1), 2020



Source: Own work

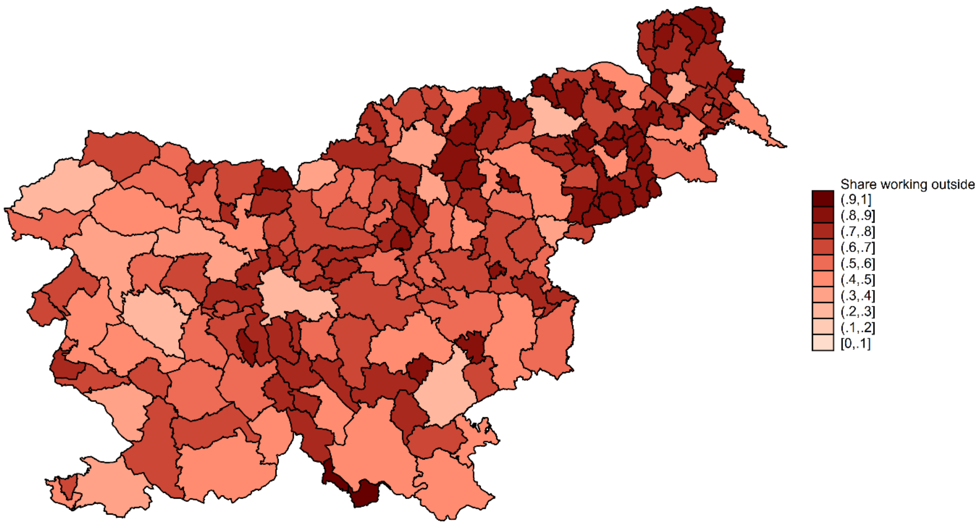
Figures 17 and 18 show the results for workers with high-school education. Obviously, the share of persons working outside their county of residence is generally higher for persons with high school education, compared to persons with elementary school education. Similarly to Figures 15-16, we can see the share of workers increase in the year 2020, especially around major employment centers (Ljubljana, Maribor, Murska Sobota, etc.). The share of people working outside their county of residence is also higher around the country's borders, meaning that people are commuting to these employment centers from farther away.

Figure 17: Share of persons working outside county of residence in Slovenia (klasius 2), 2010



Source: Own work

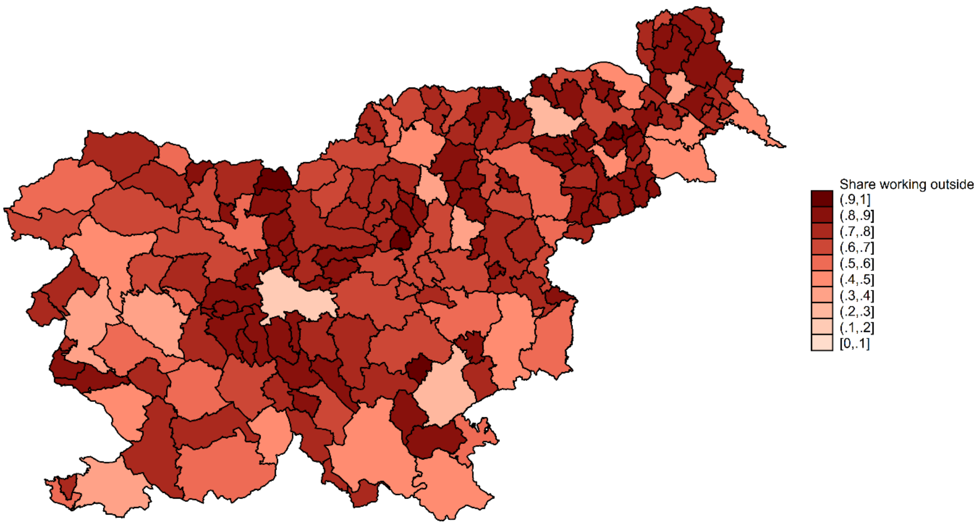
Figure 18: Share of persons working outside county of residence in Slovenia (klasius 2), 2020



Source: Own work

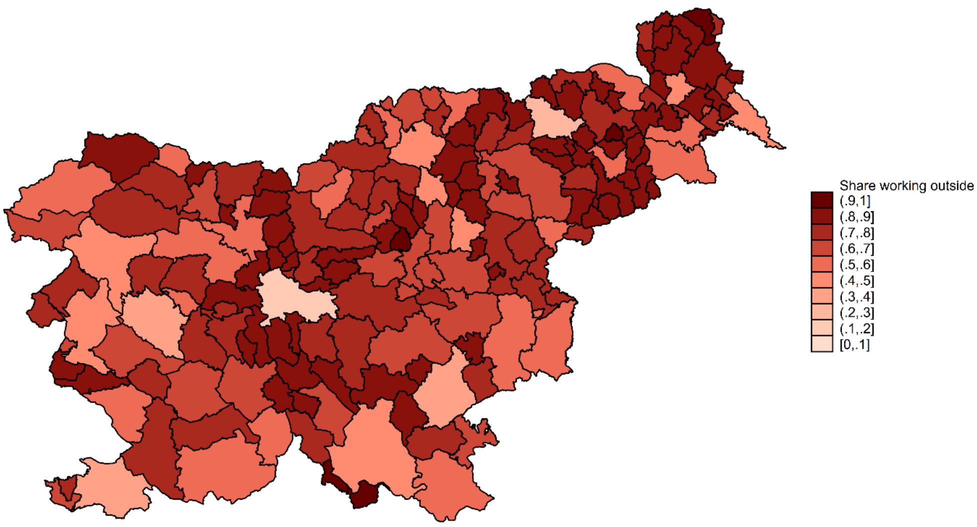
Finally, looking at Figures 19 and 20 we can see the data for workers with higher education for the years 2010 and 2020. The figures show that the share of workers in klasius 3 is significantly higher than in the previous two. Notably, the areas surrounding the major employment centers are not only darker in color, but also larger. The analysis shows us a clear disparity in commuting patterns for workers with different education levels, especially with workers who are highly educated. The data shows that educated workers are more likely to work outside their county of residence and commute longer distances to employment centers that offer the types of employment that match their education level.igh-skilled jobs are not equally dispersed throughout the country, but instead highly concentrated in larger employment centers (specifically Ljubljana), which makes commuting to work a necessity for a large share of workers with higher education.

Figure 19: Share of persons working outside county of residence in Slovenia (klasius 3), 2010



Source: Own work

Figure 20: Share of persons working outside county of residence in Slovenia (klasius 3), 2020



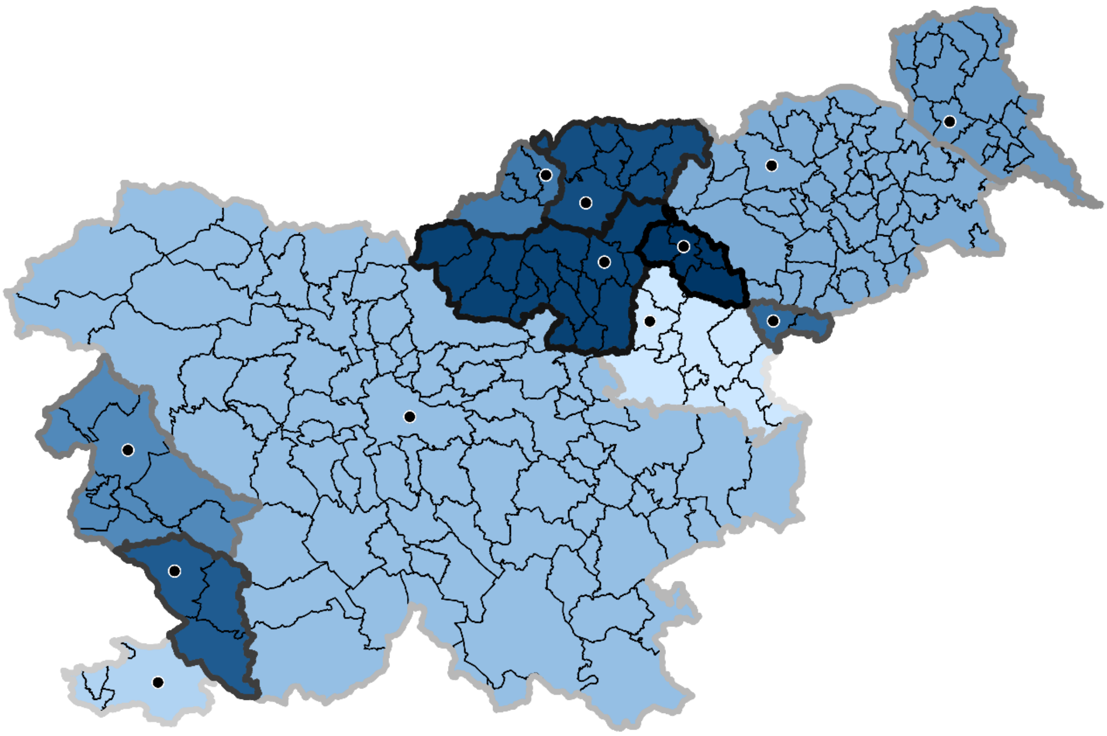
Source: Own work

4.3.1. Cluster Analysis

Since the data on commuting flows by education level is structurally identical to the data on general commuting flows, we were able to perform a cluster analysis using the same clustering algorithm. We were interested to see if there were any differences in the identified clusters for different levels of education.

Figure 21 is the result of applying the algorithm to the data for 2010, for persons with an elementary school education or less (klasius 1). The 12 employment centers identified were Celje, Koper, Ljubljana, Maribor, Murska Sobota, Nova Gorica, Ravne na Koroškem, Rogaška Slatina, Sežana, Slovenj Gradec, Velenje and Zreče. The results are quite similar to the general analysis in Figure 6, with a few notable differences. Celje, Koper, Ljubljana, Maribor and Murska Sobota represent the dominant employment centers in their corresponding statistical regions, with the highest geographical reach attributed to Ljubljana in the Osrednjeslovenska region. The Goriška region and parts of Obalno-kraška, which have, due to their geographical location and infrastructural limitations, a difficult time accessing the capital, form their employment centers in Nova Gorica and Sežana. Similarly, Koroška and Savinjska region form their own employment centers in Celje, Ravne na Koroškem, Rogaška Slatina, Slovenj Gradec, Velenje and Zreče. Most of these cities are known for being centers of heavy industry, with the steel factory in Ravne na Koroškem, Unior d.o.o. and Menerga d.o.o. in Zreče and Steklarna Rogaška d.o.o. in Rogaška Slatina, to name a few.

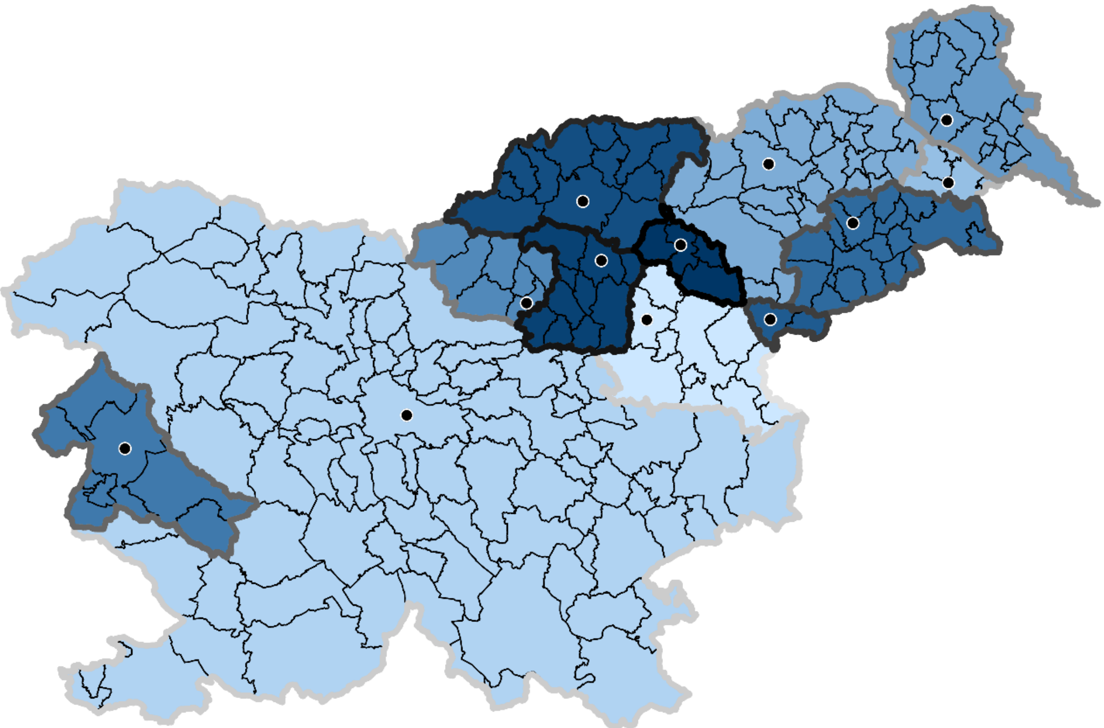
Figure 21: 12 Clusters based on commuting flows and education level in Slovenia (klasius 1), 2010



Source: Own work

Figure 22 represents the 12 clusters generated for persons with an elementary school education or less, for the year 2020. The 12 employment centers identified were Celje, Ljubljana, Ljutomer, Maribor, Murska Sobota, Nazarje, Nova Gorica, Ptuj, Rogaška Slatina, Slovenj Gradec, Velenje and Zreče. There is an interesting shift in clusters, compared to data for 2010, with several new employment centers emerging, such as Ljutomer, Nazarje and Ptuj. Moreover, the area surrounding Ljubljana has increased, now including the whole Obalno-kraška region, including Sežana. Nova Gorica remains the only employment center in the western region, apart from Ljubljana. Furthermore, there are several changes in the eastern part of Slovenia, with new clusters forming in the Koroška region, specifically Nazarje, which replaced Ravne na Koroškem as a major employment center. The Podravska and Pomurska region are now fragmented into four employment centers, with Ljutomer and Ptuj now fragmenting the Podravska region.

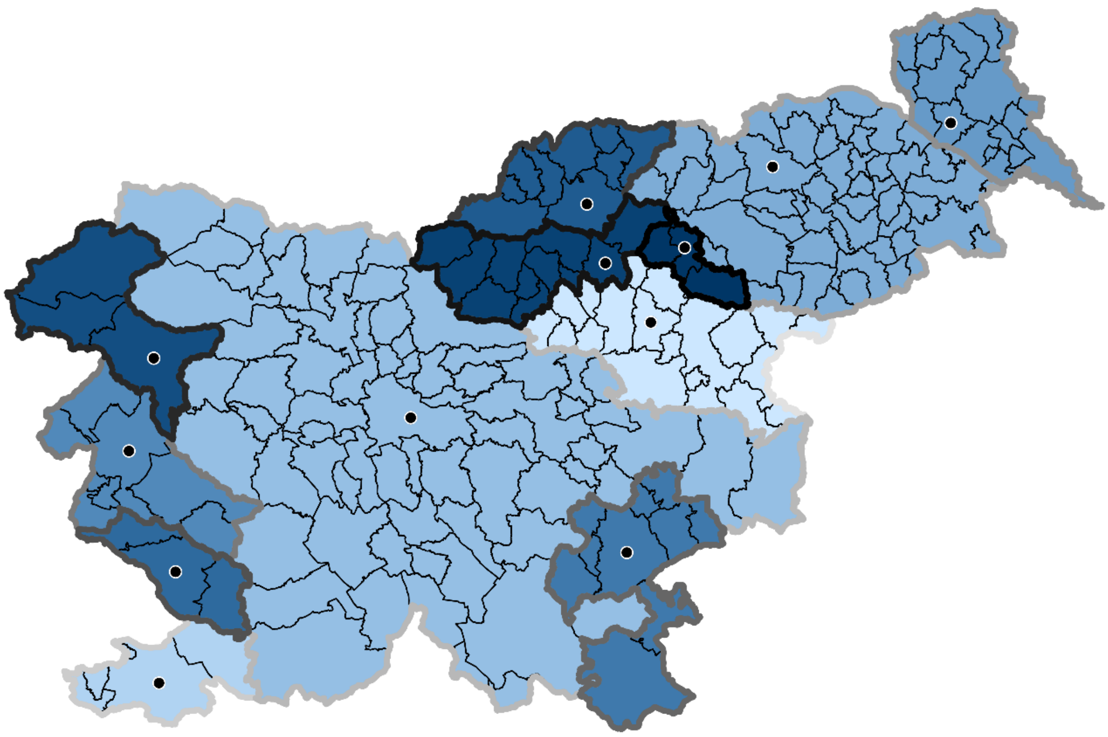
Figure 22: 12 Clusters based on commuting flows and education level in Slovenia (klasius 1), 2020



Source: Own work

Figures 23 and 24 represent the 12 generated clusters for persons with high school education for 2010 and 2020. For 2010 data, the 12 employment centers identified were Celje, Koper, Ljubljana, Maribor, Murska Sobota, Nova Gorica, Novo mesto, Sežana, Slovenj Gradec, Tolmin, Velenje and Zreče. The structure of employment centers is quite similar to that in Figure 21, with a few key differences. The Goriška region has one additional cluster, Tolmin and the Ljubljana region is more fragmented, with Novo mesto now established as an employment center in Jugo-vzhodna Slovenija. The Koroška and Savinjska regions are less fragmented, with only four employment centers recognized in the area.

Figure 23: 12 Cluster based on commuting flows and education level in Slovenia (klasius 2), 2010

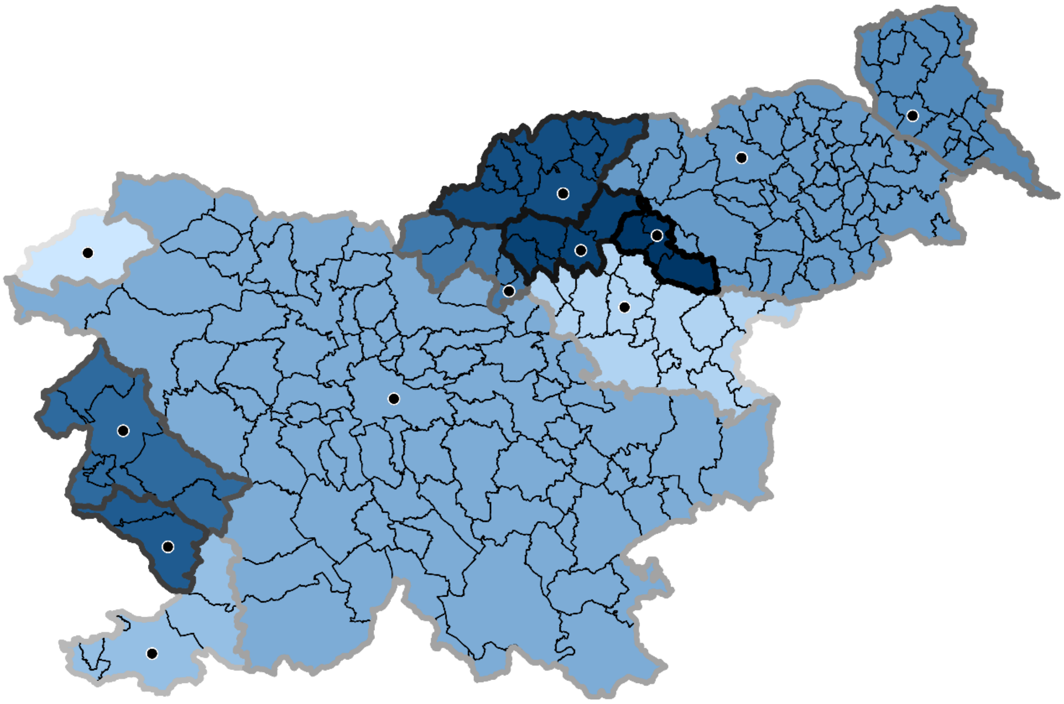


Source: Own work

In Figure 24 we can see the 12 generated clusters for the year 2020. Here there are noticeably fewer changes to the structure of the market, compared to the analysis on klasius 1. We can see the employment clusters in the western part of the country remaining intact, with the center in Tolmin now shifting to Bovec, as well as a new employment center forming in the Koroška region (Nazarje), replacing Novo mesto. The rest remain structurally the same.

Looking at the results of the analysis for klasius 1 and klasius 2, we can see that the structure of clusters is quite similar. Comparing the results to the general analysis, we can see several similarities. Ljubljana, Maribor and Murska Sobota appear to be very homogeneous, with relatively small changes in all three examples. The most diverse clusters seem to be in the Goriška and Koroško-Savinjska regions. Here we see several changes in education level as well as in time. These regions are located far from the main highway networks, thus making the commute to larger employment centers (Ljubljana, Maribor, etc.) less likely. Instead, smaller, regional clusters form in those areas, which are more likely to change given different constraints. Nevertheless, Figures 21-24 show us that commuting flow patterns between persons with elementary school education and high school education are not that diverse.

Figure 24: 12 Clusters based on commuting flows and education level in Slovenia (klasius 2), 2020



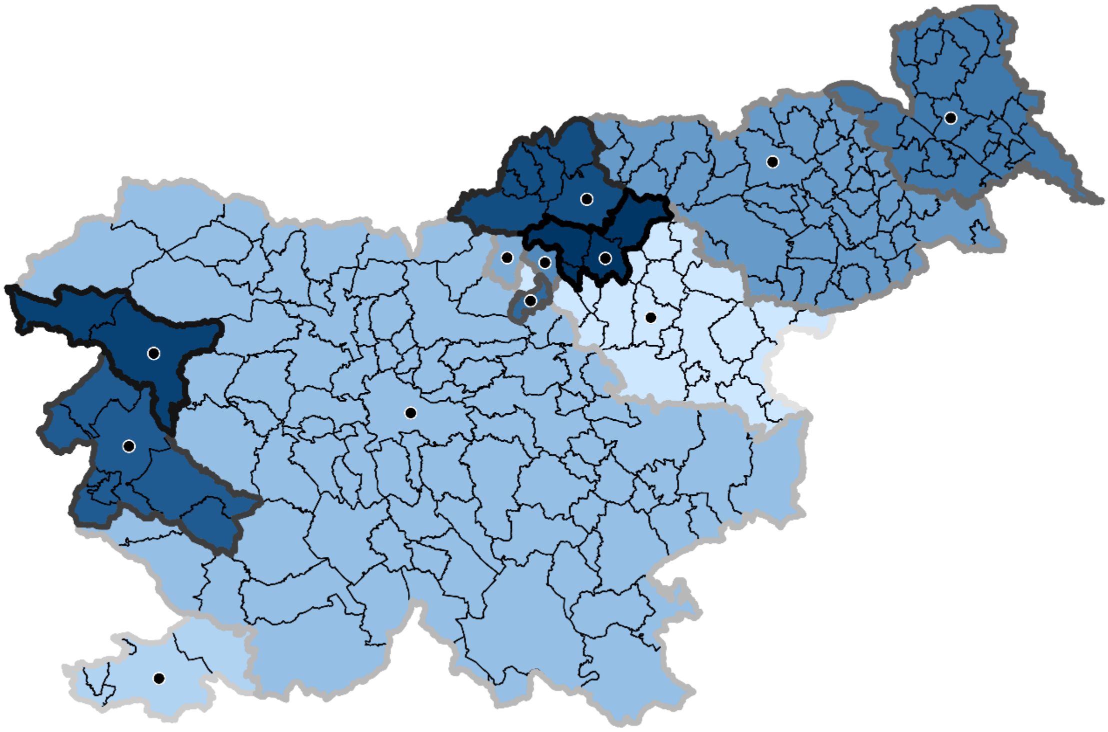
Source: Own work

Finally, in Figures 25 and 26 show the results of the cluster analysis for highly educated persons for 2010 and 2020. Firstly, Figure 25 depicts the 12 generated clusters for persons with higher education for the year 2010. The 12 identified employment centers are Celje, Koper, Ljubljana, Ljubno, Maribor, Mozirje, Murska Sobota, Nazarje, Nova Gorica, Slovenj Gradec, Tolmin and Velenje. The three major employment clusters Ljubljana, Maribor and Murska Sobota remain mostly unchanged, with Ljubljana expanding into the Obalno-kraška region. We only have two centers in the Goriška region, Nova Gorica and Tolmin, and several small employment clusters forming in the Savinjska region. Both Ljubno and Nazarje are attractive locations for high-skilled employees, due to the presence of large companies that employ a large percentage of the regions' workers (KLS Ljubno d.o.o. and Bosch in Nazarje).

Moving on to Figure 26, which presents the 12 clusters generated for highly educated people in the year 2020. The 12 clusters identified were Dravograd, Koper, Ljubljana, Ljutomer, Maribor, Murska Sobota, Muta, Nova Gorica, Radlje ob Dravi, Ravne na Koroškem, Tolmin and Vuzenica. We can see the area surrounding Ljubljana expanding significantly in the past 10 years, now covering the whole Savinjska region as well as a third of the Koroška region. Due to Ljubljana taking up more than 65% of the geographical area of Slovenia, we see several small employment clusters forming next to the border in order to fill the 12-cluster constraint.[17] The results show that in 2020 highly educated people are forced to take longer commutes to major employment centers (predominantly Ljubljana) to find appropriate work.

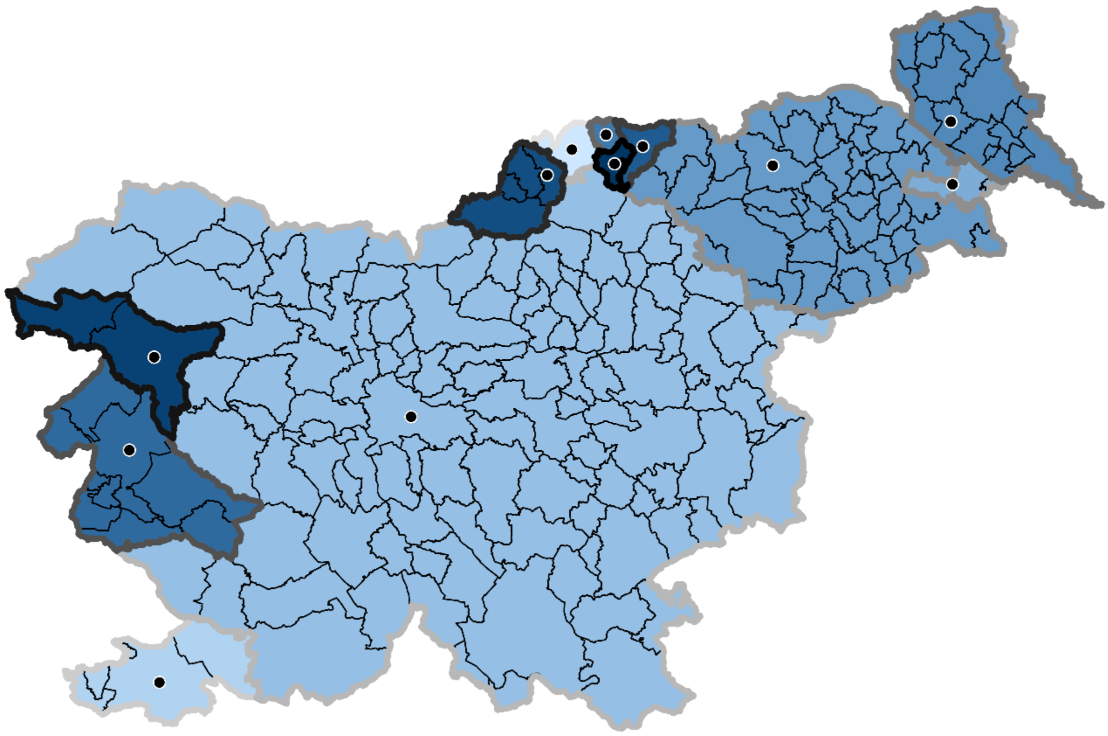
[17] If we look at the results in Figures 7-8, we can see that given a constraint of internal relative commuting flows of 50%, that is not subject to a certain number of clusters, the number of identified clusters drops from 17 to 11 from 2010-2020. We can see a similar trend in Figures 9-10.

Figure 25: 12 Clusters based on commuting flows and education level in Slovenia (klasius 3), 2010



Source: Own work

Figure 26: 12 Clusters based on commuting flows and education level in Slovenia (klasius 3), 2020



Source: Own work

4.4 Gender

Using data on commuting flows from 2000-2020, we can analyze the differences in commuting flow patterns based on gender. Table 6 represents the share of males and females working outside their county of residence from 2000 to 2020. We can see that the weighted share of male workers in the year 2000 amounted to 41.1% and increased to 53.2% in 2020. This amounts to an increase of almost 30% in total. Furthermore, the weighted share of females is smaller in 2000 compared to the share of males, amounting to 38.5% of all females. That share increased to 52.9% in 2020, amounting to an increase of 37.4%. Though the increase is higher for females, the weighted share of persons working outside the county of residence in 2020 is quite similar for both genders.

Table 6: Share of males and females working outside county of residence in Slovenia, 2000-2020

	Females				Males			
	Mean				Mean			
Year	Weighted	Unweighted	Median	Std. Dev.	Weighted	Unweighted	Median	Std. Dev.
2000	0.385	0.606	0.655	0.227	0.411	0.612	0.663	0.207
2001	0.395	0.613	0.658	0.223	0.419	0.617	0.671	0.204
2002	0.405	0.619	0.680	0.219	0.426	0.621	0.668	0.199
2003	0.415	0.625	0.674	0.213	0.438	0.629	0.675	0.196
2004	0.426	0.634	0.673	0.210	0.450	0.638	0.684	0.191
2005	0.438	0.644	0.699	0.205	0.460	0.646	0.693	0.187
2006	0.449	0.651	0.698	0.200	0.467	0.652	0.694	0.183
2007	0.460	0.660	0.704	0.196	0.475	0.657	0.701	0.178
2008	0.468	0.664	0.709	0.191	0.478	0.658	0.695	0.174
2009	0.473	0.660	0.712	0.185	0.483	0.654	0.693	0.168
2010	0.481	0.669	0.718	0.181	0.489	0.657	0.696	0.164
2011	0.487	0.673	0.716	0.178	0.497	0.661	0.703	0.162
2012	0.489	0.672	0.715	0.177	0.501	0.660	0.694	0.159
2013	0.489	0.670	0.709	0.175	0.503	0.662	0.698	0.161
2014	0.492	0.669	0.714	0.172	0.506	0.622	0.696	0.159
2015	0.496	0.673	0.717	0.170	0.508	0.665	0.694	0.160
2016	0.502	0.676	0.712	0.168	0.512	0.666	0.695	0.160
2017	0.509	0.683	0.724	0.166	0.515	0.671	0.701	0.159
2018	0.515	0.689	0.732	0.164	0.519	0.673	0.702	0.158
2019	0.518	0.693	0.732	0.164	0.521	0.675	0.707	0.159
2020	0.529	0.701	0.739	0.158	0.532	0.684	0.719	0.155

Source: Own work

4.5 Wage

Using the data on gross and net average monthly wages per county for the years 2005-2020, we were able to introduce another important variable in our empirical analysis of commuting flows. Table 7 shows the weighted average gross and net monthly wages in Slovenia for the 2005-2020 time period. The weighted average gross wage in 2005 amounted to 1,020.2 EUR and increased to 1,635.2 EUR in 2020, which corresponds with an increase of 60%. The weighted average net wage in 2005 amounted to 668.8 EUR and increased to 1084.3 EUR in 20 years. This corresponds with an increase of 61.6%.

Table 7: Average gross and net wages in Slovenia, 2005-2020

Gross					Net				
	Mean					Mean			
Year	Weighted	Unweighted	Median	Std. Dev.	Year	Weighted	Unweighted	Median	Std. Dev.
2005	1020.2	1095.0	1005.8	117.9	2005	668.8	704.2	660.7	65.4
2006	1070.6	1148.2	1060.5	117.5	2006	703.1	740.4	696.3	65.3
2007	1136.3	1218.0	1131.6	120.1	2007	757.6	798.9	757.0	68.6
2008	1228.4	1318.2	1220.6	132.8	2008	815.5	860.7	813.7	74.2
2009	1257.1	1359.9	1251.3	134.8	2009	835.4	887.7	833.3	75.2
2010	1308.9	1414.3	1309.2	131.1	2010	869.9	923.7	868.9	72.3
2011	1332.7	1443.8	1324.2	131.4	2011	887.9	944.9	884.6	73.1
2012	1332.5	1440.7	1328.7	129.1	2012	889.9	946.1	891.9	71.7
2013	1334.7	1439.4	1325.5	137.5	2013	895.8	951.1	890.7	76.1
2014	1347.0	1453.8	1333.1	140.8	2014	902.1	958.4	898.4	77.9
2015	1363.8	1469.4	1357.6	144.2	2015	910.7	966.1	906.5	79.7
2016	1391.3	1498.5	1380.4	142.5	2016	926.2	983.1	923.3	78.6
2017	1432.7	1540.2	1419.2	145.7	2017	954.0	1013.1	948.9	83.1
2018	1481.7	1593.5	1469.8	150.7	2018	982.4	1043.5	977.8	85.8
2019	1548.1	1662.0	1531.6	152.4	2019	1020.7	1082.3	1013.6	86.7
2020	1635.2	1754.6	1617.2	151.9	2020	1084.3	1150.8	1075.7	88.3

Source: Own work

5 RESULTS OF THE EMPIRICAL ANALYSIS

In this chapter, we present the results of our regression analysis using the methodology described in the previous chapters. The analysis is split in three parts. The first part of our analysis focuses on the estimation of commuting flow elasticities, for all years, using different specifications of the PPML estimator, as well as a comparison to the OLS estimation results. Secondly, we perform a PPML estimation of commuting flow elasticities by year based on gender and lastly, based on education level.

Our analysis begins with a standard OLS estimation of the equation:

$$\ln X_{ij,t} = \beta_0 + \beta_1 \ln DIST_{ij,t} + \alpha_i + \alpha_j + \alpha_t + \varepsilon_{ij,t} \tag{14}$$

Where $\ln X_{ij,t}$ corresponds to the logarithm of nominal bilateral commuting flows from the county of residence *i* to county of work *j* at time t. $\ln DIST_{ij,t}$ represents the logarithm of commuting time from county of residence *i* to county of work *j* and back, while β_0 is a constant. α_i, α_j and α_t are residence, destination, and time fixed effects. We perform the OLS estimation using residence, destination, and time fixed effects, using data from 1997-2020. The results of the analysis are shown in column (1) of Table 8.

Next, we perform a PPML estimation using three different specifications. Firstly, we perform the standard PPML estimation of equation (14) using residence, destination, and time fixed effects. Secondly, we analyze the interaction of these elasticities with time using residence-time and destination-time fixed effects. Finally, we measure the effect of wages, using the specification:

$$X_{ij,t} = \exp\left(\beta_0 + \beta_1 + \ln DIST_{ij,t} + \alpha_i + \alpha_j + \alpha_t + \ln f(w_{i,t}, w_{j,t}) + \varepsilon_{ij,t}\right) \tag{15}$$

where:

$$f\left(w_{i,t}, w_{j,t}\right) = \frac{w_{j,t}}{w_{i,t}} \tag{16}$$

The results of these analyses can be found in Table 8, under (2), (3) and (4), respectively. Looking at the results in Table 8, we can see that the estimated elasticities using the OLS method are considerably different from those corresponding to the PPML method. This confirms our initial motivation that the OLS estimator leads to empirically different results. On the contrary estimates of commuting elasticities in columns (2)–(4) in Table 8 appear to be quite similar. Our analysis shows that the elasticity of the number of daily commuters with respect to commuting time (to and back) is negative at -2.891, meaning that if commuting time increases by 1%, the number of daily commuters will drop by almost 2.9%. We see a similar result in column (3). All results are statistically significant with a p-value less than 0. 1%, except for Wage in (4), which is proven to be statistically insignificant.

Table 8: OLS and PPML estimations of commuting elasticities

Estimator	OLS	PPML		
Dependant Variable	(1)	(2)	(3)	(4)
Time Back	-2.232***	-2.891***	-2.895***	-2.863***
	(0.0169)	(0.0359)	(0.0361)	(0.0359)
Wage				-0.0657
				(0.0586)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	473,597	1,487,628	1,487,628	1,313,535
R-squared	0.735			
Chi-squared		6,476.2	6,438.9	6,357.5
Log likelihood	-621,488	-7,892,419	-7,769,135	-6,966,804
AIC	1,242,979	15,784,841	15,538,274	13,933,614

Note. * p<0.05; ** p<0.01; *** p<0.001

Source: Own work

In the second part of our analysis, we look at the elasticities of commuting flows by gender, using the same specifications as in (3) using county of residence, county of destination and time fixed effects. In order to measure the change of elasticities in time, we performed our analysis on a by-year basis, generating elasticities for each year separately. We applied the model to data from 2000-2020, eliminating the years 1997-1999[18] from our estimations. Another independent variable introduced to the model is *Time Back Intra*, which gives us information on the elasticity of daily commuting flows for persons for whom the county of work is the same as the county of residence. This extension is justified by Yotov (2012), who argues that time variation in commuting cost may only be identified by separating the intra- and extra-county commuting cost.

The results of the analysis are presented in Table 9. The results show that commuting elasticities are increasing in time, for both genders. Looking at the elasticities for males, we can see the coefficients increasing in value through the years from -2.783 in 2000 to -2.611 in 2020. We see a similar trend in the results for females. In 2000 the coefficient was -3.017 and increased to -2.819 in 2020. This implies that commuting elasticities are increasing in time, meaning that commuting times have a decreasing effect on the number of commuters in Slovenia. Similarly, we can see that women have lower commuting elasticities than men, implying there are less flexible in their choice of commute. All results are statistically significant with a p value of less than 0.1%.

[18] We eliminated the years 1997-1999 from our analysis because the data on commuting flows is only available after the year 2000. Thus, estimating elasticities prior to 2000 would provide no additional information to our analysis.

Table 9: PPML estimations of commuting elasticities by year based on gender

Estimator:	(1)	(2)	(3)
Dependant Variable:	All	Male	Female
Time Back 2000	-2.881***	-2.783***	-3.017***
	(0.0401)	(0.0392)	(0.0427)
Time Back 2001	-2.866***	-2.771***	-3.000***
	(0.0400)	(0.0389)	(0.0428)
Time Back 2002	-2.851***	-2.755***	-2.984***
	(0.0400)	(0.0390)	(0.0427)
Time Back 2003	-2.832***	-2.736***	-2.965***
	(0.0402)	(0.0392)	(0.0429)
Time Back 2004	-2.815***	-2.720***	-2.947***
	(0.0402)	(0.0392)	(0.0428)
Time Back 2005	-2.798***	-2.704***	-2.929***
	(0.0402)	(0.0391)	(0.0428)
Time Back 2006	-2.783***	-2.690***	-2.912***
	(0.0405)	(0.0394)	(0.0432)
Time Back 2007	-2.797***	-2.705***	-2.926***
	(0.0403)	(0.0390)	(0.0433)
Time Back 2008	-2.786***	-2.695***	-2.914***
	(0.0404)	(0.0392)	(0.0435)
Time Back 2009	-2.781***	-2.692***	-2.904***
	(0.0404)	(0.0392)	(0.0432)
Time Back 2010	-2.771***	-2.685***	-2.889***
	(0.0404)	(0.0392)	(0.0431)
Time Back 2011	-2.760***	-2.674***	-2.879***
	(0.0409)	(0.0398)	(0.0437)
Time Back 2012	-2.756***	-2.668***	-2.876***
	(0.0409)	(0.0394)	(0.0441)
Time Back 2013	-2.752***	-2.664***	-2.873***
	(0.0410)	(0.0396)	(0.0442)
Time Back 2014	-2.744***	-2.655***	-2.866***
	(0.0411)	(0.0396)	(0.0443)
Time Back 2015	-2.735***	-2.647***	-2.856***
	(0.0411)	(0.0397)	(0.0441)
Time Back 2016	-2.725***	-2.637***	-2.846***
	(0.0411)	(0.0397)	(0.0443)
Time Back 2017	-2.731***	-2.643***	-2.851***
	(0.0413)	(0.0396)	(0.0448)
Time Back 2018	-2.720***	-2.633***	-2.840***
	(0.0415)	(0.0399)	(0.0448)
Time Back 2019	-2.712***	-2.623***	-2.832***
	(0.0415)	(0.0400)	(0.0448)
Time Back 2020	-2.698***	-2.611***	-2.819***
	(0.0415)	(0.0399)	(0.0449)
Time Back Intra	-2.635***	-2.509***	-2.802***
	(0.0634)	(0.0613)	(0.0681)
Fixed Effects	Yes	Yes	Yes
Observations	1,487,628	1,487,628	1,487,628
Chi-squared	14,086.4	13,161.5	13,169.9
Log likelihood	-7,649,342	-4,688,525	-3,782,447
AIC	15,298,730	9,377,095	7,564,939

Note. * p<0.05; ** p<0.01; *** p<0.001

Source: Own work

Finally, using the same specification, we analyze commuting elasticities by education level. We applied the model to data from 2010-2020, which corresponds to the timeframe of our data set. Looking at the results in Table 10 we can see that the coefficients are increasing in value through the years, for all levels of education. For the least skilled workers (klasius 1) the coefficient was -2.950 in 2010 and increased to -2.890 in 2020. The coefficient for intermediate skill level (klasius 2) is slightly higher at -2.841 for the year 2010 and increases to -2.779 in 2020. The coefficient for the high skilled (klasius 3) is the highest at -2.631 in 2010 and increased to -2.595 in 2020. According to our results, we can conclude that the elasticities are increasing in time for all levels of education. Furthermore, highly educated people have higher elasticities, meaning that they are more flexible in their choice of commute. All results are statistically significant at p-value less than 0.1%.

Table 10: PPML estimation of commuting elasticities by year based on education

Estimator:	(1)	(2)	(3)	(4)
Dependant Variable:	All	Klasius 1	Klasius 2	Klasius 3
Time Back 2010	-2.745***	-2.950***	-2.841***	-2.631***
	(0.0395)	(0.0496)	(0.0343)	(0.0374)
Time Back 2011	-2.734***	-2.948***	-2.832***	-2.625***
	(0.0401)	(0.0510)	(0.0347)	(0.0382)
Time Back 2012	-2.730***	-2.952***	-2.827***	-2.625***
	(0.0401)	(0.0515)	(0.0348)	(0.0377)
Time Back 2013	-2.726***	-2.953***	-2.825***	-2.625***
	(0.0402)	(0.0517)	(0.0350)	(0.0376)
Time Back 2014	-2.718***	-2.943***	-2.817***	-2.622***
	(0.0403)	(0.0516)	(0.0350)	(0.0379)
Time Back 2015	-2.709***	-2.934***	-2.808***	-2.619***
	(0.0403)	(0.0515)	(0.0351)	(0.0380)
Time Back 2016	-2.699***	-2.929***	-2.798***	-2.611***
	(0.0403)	(0.0514)	(0.0352)	(0.0381)
Time Back 2017	-2.705***	-2.927***	-2.805***	-2.622***
	(0.0405)	(0.0518)	(0.0352)	(0.0385)
Time Back 2018	-2.695***	-2.918***	-2.796***	-2.614***
	(0.0406)	(0.0518)	(0.0354)	(0.0386)
Time Back 2019	-2.686***	-2.898***	-2.788***	-2.609***
	(0.0407)	(0.0513)	(0.0356)	(0.0386)
Time Back 2020	-2.673***	-2.890***	-2.779***	-2.595***
	(0.0407)	(0.0514)	(0.0356)	(0.0387)
Time Back Intra	-2.593***	-2.743***	-2.674***	-2.586***
	(0.0624)	(0.0738)	(0.0549)	(0.0606)
Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,141,668	377,026	377,026	377,026
Chi-squared	11,348.7	10,337.1	18,389.7	9,729.1
Log likelihood	-5,918,589	-285,293	-1,033,132	-568,690
AIC	11,837,204	570,612	2,066,289	1,137,405

Note. * p<0.05; ** p<0.01; *** p<0.001

Source: Own work

6 DISCUSSION

Our empirical analysis of the commuting patterns of workers in Slovenia uncovered several key results. The data on commuting flows shows that the share of persons working outside the county of residence has been increasing for the past 20 years. This holds more so for women than for men, however, the share of persons working outside the county of residence continues to be higher for men than for women. Similar conclusions can be drawn from the data on commuting flows based on education level. The share of persons working outside the county of residence has been steadily increasing for all levels of education but remains the highest for persons with higher education. In sum, men are more to likely commute to work than women and highly educated people are more likely to commute than persons with a high school education or less.

Our analysis of the geographical distribution of commuters showed that the number of commuters has increased proportionally across the whole country, but most notably around areas surrounding major employment centers such as Ljubljana, Maribor, Murska Sobota, etc. What is more, the data shows that the number of commuters also increased in areas located farther away from these centers, even in counties located next to the borders. This would imply that not only are regional employment centers attracting more workers from neighboring counties but that their reach is expanding to counties located further away. This is further supported by the fact that the average commuting time has increased from 35 min to 40 min in the past 20 years, even though infrastructural developments and improvements shortened commuting times between many locations. Another possible implication of these findings is that not only are established employment centers becoming stronger and expanding their reach, but smaller more local employment markets are losing their strength and appeal or are shrinking in size.

Looking at the results of our cluster analysis, we can see that not only has the number of commuters increased in the past years but so has the interconnectedness of the counties. In 2010, when restricted by a 50% internal commuting flows stopping criterion, the algorithm produced 17 distinct employment clusters that matched that specification. In 2020 the number of clusters dropped to 12. This means that in the past 10 years there has been an increased flow of commuters to neighboring counties of employment, thus decreasing the total number of employment clusters. A similar result was found when applying a minimum external commuting flow stopping criterion. In 2010, when applying this specification, the algorithm produced 11 distinct clusters, while in 2020 the algorithm identified only 4. This result is especially concerning since it implies that there exists a minimal commuting flow to these four employment centers from all counties, even those located in more remote parts of the country.

When imposing a 12-cluster constraint on the data on commuting flows, we can conclude that the generated employment centers and their corresponding geographical reach do not coincide with the 12 statistical regions. The most notable result of this analysis regards the employment reach surrounding the country's capital, Ljubljana. The data shows that the geographical reach of Ljubljana not only encircles the Osrednjeslovenska region but expands to the north-western and south-eastern parts of the country. This is not the case for the north-eastern and the western parts of the country, where we see the formation of a diverse array of employment centers. This is most likely due to the structure of Slovenia's motorway network. If we look back at the geographical reach of Ljubljana, we can see that all three regions spanning this cluster have easy access to the A1 and A2 motorway networks, connecting each region to the capital. This allows workers a faster and easier commute to the capital, compared to other regions with limited access to this network. For example, if we look at the Goriška and Koroška regions, they might seem geographically close to Ljubljana, but when we factor in motorway access, we can see that commuting to the capital from these two regions would be inefficient and time-consuming. The same holds for clusters forming around Celje, Maribor and Murska Sobota. Even though these regions are connected to the capital through the A1 motorway network, the workers would still have to commute upwards of 1 hour each way to reach their place of employment.

When we compare the results of the 12-cluster analysis for the years 2010 and 2020, we can see the most noticeable change in the geographical reach of Ljubljana, which has expanded to include the Obalno-kraška region as well as several other bordering counties. Due to the expansion of this cluster, we see a shift in the structure of other employment centers and their employment reach in order to satisfy the constraint. We note several minuscule changes to the employment centers in the Koroška, Savinjska and Goriška regions, while the clusters surrounding Maribor and Murska Sobota remain homogeneous. In order to achieve a deeper understanding of the structure of the labor market, we performed the same 12-cluster analysis for each klasius of education.

The results of the 12-cluster analysis for persons with elementary school education or less (klasius 1) for the year 2010 show a similar result to those based on data for all education levels. The employment reach of Ljubljana remains high, while Maribor and Murska Sobota remain the dominant employment centers in their region. In the remaining parts of the country, we see the employment structure shift to centers of heavy industry such as Ravne na Koroškem, Rogaška Slatina, Slovenj Gradec, Velenje and Zreče, around which clusters are formed. In 2020 we once again see the geographical reach of Ljubljana expand to the west, absorbing several regional employment centers. Interestingly, in 2020 the north-eastern part of Slovenia seems to be more fragmented than in 2010, with several new employment centers emerging (Ljutomer, Nazarje and Ptuj). The structure of the labor market in 2020 seems to have significant differences compared to the one produced in the general analysis, which is a valid result when we consider the fact that these results apply to only a fragment of the working population.

The results for the 12-cluster analysis for persons with high school education (klasius 2) for the year 2010 seem to be very much in line with the results based on all employees, with a few notable differences. For the first time, we see the formation of an employment cluster in the south-eastern part of the country, with the center located in Novo mesto. This implies a smaller employment reach of Ljubljana, which is now fragmented not only in the western but also in the eastern part of the country. The rest of the results seem to be in line with the general analysis apart from the Koroška region, which is now considered a singular employment cluster due to the 12-cluster constraint. In 2020 we see the geographical reach of Ljubljana expand and absorb the new cluster surrounding Novo mesto as well as Tolmin in the west. What is more, we see far fewer changes between the years 2010 and 2020 compared to the results for klasius 1, implying that the regional employment centers for people with high school education are much more firmly positioned in their regions and far less susceptible to change.

The strength of Ljubljana's employment market is even more pronounced when looking at the results of the 12-cluster analysis for highly educated people (klasius 3). Already in 2010, we see that the cluster surrounding the capital is much larger than in the case of klasius 1 and 2, but it also seems to shrink the geographical reach of other nearby clusters. For example, we see the area of Ljubljana's cluster dig into the Goriška, Obalno-kraška as well as parts of the Savinjska region, expanding its reach into more remote parts of the country, where commutes to the capital would be considerably longer. These findings are even more worrisome when we look at the results for 2020. Here we can see the cluster surrounding Ljubljana absorb most of the Savinjska and Koroška regions, expanding to the northern border. These results imply that there is only a handful of employment centers in Slovenia that satisfy the demand for high-skilled jobs, forcing people with higher education to commute longer distances for appropriate work. These findings speak to the vast differences in commuting patterns between education levels, as well as the noticeable developmental gap between Ljubljana and its surrounding counties compared to the rest of the country. These trends have already been documented in Bole (2004) and Bole and Gabrovec (2012), the results of which are further supported by our analysis.

Finally looking at the results of our empirical analysis we can conclude that the elasticity of the number of daily commuters with respect to commuting time (to and back) is negative at -2.9%, implying that given a 1% increase in commuting time, the number of daily commuters would drop by almost 2.9%. We can further conclude that net wage is not a statistically significant variable in determining the commuting choice. Furthermore, our analysis shows that commuting elasticities are different for men and women. Namely, men have higher (lower in absolute terms) commuting elasticities than women, meaning that women are less flexible in their commuting choices than men. Looking at the commuting elasticities in time, we can see that the elasticities are increasing in time for both men and women, implying that workers of both genders are becoming more flexible in their choice of commute. We find similar results for education. We see commuting elasticities for all klasius of education increasing in time, implying that all workers regardless of education are increasing flexibility in their choice of commute. Finally, looking at the results for each klasius of education we can see that the elasticities are increasing with education, implying that more educated people are more flexible in their choice of commute than less educated people.

CONCLUSION

It was the purpose of this thesis to examine the commuting patterns of workers in Slovenia and characterize the geographical structure of Slovenia's labor market. We analyzed the characteristics of commuting flows in Slovenia based on data at the level of the county of residence and work for all employed persons in Slovenia, generally, by gender and educational attainment. Our analysis of the geographical distribution of commuters showed that the number of commuters has increased proportionally across the whole country, but most notably around areas surrounding major employment centers such as Ljubljana, Maribor, Murska Sobota, etc. What is more, the data showed that the number of commuters also increased in areas located farther away from these centers, even in counties located next to the borders. Furthermore, our analysis showed that the share of persons working outside the county of residence has been steadily increasing for all levels of education but remains the highest for persons with higher education. The data also showed that the share of persons working outside the county of residence has been increasing for both genders, however, the share continues to be higher for men than for women. In sum, men are more to likely commute to work than women and highly educated people are more likely to commute than persons with a high school education or less.

Secondly, we examined the structure of Slovenia's labor market by generating flow-based clusters dependent on the identification of regional employment centers, using a clustering algorithm applied to data on bilateral commuting flows. Our analysis produced several key results the most notable of which is the employment reach surrounding the country's capital, Ljubljana. These results imply that there is only a handful of employment centers in Slovenia that satisfy the demand for high-skilled jobs, forcing people with higher education to commute longer distances for appropriate work. These findings speak to the vast differences in commuting patterns between education levels, as well as the noticeable developmental gap between Ljubljana and its surrounding counties compared to the rest of the country.

Lastly, we modelled the distance elasticities for the labor market in Slovenia, using the Poisson Pseudo-Maximum-Likelihood estimator. We can conclude that the elasticity of the number of daily commuters with respect to commuting time (to and back) is negative at -2.9%. We can further conclude that net wage is not a statistically significant variable in determining commuting choice. Furthermore, our analysis shows that commuting elasticities are different for men and women. Namely, men have higher (lower in absolute terms) commuting elasticities than women, meaning that women are less flexible in their commuting choices than men. We find similar results for education. We see commuting elasticities for all klasius of education increasing in time, implying that all workers regardless of education are increasing flexibility in their choice of commute. Finally, looking at the results for each klasius of education we can see that the elasticities are increasing with education, implying that more educated people are more flexible in their choice of commute than less educated people.

REFERENCE LIST

1. Ahlfeldt, G. M. (2011). If Alonso Was Right: Modeling Accessibility and Explaining the Residential Land Gradient. *Journal of Regional Science*, *51*(2), 318–338. https://doi.org/10.1111/j.1467-9787.2010.00694.x
2. Ahlfeldt, G. M., & Wendland, N. (2016). The Spatial Decay in Commuting Probabilities: Employment Potential vs. Commuting Gravity. *Economics Letters*, *143*, 125–129. https://doi.org/10.1016/j.econlet.2016.04.004
3. Ahrens, A., & Lyons, S. (2021). Do Rising Rents Lead to Longer Commutes? A Gravity Model of Commuting Flows in Ireland. *Urban Studies*, *58*(2), 264–279. https://doi.org/10.1177/0042098020910698
4. Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press.
5. Artuç, E., Chaudhuri, S., & McLaren, J. (2010). Trade Shocks and Labor Adjustment: A Structural Empirical Approach. *American Economic Review*, *100*(3), 1008–1045. https://doi.org/10.1257/aer.100.3.1008
6. Artuç, E., & McLaren, J. (2015). Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility. *Journal of International Economics*, *97*(2), 278–294. https://doi.org/10.1016/j.jinteco.2015.06.001
7. Bartolj, T., Murovec, N., & Polanec, S. (2022). Reported Time Allocation and Emotional Exhaustion During Covid-19 Pandemic Lockdown in Slovenia. *The Economic and Labour Relations Review*, *33*(1), 117–137. https://doi.org/10.1177/10353046211070042
8. Boarnet, M. G. (2005). An Empirical Model of Intrametropolitan Population and Employment Growth. *Papers in Regional Science*, *73*(2), 135–152. https://doi.org/10.1111/j.1435-5597.1994.tb00607.x
9. Bole, D. (2004). Daily Mobility of Workers in Slovenia. *Acta Geographica Slovenica*, *44*(1), 25–45. https://doi.org/10.3986/AGS44102
10. Bole, D. (2011). Changes in Employee Commuting: A Comparative Analysis of Employee Commuting to Major Slovenian Employment Centers from 2000 to 2009. *Acta Geographica Slovenica*, *51*(1), 93–108. https://doi.org/10.3986/AGS51104

11. Bole, D., & Gabrovec, M. (2012). Daily Commuters in Slovenia. *Geografski Vestnik*, *84*(1), 171–185. https://doi.org/10.3986/AGS44102

12. Brezina, T., Tiran, J., Ogrin, M., & Laa, B. (2021). COVID-19 Impact on Daily Mobility in Slovenia. *Acta Geographica Slovenica*, *61*(2), 91–107. https://doi.org/10.3986/AGS.9390

13. Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The Scaling Laws of Human Travel. *Nature*, *439*(7075), 462–465. https://doi.org/10.1038/nature04292

14. Carlino, G. A., & Mills, E. S. (1987). The Determinants of County Growth. *Journal of Regional Science*, *27*(1), 39–54. https://doi.org/10.1111/j.1467-9787.1987.tb01143.x

15. Choukroun, J.-M. (1975). A General Framework for the Development of Gravity-Type Trip Distribution Models. *Regional Science and Urban Economics*, *5*(2), 177–202. https://doi.org/10.1016/0166-0462(75)90003-4

16. Commins, N., & Nolan, A. (2011). The Determinants of Mode of Transport to Work in the Greater Dublin Area. *Transport Policy*, *18*(1), 259–268. https://doi.org/10.1016/j.tranpol.2010.08.009

17. Crane, R. (2007). Is There a Quiet Revolution in Women's Travel? Revisiting the Gender Gap in Commuting. *Journal of the American Planning Association*, *73*(3), 298–316. https://doi.org/10.1080/01944360708977979

18. Drobne, S., Bogataj, M., & Lisec, A. (2008). *The Influence of Accessibility to Inter-Regional Commuting Flows in Slovenia*. Retrieved November 20, 2022 from https://www.researchgate.net/publication/252858582_The_Influence_of_Accessibility_to_Inter-Regional_Commuting_Flows_in_Slovenia

19. Drobne, S., Bogataj, M., & Lisec, A. (2012). Dynamics and Local Policy in Labour Commuting. *BSRJ*, *3*(2), 14–26. https://doi.org/10.2478/v10305-012-0009-x

20. Drobne, S., Bogataj, M., Zupan, M., & Lisec, A. (2011). Dynamics and Local Policy in Commuting: Attractiveness and Stickiness of Slovenian Municipalities. *Proceedings of the 11th International Symposium on Operational Research in Slovenia, SOR 2011*, 323–328. https://www.researchgate.net/publication/287434723_Dynamics_and_local_policy_in_commuting_Attractiveness_and_stickiness_of_Slovenian_municipalities

21. Drobne, S., & Lakner, M. (2014). Which Distance-Decay Function for Migration and Which One for Commuting? A Case Study of Slovenia. *Croatian Operational Research Review*, *5*(2), 259–272. https://doi.org/10.17535/crorr.2014.0012

22. Drobne, S., Rajar, T., & Lisec, A. (2013). Dynamics of Migration and Commuting to the Urban Centres of Slovenia, 2000-2011. *Geodetski Vestnik*, *57*(02), 313–332. https://doi.org/10.15292/geodetski-vestnik.2013.02.313-332

23. Evers, G. H. M. (1989). Simultaneous Models for Migration and Commuting: Macro and Micro Economic Approaches. In J. Van Dijk, H. Folmer, H. W. Herzog, & A. M. Schlottmann (Eds.), *Migration and Labor Market Adjustment* (pp. 177–197). Springer Netherlands. https://doi.org/10.1007/978-94-015-7846-2_8

24. Fotheringham, A. S. (2001). Spatial Interaction Models. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 14794–14800). Pergamon: Elsevier. https://doi.org/10.1016/B0-08-043076-7/02519-5

25. Freedman, O., & Kern, C. R. (1997). A Model of Workplace and Residence Choice in Two-Worker Households. *Regional Science and Urban Economics*, *27*(3), 241–260. https://doi.org/10.1016/S0166-0462(96)02158-8

26. Giménez-Nadal, J. I., Molina Chueca, J. A., & Velilla, J. (2020). Trends in Commuting Time of European Workers: A Cross-Country Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3525244

27. Giménez-Nadal, J. I., Molina, J. A., & Velilla, J. (2019). Modeling Commuting Time in the Us: Bootstrapping Techniques to Avoid Overfitting. *Papers in Regional Science*, *98*(4), 1667–1684. https://doi.org/10.1111/pirs.12424

28. González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008). Understanding Individual Human Mobility Patterns. *Nature*, *453*(7196), 779–782. https://doi.org/10.1038/nature06958

29. Gutiérrez-i-Puigarnau, E., & van Ommeren, J. N. (2015). Commuting and Labour Supply Revisited. *Urban Studies*, *52*(14), 2551–2563. https://doi.org/10.1177/0042098014550452

30. Han, X.-P., Hao, Q., Wang, B.-H., & Zhou, T. (2011). Origin of the Scaling Law in Human Mobility: Hierarchy of Traffic Systems. *Physical Review E*, *83*(3), 1–4. https://doi.org/10.1103/PhysRevE.83.036117

31. Hidalgo, C. A., & Rodriguez-Sickert, C. (2008). The Dynamics of a Mobile Phone Network. *Physica A: Statistical Mechanics and Its Applications*, *387*(12), 3017–3024. https://doi.org/10.1016/j.physa.2008.01.073

32. Horton, F. E., & Wittick, R. I. (1969). A Spatial Model for Examining the Journey-to-Work in a Planning Context. *The Professional Geographer*, *21*(4), 223–226. https://doi.org/10.1111/j.0033-0124.1969.00223.x

33. Hsieh, C.-T., & Moretti, E. (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics*, *11*(2), 1–39. https://doi.org/10.1257/mac.20170388

34. Jacob, N., Munford, L., Rice, N., & Roberts, J. (2019). The Disutility of Commuting? The Effect of Gender and Local Labor Markets. *Regional Science and Urban Economics*, *77*, 264–275. https://doi.org/10.1016/j.regsciurbeco.2019.06.001

35. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science*, *306*(5702), 1776–1780. https://doi.org/10.1126/science.1103572

36. Kareiva, P. M., & Shigesada, N. (1983). Analyzing Insect Movement as a Correlated Random Walk. *Oecologia*, *56*(2–3), 234–238. https://doi.org/10.1007/BF00379695

37. Klafter, J., Shlesinger, M. F., & Zumofen, G. (1996). Beyond Brownian Motion. *Physics Today*, *49*(2), 33–39. https://doi.org/10.1063/1.881487

38. Kreindler, G. E., & Miyauchi, Y. (2021). Measuring Commuting and Economic Activity Inside Cities with Cell Phone Records. *The Review of Economics and Statistics*, 1–48. https://doi.org/10.1162/rest_a_01085

39. Lenormand, M., Huet, S., Gargiulo, F., & Deffuant, G. (2012). A Universal Model of Commuting Networks. *PLoS ONE*, *7*(10), 1–11. https://doi.org/10.1371/journal.pone.0045985

40. Lyons, G., & Chatterjee, K. (2008). A Human Perspective on the Daily Commute: Costs, Benefits and Trade-offs. *Transport Reviews*, *28*(2), 181–198. https://doi.org/10.1080/01441640701559484

41. Ma, K.-R., & Banister, D. (2006). Extended Excess Commuting: A Measure of the Jobs-Housing Imbalance in Seoul. *Urban Studies*, *43*(11), 2099–2113. https://doi.org/10.1080/00420980600945245

42. Manning, A., & Petrongolo, B. (2017). How Local Are Labor Markets? Evidence from a Spatial Job Search Model. *American Economic Review*, *107*(10), 2877–2907. https://doi.org/10.1257/aer.20131026

43. Mantegna, R. N., & Stanley, H. E. (1994). Stochastic Process with Ultraslow Convergence to a Gaussian: The Truncated Lévy Flight. *Physical Review Letters*, *73*(22), 2946–2949. https://doi.org/10.1103/PhysRevLett.73.2946

44. McArthur, D. P., Kleppe, G., Thorsen, I., & Ubøe, J. (2011). The Impact of Monetary Costs on Commuting Flows: The Impact of Monetary Costs on Commuting Flows. *Papers in Regional Science*, *92*(1), 69–86. https://doi.org/10.1111/j.1435-5957.2011.00405.x

45. Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area. *The American Economic Review*, *57*(2), 197–210.

46. Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, *108*(12), 3855–3890. https://doi.org/10.1257/aer.20151507

47. Muth, R. F. (1969). *Cities and Housing; the Spatial Pattern of Urban Residential Land Use*. Chicago: University of Chicago Press.

48. Novaco, R. W., & Gonzalez, O. I. (2009). Commuting and Well-Being. In *Technology and psychological well-being* (pp. 174–205). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511635373.008

49. O'Kelly, M. E. (2009). Spatial Interaction Models. In *International Encyclopedia of Human Geography* (pp. 365–368). Pergamon: Elsevier. https://doi.org/10.1016/B978-008044910-4.00529-0

50. O'Kelly, M. E., Niedzielski, M. A., & Gleeson, J. (2012). Spatial Interaction Models from Irish Commuting Data: Variations in Trip Length by Occupation and Gender. *Journal of Geographical Systems*, *14*(4), 357–387. https://doi.org/10.1007/s10109-011-0159-3

51. Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., & Barabási, A.-L. (2007). Structure and Tie Strengths in Mobile Communication Networks. *Proceedings of the National Academy of Sciences*, *104*(18), 7332–7336. https://doi.org/10.1073/pnas.0610245104

52. Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying Social Group Evolution. *Nature*, *446*(7136), 664–667. https://doi.org/10.1038/nature05670

53. Persyn, D., & Torfs, W. (2016). A Gravity Equation for Commuting with an Application to Estimating Regional Border Effects in Belgium. *Journal of Economic Geography*, *16*(1), 155–175. https://doi.org/10.1093/jeg/lbv003

54. Ramos, R. (2016). Gravity Models: A Tool for Migration Analysis. *IZA World of Labor*. https://doi.org/10.15185/izawol.239

55. Ren, Y., Ercsey-Ravasz, M., Wang, P., González, M. C., & Toroczkai, Z. (2014). Predicting Commuter Flows in Spatial Networks Using a Radiation Model Based on Temporal Ranges. *Nature Communications*, *5*(1), 1–9. https://doi.org/10.1038/ncomms6347

56. Renkow, M., & Hoover, D. (2000). Commuting, Migration, and Rural-Urban Population Dynamics. *Journal of Regional Science*, *40*(2), 261–287. https://doi.org/10.1111/0022-4146.00174

57. Reynolds, A. M. (2008). Optimal Random Lévy-Loop Searching: New Insights into the Searching Behaviours of Central-Place Foragers. *EPL (Europhysics Letters)*, *82*(2). https://doi.org/10.1209/0295-5075/82/20001

58. Reynolds, A. M., & Rhodes, C. J. (2009). The Lévy Flight Paradigm: Random Search Patterns and Mechanisms. *Ecology*, *90*(4), 877–887. https://doi.org/10.1890/08-0153.1

59. Rodrigue, J.-P. (2020). *The Geography of Transport Systems* (5th ed.). New York: Routledge. https://doi.org/10.4324/9780429346323

60. Rovan, J., & Sambt, J. (2003). Socio-Economic Differences Among Slovenian Municipalities: A Cluster Analysis Approach. *Metodološki Zvezki*, *19*, 265–278.

61. Santos Silva, J. M. C., & Tenreyro, S. (2006). The Log of Gravity. *Review of Economics and Statistics*, *88*(4), 641–658. https://doi.org/10.1162/rest.88.4.641

62. Shibayama, T., Sandholzer, F., Laa, B., & Brezina, T. (2021). Impact of Covid-19 Lockdown on Commuting. *European Journal of Transport and Infrastructure Research, 21*(1), 70–93. https://doi.org/10.18757/EJTIR.2021.21.1.5135

63. Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A Universal Model for Mobility and Migration Patterns. *Nature*, *484*(7392), 96–100. https://doi.org/10.1038/nature10856

64. So, K. S., Orazem, P. F., & Otto, D. M. (2001). The Effects of Housing Prices, Wages, and Commuting Time on Joint Residential and Job Location Choices. *American Journal of Agricultural Economics*, *83*(4), 1036–1048. https://doi.org/10.1111/0002-9092.00228

65. Song, C., Koren, T., Wang, P., & Barabási, A.-L. (2010). Modelling the Scaling Properties of Human Mobility. *Nature Physics*, *6*(10), 818–823. https://doi.org/10.1038/nphys1760

66. Stone, A. A., & Schneider, S. (2016). Commuting Episodes in the United States: Their Correlates with Experiential Wellbeing from the American Time Use Survey. *Transportation Research Part F: Traffic Psychology and Behaviour*, *42*(Part 1), 117–124. https://doi.org/10.1016/j.trf.2016.07.004

67. Tinbergen, J. (1962). Shaping the World Economy; Suggestions for an International Economic Policy. *Twentieth Century Fund*, *5*(1), 27–30. https://doi.org/10.1002/tie.5060050113

68. Tiwari, P., & Kawakami, T. (2001). Modes of Commuting in Mumbai: A Discrete Choice Analysis. *Review of Urban and Regional Development Studies*, *13*(1), 34–45. https://doi.org/10.1111/1467-940X.00030

69. Travisi, C. M., Camagni, R., & Nijkamp, P. (2010). Impacts of Urban Sprawl and Commuting: A Modelling Study for Italy. *Journal of Transport Geography*, *18*(3), 382–392. https://doi.org/10.1016/j.jtrangeo.2009.08.008

70. United Nations. (2014). *World Urbanization Prospects The 2014 Revision: Highlights*. Retrieved November 20, 2022 from https://population.un.org/wup/Publications/Files/WUP2014-Highlights.pdf

71. Vega, A., & Reynolds-Feighan, A. (2008). Employment Sub-Centres and Travel-to-Work Mode Choice in the Dublin Region. *Urban Studies*, *45*(9), 1747–1768. https://doi.org/10.1177/0042098008093377

72. Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., & Stanley, H. E. (1996). Lévy Flight Search Patterns of Wandering Albatrosses. *Nature, 381*(6581), 413–415. https://doi.org/10.1038/381413a0

73. Vodeb, V., Mladenović, L., Nikšič, M., Cotič, B., Šuklje Erjavec, I., Praper, S., Gulič, A., & Sendi, R. (2016). *National Report on Urban Development – Habitat III*. Retreived November 20, 2022 from https://www.gov.si/assets/ministrstva/MOP/Publikacije/9d3d8bd710/porocilo_urbani_razvoj_HabitatIII_en.pdf

74. Wilson, A. G. (1998). Land-Use/Transport Interaction Models: Past and Future. *Journal of Transport Economics and Policy*, *32*(1), 3–26.

75. Yan, X.-Y., Wang, W.-X., Gao, Z.-Y., & Lai, Y.-C. (2017). Universal Model of Individual and Population Mobility on Diverse Spatial Scales. *Nature Communications*, *8*(1–9), 1639. https://doi.org/10.1038/s41467-017-01892-8

76. Yotov, Y. V. (2012). A Simple Solution to the Distance Puzzle in International Trade. *Economics Letters*, *117*(3), 794–798. https://doi.org/10.1016/j.econlet.2012.08.032

77. Yotov, Y. V., Piermartini, R., Monteiro, J.-A., & Larch, M. (2017). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*. United Nations. https://doi.org/10.18356/57a768e5-en

78. Zipf, G. K. (1946). The P1 P2/D Hypothesis: On the Intercity Movement of Persons. *American Sociological Review*, *11*(6), 677–686. https://doi.org/10.2307/2087063

APPENDICES

Appendix 1: Povzetek (Summary in Slovene language)

Namen diplomske naloge je preučiti migracijske vzorce delavcev v Sloveniji in opisati geografsko strukturo slovenskega trga dela. Sprva bomo predstavili značilnosti migracijskih tokov v Sloveniji na podlagi podatkov o kraju prebivališča in kraju dela za vse zaposlene osebe v Sloveniji, na ravni občin. Migracijske tokove bomo prikazali na splošno, kot tudi po spolu in izobrazbi. Strukturo slovenskega trga dela bomo preučili s pomočjo algoritma, ki omogoča generiranje skupkov (ang. clustrov) in identifikacijo regionalnih zaposlitvenih centrov za posamezen skupek, na podlagi informacij o bilateralnih tokovih dnevnih migracij, na splošno, po spolu in izobrazbi. Na koncu bomo podali oceno elastičnosti delovnih migracijskih tokov z uporabo PPML (Poisson Pseudo-Maximum-Likelihood) metode. Naš primarni cilj je oceniti učinek časa vožnje na pripravljenost delavcev za vožnjo na delo, v analizo pa vključujemo učinek izobrazbe, spola in plač na vzorce dnevnih migracij v Sloveniji.