CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

Population Characteristics of Voters: Evidence from the Czech Parliamentary Election

Bachelor thesis

Author: Jakub Černý

Study program: Ekonomické teorie

Supervisor: doc. Petr Janský, Ph.D.

Year of defense: 2019

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 9, 2019

Jakub Cerny

Abstract

In 2017, nine political parties were elected into the Czech Parliament, which is the greatest number in the history of the country. This thesis analyses the voter turnout and results of particular parties, using aggregated data on the municipal level. The goal is to find a spectrum of variables that would uncover connections between the population characteristics of voters and their decisions to vote and to choose a political party. The method of weighted least squares is used for estimation and the results are tested for the presence of spatial autocorrelation. Subsequently, a spatial error model is used for the same analysis in order to observe spatial effects in the voting results and provide a comparison between the methods. There is found a significant negative connection between the voter turnout and the share of people facing distraints, unemployment, and the share of people with no education. Concerning the election results, the parties ANO, SPD, and KSCM receive greater support in municipalities with greater unemployment and lower shares of businessmen and people with university education. On the other hand, the parties ODS, Pirati, and TOP 09 evince exactly the opposite trends in these explanatory variables.

JEL Classification	H70, I21, I30, J10, J11
Keywords	voting behaviour, voter turnout, spatial analysis, parliamentary elections, weighted least squares, spatial error model, Czech Republic
Title	Population Characteristics of Voters: Evidence from the Czech Parliamentary Election

Author's e-mail	jakub3456@seznam.cz
Supervisor's e-mail	petr.jansky@fsv.cuni.cz

Abstrakt

V roce 2017 bylo do Poslanecké sněmovny v České republice zvoleno devět politických stran, což je největší počet v historii. Tato práce zkoumá volební účast a výsledky jednotlivých stran za použití agregovaných dat na obecní úrovni. Cílem je najít vhodné proměnné, které by umožnily najít souvislost mezi populačními charakteristikami obyvatel a jejich volební účastí a výběrem politické strany. Pro analýzu je použita metoda vážených nejmenších čtverců a její

výsledky jsou testovány na přítomnost prostorové autokorelace. Následně je pro stejná data použit prostorový chybový model za účelem pozorování prostorových efektů ve volebních výsledcích a za účelem porovnání obou metod. U volební účasti je nalezen negativní vztah s podílem lidí v exekuci, nezaměstnaností a podílem lidí bez základního vzdělání. Strany ANO, SPD a KSČM dosahují větší podpory v obcích s vyšší nezaměstnaností a nižšími podíly vysokoškoláků a podnikatelů. Na druhé straně, ODS, Piráti a TOP 09 vykazují opačné trendy v těchto nezávislých proměnných.

Klasifikace JEL	H70, I21, I30, J10, J11
Klíčová slova	volební chování, volební účast, prostorová analýza, parlamentní volby, metoda vážených nejmenších čtverců, prostorový chybový model, Česká republika
Název práce	Demografické charakteristiky voličů: pozorování z voleb do Poslanecké sněmovny v České republice
E-mail autora	jakub3456@seznam.cz
E-mail vedoucího práce	petr.jansky@fsv.cuni.cz

Acknowledgments

I am especially grateful to doc. Petr Janský, Ph.D., who provided initial inspiration for this topic and many useful comments during the analysis. Thanks belongs also to a dear friend of mine Emily Young who helped me with the language part of the thesis. Finally, I would like to express my gratitude to my family, whose support is really important to me, and all my friends, who discussed this topic with me and contributed to my work.

Typeset in FSV LaTeX template with great thanks to prof. Zuzana Havrankova and prof. Tomas Havranek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Černý, Jakub: *Population Characteristics of Voters: Evidence from the Czech Parliamentary Election.* Bachelor thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2019, pages 62. Advisor: doc. Petr Janský, Ph.D.

Contents

List of Tables

List of Figures

Acronyms

WLS Weighted Least Squares

SEM Spatial Error Model

ML Maximum Likelihood

OLS Ordinary Least Squares

2SLS Two-Stage Least Squares

ANO political party

ODS The Civic Democratic Party

Pirati Czech Pirate Party

SPD Freedom and Direct Democracy, political party

KSCM Communist Party of Bohemia and Moravia

CSSD Czech Social Democratic Party

TOP 09 political party

STAN Mayors and Independents, political party

Master's Thesis Proposal

Author	Jakub Černý
Supervisor	doc. Petr Janský, Ph.D.
Proposed topic	Population Characteristics of Voters: Evidence from the Czech Parliamentary Election

Research question and motivation In my bachelor thesis, I would like to investigate which population characteristics have certain interconnections with results of elections. I want to propose a spectrum of economic, socio-economic and demographic characteristics in order to find a correlation (or possible causality) between these data and results of the Czech Parliamentary election in 2017. I assume that specifically profiled parties have bigger support in regions with a greater share of voters from their target groups, for instance Christian Democrats in regions with greater number of believers and so on.

Analysing the data, I will determine which areas, compared to average, do not perform well in terms of unemployment, income, etc. and these areas I will subsequently analyse in further detail. I will attempt to find out whether a voter turnout in economically or socially excluded areas is lower or not. I assume that inhabitants of regions with worse accessibility to education and less developed services will evince lower voter turnout, because they lack either motivation to vote or awareness of purpose of the elections. Gallego (2010) studies the voter turnout with respect to education and claims that this relationship substantially varies across the countries. Nevertheless, she states that in the Czech Republic there is a large effect of education on voting and this statement I would like to verify.

Another question is whether people in economically or socially excluded areas attempt to somehow revenge through a ballot as Rodríguez-Pose (2018) suggests in his article. I would like to observe voting behaviour of people and simultaneously find determinants of support of particular parties, especially the extremist ones. Algan et al. (2017) analyse global rise of extremism and as possible roots they consider unemployment, globalisation and immigration. I suppose that people from regions with higher unemployment and lower-income tend to support the extremists more

because they offer more radical solutions to problems in order to attract voters from excluded areas. According to Ezrow et al.(2014), the support of the extremists should be also related to polarization of voters and their attachment to political parties.

Contribution There are many attempts to evaluate the results of elections and to conclude something about voters, however, analyses for commercial media are usually broad and not sophisticated and on the other hand, among academic articles there are many topics which can be elaborated in further detail or with greater accuracy. In my case, I would like to focus especially on economic characteristics of the population, which are according to Guiso et al. (2017) the most important with respect to the rise of populism in western Europe.

I was inspired by a journal article written by Karel Kouba (2007) who performed spatial analysis of voting results of five Parliamentary parties between 1990 and 2006, using aggregated data on district level. I would like work with actual data and provide certain conclusions about contemporary voting patterns in the Czech Republic and therefore I decided to analyse all the parties who nominated any candidates for the Parliamentary elections in 2017. In contrast with the year 2007, nowadays it is possible to obtain more detailed data about population which allows me to perform the analysis on municipal level and accomplish much greater accuracy of the results.

Similarly to bachelor thesis of Michal Kotrč (2017), which is concerned with Slovak extremists, I would like to find determinants of support of Czech extremist parties because the Czech political scene may be slightly neglected in terms of these analyses. The results of my work might explain certain part of voting behaviour in different regions and also provide more detailed information about voters of extremist parties. With respect to the fact that authors of academic papers, for instance Kubát (2006), consider Czech Communist Party to be an extremist party, it is a very sensitive topic for former communist country. Furthermore, people often make claims about voters of Czech Communist Party but they do not look for evidence from the results of elections and that is another reason why the structure of this electorate should be studied in detail.

Data My analysis will be based on voting results from all electoral districts in the Czech Republic but since it is not possible to obtain data about various population characteristics on this level, I want to perform the analysis on municipal level, i.e. using voting results and data from approximately 6250 municipalities in the Czech Republic. The data will be obtained mostly from the website of Czech Statistical Office and possibly on the same or closest possible level.

Methodology In my bachelor thesis, I would like to look for correlation coefficients between the results of elections and individual population characteristics.

Subsequently I will develop an econometric model in order to quantify significance and influence of particular variables, using Ordinary Least Squares method. An explained variable will be the percentage result of a party within a municipality and explanatory variables will be economic, socio-economic and demographic data from the same area. Sample size should be large enough in order to provide relevant information but the explanatory variables must be chosen carefully so that there would not appear any violation of the assumption of OLS estimators.

Outline

1. Introduction
2. Literature review
3. Hypotheses
4. Data and methodology
5. Results and discussion
6. Conclusion
7. Bibliography

Core bibliography

Kouba, K., 2007. Spatial Analysis of the Czech Party System. Institutionalization and Spatial Regimes. Czech Sociological Review. 43, 1017-1037.

Ezrow, L., Tavits, M., Homola, J., 2014. Voter Polarization, Strength of Partisanship, and Support for Extremist Parties. Comparative Political Studies. 47, 1558-1583.

Gallego, A., 2010. Understanding unequal turnout: education and voting in comparative perspective. Electoral Studies. 29(2), 239-247.

Rodríguez-Pose, A., 2018. The revenge of the places that don't matter.

Algan, Y., Guriev, S., Papaioannou, E., Passari, E., 2017. The European Trust Crisis and the Rise of Populism. BPEA Fall Conference

Kotrč, M., 2017. The Rise of Extreme Right: Evidence from the Slovak Parliamentary Election. Bachelor thesis. Charles University.

Guiso, L., Herrera, H., Morelli, M., Sonno, T., 2017. The spread of populism in Western countries. Einaudi Institute for Economics and Finance.

Queyranne, M., Hallaert J., Yackovlev, I., 2018. Inequality and Poverty Across Generations in the European Union: IMF STAFF DISCUSSION NOTE. 2018.

Kubát, M., 2006. Teorie antisystémovosti a Komunistická strana Čech a Moravy. 548-556.

Author

Supervisor

Chapter 1

Introduction

Elections are an important part of democratic regimes across the world since they significantly influence the functioning of countries and their administrative units. Many academic papers analyse voting procedures in order to find the most effective and fair parameters for specific types of voting. Another part of the academic literature is concerned with voting behaviour and the description of voters. The latter is an interesting topic for sociologists and also for politicians since they analyse the electorate in order to target specific groups of voters. Regarding the literature that analyses support of political parties, Kouba (2007) describes voting results of Czech political parties, using aggregated data on a district level. A similar analysis is provided by Kotrč (2017) who studies the rise of the extreme right based on data from Slovak municipalities.

The Czech Republic is a democratic country, whose history was significantly influenced by the communist regime in the 20th century. Nowadays, there is a broad spectrum of political subjects that successfully strive for the support of voters, including the communist party KSCM that is controversial for many people. The diversity of the political scene is demonstrated on the results of the parliamentary elections in 2017, since 9 parties received over 5% of total votes and became members of the Parliament. All these attributes provide interesting material for studying the voting results of particular parties.

The objective of the thesis is to analyse voting behaviour based on aggregated data on the municipal level. Various analyses in this field usually work with individual data from a standardized sample which implies they have more specific

data from a smaller part of the population. On the other hand, this analysis includes the whole population of the Czech Republic, which represents over 10 million people living in 6248 municipalities. The important part of the thesis is to select a spectrum of socio-economic and demographic variables that should be included in econometric models in order to describe the decision-making process of voters. The academic literature concerning voting behaviour often emphasises an important role of a social group to which an individual belongs. Based on this argument, the variables such as education, unemployment, or age are incorporated into the analysis since they might be related to social status.

Regarding the type of data being used in the thesis, the weighted least squares method is used instead of ordinary least squares since it enables to assign weights to particular observations. These observations, including mostly averages and shares of inhabitants within the municipality, are based on different numbers of people where there might be a bias in the estimates of OLS if certain statistics from a city differ from a small village. The problem is there might be many villages which in total include fewer inhabitants, however, they influence the dependent variable way more than the city. Therefore, the weights assigned according to a number of inhabitants should provide more reliable estimates in terms of the whole population.

The analysis is divided into two parts; the first one is related to the voter turnout and examines which variables might have certain connections with the participation of people in the voting procedure. The second part aims to find out which factors determine the support of political parties and decide which variables are the most significant. This analysis includes voting results of all political parties which were elected into the Parliament in 2017. The thesis works with aggregated data and this fact needs to be taken into account during the interpretation of the results. Subsequently, the data are subject to test for the presence of spatial autocorrelation and then a spatial error model is introduced. This method, considering spatial effects, is used for the same data as the WLS in order to compare the results. As a weighting scheme, there is used an inverse-distance matrix that is introduced in Chapter 5.

Chapter 2 introduces articles that provided the initial idea of this thesis and describes the literature that was used during the work. Chapter 3 summarizes the political situation in the Czech Republic in order to provide an insight into

the topic. Chapter 4 describes and summarizes the data that are used for the estimation processes. Chapter 5 is concerned with the theory behind econometric models that are applied in the computational part. Chapter 6 interprets the results of the WLS method and provides tests for the use of the spatial econometrics. Subsequently, the SEM is used and its results are compared to the first method. Chapter 7 summarizes the results and provides a conclusion of the thesis.

Chapter 2

Literature Review

This chapter introduces the literature that is relevant in terms of analysis of voting behaviour. Firstly, it describes the role of particular papers and the process in which the variables are selected into models. Secondly, there is introduced the literature that concerns finding determinants of increased support of extremist and populist parties. The latter part is important primarily for the provision of hypothesis to verify and for interpretation of final results. The very basic topic for this thesis was provided by papers regarding the idea that people from certain worse-performing regions tend to revenge to the authorities through the ballot by voting for anti-establishment parties, alternatively for extremists. Finally, the literature from the Czech Republic or Slovakia concerning a similar topic as this thesis is introduced in order to follow up with the work which has been already done. These analyses work with models regarding spatial effects and therefore provide a methodological framework for the thesis.

Selection of Variables

Academic literature related to elections and voting behaviour has a long history and there are many attempts to analyse this field in order to better understand the mechanism of the decision-making process and other determinants of final results. Hayek (1945) uses an example of the price system and argues that people sometimes develop and incorporate a functioning system even though they do not have proper knowledge of all its attributes, which might not cause harm, however, there might remain an unused potential or various malfunctions. Concerning the literature about voting behaviour, Antunes (2010) summarizes theoretical models in this field and divides them into three categories:

sociological model, psychological model, and rational choice theory. With respect to the topic of this bachelor thesis and the type of data being used for the analysis, the sociological approach, which is mainly based on the work of an Austrian-American sociologist Paul Lazarsfeld, will provide a fundamental framework for the selection of particular variables.

Lazarsfeld et al. (1944) monitor 600 individuals for a couple of months before a presidential election in the United States, analysing their voting preferences and looking for possible factors that might influence the decision making process. The initial assumption is that voting is mostly based on the individual and his or her opinion, which might be potentially affected by information from the media. However, the results of the analysis indicate that a very important factor is a social group to which the person belongs and that the campaign of a political subject, in newspapers or on the radio, primarily serves for strengthening the relationship of current voters rather than for engaging new ones. Another argument states that information given by media might not be distributed to all people in the same way but vicariously by so-called "opinion leaders" who receive the information and interpret it to the people in their surroundings. The authors say that the connection between the voting behaviour and the social group is so significant that it might be possible to use only socio-economic status, religion and the area where the person lives in order to be able to describe his or her electoral choice.

Among the variables related to the socio-economic status belong education and age, factors used by Gallego (2010), who analyses over 52 000 individuals from 28 advanced democratic countries in order to understand the reasons of participation of people in a voting procedure. The main attempt is to explain what contributes to unequal voting turnout and why there exist any differences among countries. The author introduces a theory which says that socially disadvantaged people can under certain circumstances vote in a similar proportion as the advantaged ones, which is not an expected scenario because they are considered to be less informed about politics, less educated and consequently assumed to evince lower voting turnout. The theory says that if there is a strong left-wing political party that manages to motivate and mobilise these voters, it can result both in greater overall turnout and in lower turnout inequality. However, the empirical test does not evince this trend and the author suggests a re-examination of the idea.

The primary focus in terms of analysing unequal voting turnout is the impact of education because it is a good proxy of social status. The author finds a relatively strong influence of education on the voter turnout in the USA, in Germany and in the Czech Republic, which means that in these countries there is an augmented turnout inequality between highly-educated and poorly-educated people. On the other hand, other countries evince a similar turnout level for different social groups, which suggests there might be other important factors, for instance, a structure of a voting system or compulsory voting. The author finds out that voter turnout inequality is lower when the process is less complicated, i.e. fewer rules, fewer parties to choose from and the registrations initiated by the state. There is also the possibility of reducing inequality by establishing the compulsory voting system which somehow penalizes the people who abstain, but it is not a very popular solution for the society, especially in an advanced democracy.

Arzheimer (2012) is concerned with voters who support parties of the Extreme Right. He asks whether there is a social group whose members are more likely to cast votes for this kind of political parties and also why there is any variance in support of these parties among different countries in Western Europe. The author works with various analyses, using both aggregated and individual data, and pays attention to variables concerning gender, education, age, working class, immigration, unemployment, and religion. The results suggest that across Western Europe there is indeed a pattern in the social groups which vote for the parties of the Extreme Right. In particular, the typical voter of these political parties is identified as a youngish male with a moderate level of education and with a concern about immigration and immigrants. Despite the fact that specific political parties of the Extreme Right may sometimes manage to gain also the support of another social group, the above-mentioned type of voter was identified as a fundamental social group in their electorate.

Revenge through Ballot

Rodríguez-Pose (2018) points out a problem that appears in many countries across the world and is related to regions that are not performing well in economic terms and suffer from poverty and a lack of new possibilities to change their situation. If politicians focus on prosperous regions and do not care about

other places, the inhabitants of problematic regions feel isolated and overlooked. As a response, they tend to revenge the politicians through the voting ballot and cast their votes to populist parties which take the advantage of this situation. The author admits that the rise of populism was anticipated neither by politicians nor academics and argues that it might be caused by overlooking the economic potential of worse-performing areas and by concerning only the inequality on the interpersonal level instead of the inequality between the whole regions. The solution of this problem might be an improvement in quality of development policies for lagging areas which would contribute to utilize their potential and possibly help to reduce the unintended inequality between regions.

Regarding the worst-performing areas, Gabal & Víšek (2010) provide a study of socially-excluded areas in the Czech Republic. They note that these areas polarize the Czech society in economic, social, ethnic, and other aspects since they are mostly inhabited by Romany people and evince a low level of education and economic activity. The authors describe the origins of the socially-excluded areas and possible measures that should be taken in order to prevent their emergence. Related to this thesis, there are two complications. Firstly, there is a problem with counting of the Romany people that is described in Chapter 4, and secondly, the socially-excluded areas are usually a part of a municipality and therefore their influence might not be strong enough since the analysis works with aggregated data.

Algan et al. (2017) also look for possible causes of the recent rise of populism, extremism, and in general, the distrust in institutions in Europe. They analyse various types of political parties, for instance, anti-establishment, radical left, populist, etc., and look for an impact of the unemployment rate on their voting results. Using OLS, they find a positive effect of unemployment on anti-establishment and populist parties and the negative effect on the participation rate in the election. Another method, 2SLS, shows a positive effect on anti-establishment, radical left, populist and anti-European parties, meaning that regions with greater levels of unemployment are more likely to support these parties. Analysing the trust of people in various institutions, authors find negative influences of unemployment on trust in a national parliament, European Parliament and politicians in general, which probably has an interconnection with the previous analysis.

Guiso et al. (2017) ask whether the increased support of populist parties across Europe is caused by a shift of people's preferences and whether this shift could be related to (worsening) economic situation or crisis. They use analytical tools to examine the origins of the success of a populist party, which they define as "a party that champions short-term protection policies while hiding their long-term costs by using anti-elite rhetoric to manipulate beliefs." Concerning the demand side of the populism, i.e. people's preferences, and also the supply side, i.e. establishing new populist parties, they find out that there is a strong connection to the incentives related to participation in the election. Related to this, they conclude that economic insecurity, often related to the crisis, is a key determinant of increased support of populist parties, whose right-left orientation might primarily depend on accessible political space.

Ezrow et al. (2014) argue that if there is strong partisanship of voters towards established parties, the support of extreme parties is smaller, even if the electorate is polarized. However, they find out that the polarization of voters, in general, increases the probability of success for extreme parties as well as a greater number of new political parties or the fact that democracy is newly established in the given country. They also discuss the possibility that in democratic countries with a greater level of partisanship, the polarization of voters may not be caused by extremists getting more votes but traditional parties implementing more extreme policies. This claim might be important for the debate about the origin of extremism- whether it arises because of the decisions of political parties or because of people's preferences.

Spatial Analyses

Kouba (2007) uses spatial analysis in order to test whether there is a spatial effect influencing the election results in the Czech Republic. He works with aggregated data from 77 districts and the election results of 5 political parties in 1990-2006. He mentions that this kind of data is related to a problem called ecological fallacy which includes inference of individual behaviour based on aggregated data. On the other hand, he admits that recent improvements in the available methodology- spatial analyses- enable the use of this type of data, and in addition, his primary goal is to describe the regional variance in voting results rather than to explain individual voting behaviour. The author finds

out that the electoral results of political parties are relatively strongly related to certain regions, but simultaneously adds that it does not imply that it must be due to specific local factors. He analyses the regional context and the voting results and he concludes that in most cases there is no significant relationship between these variables.

Kotrč (2017) points out in his bachelor thesis that in the last years there is an increasing number of people in Western Europe who are not satisfied with established political systems and therefore right-wing extremist parties might receive greater support of voters. The author is concerned with the situation in Slovakia, in particular with the rise of the extreme right-wing party, led by Mr. Kotleba, which is now in the Slovak parliament. For the analysis, he uses economic, demographic and other data from almost 3000 municipalities and he finds out that there is a significant positive impact of the share of qualified unemployed people on the voting result of the given party. Another factor with a positive impact is, for instance, the share of young people. On the other hand, with an increasing percentage of older people or Hungarian inhabitants, the support of Kotleba's party decreases.

From the methodological point of view, the author uses an ordinary least squares model, however, in this case there is found significant evidence of spatial autocorrelation and therefore he subsequently uses a spatial error model. As weighting schemes author uses the first-order contiguity matrix and the inverse-distance matrix, which are 2927×2927 matrices representing binary variable for a shared border of two municipalities and inverse value of the distance between two municipalities in kilometres, respectively. The estimated coefficients of particular variables in the spatial error model are not substantially different from the ones from the OLS, nevertheless, they indicate a presence of the spatial effect which is considered by the author to be significant in this particular case of Kotleba's party. This work, monitoring overall population characteristics, might be a solid material for further analyses based also on individual-specific data.

Spatial analysis is provided also by Poole & Rosenthal (1984) who are concerned with the results of presidential elections in the US. They introduce new methodological approaches for this field and perform the analysis. They find out there are certain spatial effects since the distance of a candidate seems

to be more important than his or her attributes. As the main limitation of their work they consider the use of survey data and therefore they assume that the next step is to perform the analysis using aggregated data. Agnew (1996) writes about spatial effects in electoral results and asks whether they should be taken into account in electoral geography. He describes various methods and types of spatial effects that can be used. He analyses case studies from Italy and concludes that the context indeed matters in this field. However, it is challenging to determine a specific spatial effect and there is not a universal factor that would influence the voting results.

Chapter 3

Political Situation in the Czech Republic

This chapter is dedicated to a brief description of the political situation in the Czech Republic, regarding its development in time and focusing primarily on the parties which are considered to be extremist or populist. The goal is to provide an insight into the Czech political scene and find an argumentation that could be possibly used during the interpretation of results and finding conclusions.

History

After the Velvet Revolution in 1989, the communist regime in Czechoslovakia was substituted by the parliamentary democracy. A few decades under this regime enable to look back and consider the development of particular political parties and their voting results in parliamentary elections. Analysing these data, it is possible to say that there are major political subjects, such as CSSD, ODS, or KDU-CSL, that receive a sufficient number of votes for the whole time. This fact evinces stability of the voting results and simultaneously the voter partisanship which is mentioned by Ezrow et al. (2014) who say that it decreases the probability of success of extremist parties.

Novák (1996) is concerned with the parliamentary elections in 1996 and primarily with the future development of the democratic regime in the Czech Republic. He refers to a French political scientist Maurice Duverger who divides governments in European democracies into two groups, one is called decisive and the

other one powerless. The former is related to a political consensus within a country that enables an alternation of right-wing and left-wing parties without severe changes. The latter group lacks the consensus and therefore the government is based on parties close to the political centre which implies that parties further left or right remain in the opposition and the overall ability to act is smaller. The author argues that good results of CSSD and ODS in the parliamentary elections in 1996 suggest an orientation to the first group mentioned above, including a possible alternation of the parties. On the other hand, the elections, unfortunately, did not enable the creation of any decisive government and showed a significant polarization of voters, mainly due to sufficient support of extremist parties, i.e. the KSCM and the SPR-RSC, a nationalistic extreme-right party.

The KSCM is a specific example that also counts among stable parts of the Czech Parliament. Because of the fact that the Czech Republic is a former communist country whose history was significantly influenced by this regime, the analysis of current electoral support of the KSCM is a sensitive topic for society. Kubát (2006) writes about certain ambiguity related to this party since it is very often being discussed and analysed, nevertheless, its position within a political spectrum remains unspecified. He reflects on the position, using sociological theories, and concludes that there is empirical evidence which confirms that the KSCM is an anti-establishment political party. A similar opinion has also Pecková (2011) who analyses the KSCM and its particular members in detail and considers it to be an anti-system opposition party. In addition, she mentions that there are incessant disputes over its character and legitimacy since it follows the main party of the former communist regime, the KSC.

Recent Situation

Regarding the composition of the Czech Parliament, it is possible to say that new political subjects are able to attract voters since a greater diversification of the electoral support can be observed. In the past, there were on average 5 parliamentary parties, mostly including the previously-mentioned CSSD, ODS, KDU-CSL, and KSCM. Nevertheless, in 2013 there were 7 members of the parliament and the ascending trend was confirmed in 2017 when 9 parties managed to reach the minimum limit that guarantees the membership in the parliament.

The most significant change in the political scene was made by the party called ANO which was established in 2011 by A. Babiš and quickly became an important player. Another considerable person in Czech politics is T. Okamura who established and led two political subjects into the parliament. First, it was the party Dawn of Direct Democracy and then, after internal conflicts with other members, also the party SPD. Šárovec (2013) analyses the support of the ANO and the Dawn of Direct Democracy in 2013 and as the main determinants, he defines anti-establishment attitudes, strong leaders and the overall attempts to provide changes in the parliament. The other party of T. Okamura, the SPD, is also an anti-establishment subject that is known especially because of its negative attitude towards Islamism and refugees from the Middle East. The utilization of the extensive problem with migration in order to attract voters might resemble the claim of Guiso et al. (2017) who say that the orientation of populist and extremist parties is often based primarily on the available political space.

Concerning the last parliamentary elections in 2017, it is possible to notice an unequal distribution of electoral support among political parties since the difference between the winner and the second is approximately three times greater than the gap between the second and the last one. The previously-mentioned traditional parties also become parliament members, however, their results decline at the expense of the increased support of the anti-establishment parties ANO or SPD. A considerable voting result can be observed by the Pirate party which is assumed to attract younger people with a higher level of education. This trend can be found also in the analysis of the Czech organization Behavio which maps characteristics of voters based on a representative sample from the population in order to describe a typical voter of every political party.

Table 3.1: Election Results

Political party	ANO	ODS	Pirati	SPD	KSCM	CSSD	KDU-CSL	TOP 09	STAN
Share of votes	29.64 %	11.32 %	10.79 %	10.64 %	7.76 %	7.27 %	5.80 %	5.31 %	5.18 %
Number of votes	1 500 113	572 962	546 393	538 574	393 100	368 347	293 643	268 811	262 157
Mandates	78	25	22	22	15	15	10	7	6

A research Agency MEDIAN also provides analyses related to the parliamentary elections in 2017 and one of them is similar to the topic of this thesis since it works with a socio-demographic data on the municipal level and provides a graphical representation of particular variables with respect to the

voting results of political parties. The other report is, similarly to the organisation Behavio, based on a representative sample since it is concerned with more specific information about voters which cannot be collected for the whole population. The results of the econometric models used in this thesis should not be significantly different from the report of the MEDIAN using the data on the municipal level.

Chapter 4

Data

This chapter describes the data used in this bachelor thesis and introduces particular variables for the models. In the Czech Republic, there are over 6250 municipalities, however, the list of them slightly varies over the years as they arise, cease to exist, merge, etc. The analysis excludes municipalities Boletice, Brdy, Březina, Libavá, and Hradiště since they are military training areas and no inhabitants live there which means that the final dataset contains 6248 observations. There are no missing values and the variables used in the models include aggregated data, i.e. averages and shares of the population, from particular municipalities and this fact has to be taken into account while interpreting the results. All the variables, except the average age, are expressed as a number between 0 and 1. This type of data does not allow using the same methodology as for instance Gallego (2010) who works with individual data and can use binary variable "vote" or "did not vote" in her model explaining the voter turnout.

Parliamentary Election

In the Czech Republic, there are various types of elections, for instance, presidential, parliamentary, municipal, and others. Concerning the decision-making process of voters, every election has its specific attributes. Presidential candidates might be selected by their personal characteristics, and on the other hand, municipal representatives might be more supported by people who know them personally. An empirical analysis should work with election results that do not include a considerable electoral bias related to factors that are not quantifiable. Therefore, the parliamentary election was chosen for the analysis since

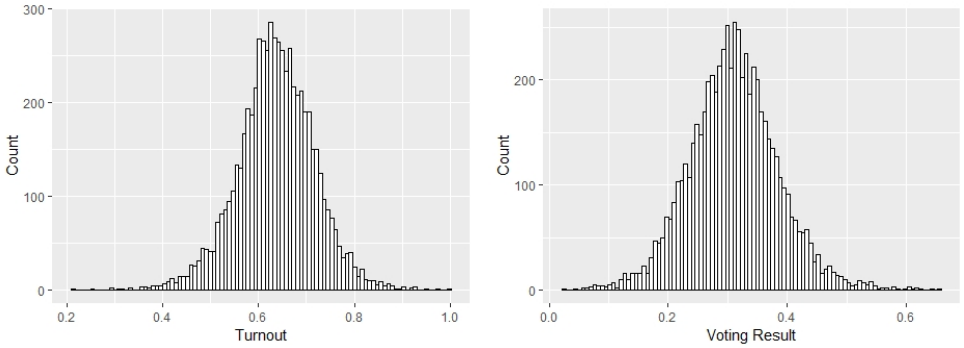
the vote for a specific political party is assumed to reflect the requirements and needs of the voters better than in the previously mentioned cases.

The results of the parliamentary election in 2017 provide information about the total number of eligible voters, people who actually voted and the number of votes for a political party within every electoral district. Since the district level is too specific and it is impossible to collect other information from this unit, the voting turnout and share of votes for particular parties are counted on the municipality level. In this election, there were 31 political parties who nominated at least one candidate and for the analysis, there were chosen 9 of them which received over 5% of total votes which means they reached a minimum limit for the entrance to the parliament.

The voter turnout in this parliamentary election was 60.84% and as the Figure 4.1 indicates, the distribution of voter turnouts is approximately normal as the most frequent values are situated close to the mean. The situation is similar also for voting results of particular parties, whose distributions of electoral support are also closed to normal and differ from each other mostly in the mean values. Several outliers can be observed in the data, however, they are usually based on voting results from the smallest villages. The probability that a party receives a majority of total votes, or that the voter turnout is extremely high, is obviously higher in the municipalities with only a few inhabitants. For instance, 23 people voted in the village called Těchařovice and the CSSD party received 73.9% of the votes. Another example is the village Čilá where all 17 eligible voters came to the elections, which resulted in the 100% voter turnout.

Independent Variables

All the data used for the analysis, including the results of the election and the voting turnout, have a specific attribute that has to be taken into account. The observations from particular municipalities are based on certain number of inhabitants which, as it is shown in Table 4.1, is not equally distributed. Because of this fact, there might be a problem with the ordinary least squares method since the statistic from a village influences the dependent variable in the same way as the statistic from a city. The city includes substantially more inhabitants and that might possibly lead to a bias in the results and their subsequent



(a) (b)

Figure 4.1: Distributions of Voter Turnouts and Voting Results of the ANO Party

interpretation with regards to the whole population. To prevent this problem, the weighted least squares method is used for the analysis because it enables to assign various weights to particular observations, in this case, according to the number of inhabitants within a municipality.

The average number of inhabitants living within a municipality in the Czech Republic is approximately 1698, however, there are many villages with a smaller number of inhabitants which implies that the graph of their distribution is strongly skewed to the right. Since there is a wide range between the majority of villages which have only hundreds of inhabitants and the capital city of Prague which has over one million inhabitants, displaying the distribution in the graph is not meaningful and therefore the table is provided instead.

Table 4.1: Distribution of Population

Inhabitants	<300	300−2000	2001−10 000	10 001−50 000	>50000	Total
Number of municipalities	2 294	3 254	569	113	18	6 248
Total inhabitants	391 772	2 463 413	2 257 817	2 285 090	3 210 430	10 608 522
Share of population	3.7 %	23.2 %	21.3 %	21.5 %	30.3 %	100 %

All the explanatory variables used in this analysis are collected on the municipal level and in the closest-possible time period to the date of parliamentary elections, i.e. October 2017. [1] Unfortunately, it is not possible to collect the data from more years in order to perform a panel data analysis. The vast majority of the data come from the website of the Czech Statistical Office and the

[1] The list of variables, including their description, date of collection, and source, is available in the Appendix B.

remaining part is from various, usually non-commercial, studies. In most cases it is possible to find data from 2017 or at least from January 2018, however, variables concerning education and religion are accessible only from the census in 2011.

Since the data about education are expected to play a significant in the analysis, it might be a problem, but on the other hand, this variable is not as volatile as, for instance, the unemployment and therefore it should not dramatically change during 6 years. Table 4.2 provides a basic summary of the explanatory variables used in the analysis. The values of mean and median are mentioned rather for an illustrative purpose since their true value is, obviously, not computed using aggregated data from the municipalities.

Table 4.2: Explanatory Variables

Statistic	Min	Median	Max	Mean	St. Dev.
unemployment	0.000	0.028	0.247	0.031	0.020
distraint	0.000	0.062	0.695	0.076	0.058
average age	31.670	41.938	65.167	42.079	2.681
men share	0.332	0.504	0.725	0.506	0.026
migration	−0.154	0.003	0.171	0.005	0.023
population growth	−0.143	0.000	0.045	−0.0003	0.010
no education	0.000	0.003	0.571	0.006	0.016
primary school	0.029	0.206	0.538	0.212	0.053
high school	0.101	0.401	0.635	0.399	0.057
university	0.000	0.067	0.324	0.074	0.040
believers	0.000	0.137	0.818	0.174	0.128
businessmen	0.018	0.116	0.500	0.119	0.035

Since the aim of the thesis is to explain voting behaviour using socio-economic variables, fixed effects of specific regions should be included in the model. Horká (2013) analyses 248 regions of the European Union in order to determine their level of socio-economic development. Focusing on the Czech Republic, four different degrees of development were identified. The capital city of Prague, which belongs to the most developed category, is considered to be an independent socio-economic region and since it is a single observation, it is not an appropriate candidate for a dummy variable.

Five regions were picked from the remaining categories, based on the level

of development: Central Bohemia, representing the average level, Southwest, Central Moravia, and Eastern Moravia, counting among the less developed, and Northwest, which is among the worst-performing regions. The assumption is that if there is an impact of these socio-economic areas on voting behaviour, the coefficients of their dummy variables should be ranked approximately according to their level determined by Horká (2013), i.e. Southwest, Central Moravia, and Eastern Moravia should be in the middle and Central Bohemia and Northwest should be situated on the opposite sides.

Concerning the independent variables, it is possible to display their values in maps, using a webpage *mapinseconds.com*. Figure 4.2 represents the unemployment rate in all 78 districts in the Czech Republic and it shows greater share of unemployed people in the northwest and other districts near borders.

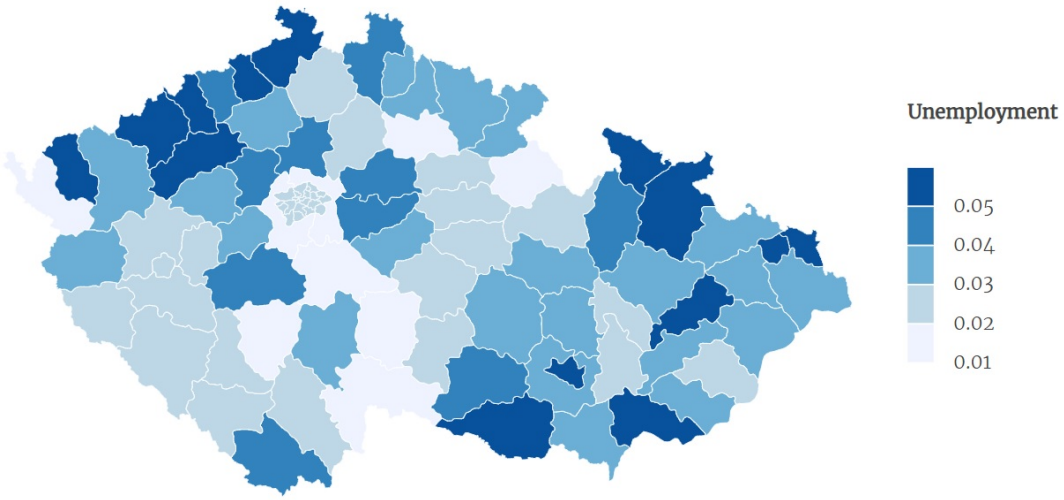


Figure 4.2: Unemployment Rate

A very similar pattern can be seen in Figure 4.3, which displays the share of people facing distraints. These maps might show which areas perform economically worse than the remaining part of the country. Furthemore, the same districts evince also lower share of people with university education, as it is shown in Figure 4.4. In the cities of Prague and Brno, there are almost 25% of people with university education, which is significantly more than other districts. Therefore, these statistics are not displayed in the map in order to make other differences observable. In this case, there is an evident impact of administrative centres since there are higher shares of university-educated people in major cities of particular regions.

The explanatory variables used in the models were chosen primarily based on

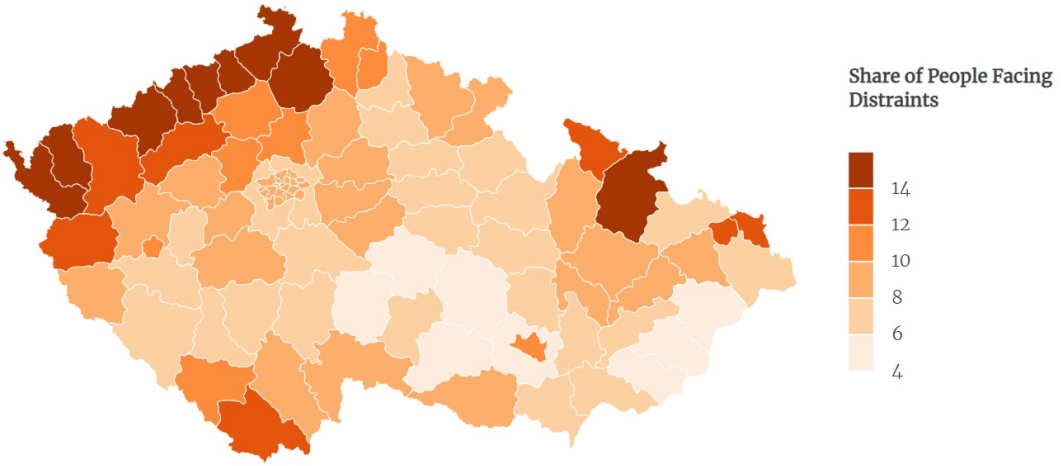


Figure 4.3: Share of People Facing Distraints

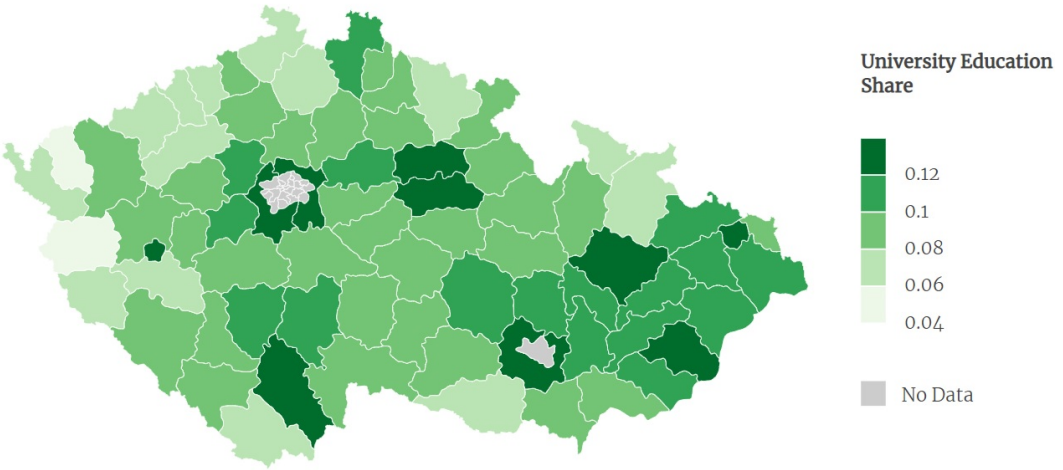


Figure 4.4: Share of People with University Education

existing literature, nevertheless, there were more variables which seemed to be potential candidates. Among these variables there were for instance data about municipal budgets, shares of Romany people, balances of municipal economic activities, dummy variables for primary school, kindergarten, healthcare institutions or utilities such as gas installation, etc. Regarding the share of Romany people within a municipality, there appears an error in the census from 2011 since the total number of inhabitants who reported this nationality was around 5000 and various studies estimate over 200 000 of Romany people in the Czech Republic. Other variables mentiones as the candidates turned out to be irrelevant in terms of the analysis.

In order to get an insight into the relevance of explanatory variables with

respect to the explained variable, the data were examined by means of data science, in this case, by using Lasso, Ridge and Elastic Net Regression. The results suggest that in the model with voting turnout there should be used variables concerning education, unemployment, age, share of believers and share of businessmen. Nevertheless, these methods, unfortunately, do not take into account the problem of different weights for particular observations and therefore they are, as well as the ordinary least squares method, not used while building the final version of models.

Weighting Matrix

Regarding the use of spatial econometrics, a weighting matrix for the analysis has to be introduced. The aim is to assign weights according to distances between all municipalities in the Czech Republic and observe whether there are certain spatial effects that cannot be explained by the explanatory variables. In this case, there is a 6248×6248 matrix that is symmetric and has zeros on the diagonal. The approach of spatial econometrics assumes that nearby villages or towns might evince similar results and vice versa which implies that smaller weights must be assigned to more distant municipalities. Therefore, the inverse values of the distances are used as weights and, for computational purposes, the matrix is row-standardized. [2] The initial distances are reported in kilometres and computed based on geographical coordinates of the municipalities.

The literature related to this topic works also with another type of matrix that is called the first-order contiguity matrix. This weighting scheme provides a matrix with binary values, one or zero, which determine whether particular municipalities share a border or not. Unfortunately, this kind of data about neighbours is not available in any dataset and the only way would be to estimate it by setting a distance that is sufficiently small to consider the municipalities to be the neighbours. Therefore, only the inverse-distance matrix is used in the analysis.

[2] Every element is divided by the sum of all elements within a corresponding row. This implies that every row of the matrix sums up to one.

Chapter 5

Methodology

This chapter describes a theory behind the models which are used in this thesis. It introduces possible reasons that might lead to the substitution of the ordinary least squares method by the weighted least squares and explains why the number of inhabitants within a municipality is used as the weight in this particular case. The other model, regarding spatial analysis, incorporates distances between municipalities in order to find out whether there are certain spatial effects in the results, i.e. whether people consistently support particular political party within specific regions.

Weighted Least Squares Model

The WLS method can be used instead of the OLS method because of various reasons. The most common case is the problem with heteroskedasticity. [1] Another reason for the use of this method might be an attempt to highlight a certain part of the data or to penalize observations that are expected to be measured imprecisely. However, it might be problematic to determine these values and to find corresponding weights that should be assigned.

Since this analysis works with aggregated data , the weights are assigned according to the population of particular municipalities in order to appropriately reflect the number of inhabitants from whom the statistic is computed. It implies that observation from a city influences the dependent variable more

[1] The variance of errors is not constant for all values of the independent variables. This situation leads to problems with standard errors and inefficient estimates of the regression. The WLS assigns greater weights to observations with a smaller variance, and vice versa, in order to compensate for the variance and through that correct for the heteroskedasticity.

than observation from a village. It is important because the main goal of the thesis is to describe voting behaviour of the people and the use of the weights improves the final interpretation of results in terms of the whole population. Nevertheless, it must be taken into account that the determination of weights has a significant impact on the results of the estimation process.

The model used in this analysis is represented by the equation

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + u, \tag{5.1}$$

where y denotes the dependent variable, β_0 stands for the intercept, $x_1, \ldots, x_k$ denote the independent variables with corresponding coefficients $\beta_1, \ldots, \beta_k$, and u represents the error term.

The WLS method is considered to be an extension of the OLS since it is usually used when the assumption regarding homoskedasticity is violated. The list of the assumptions for the OLS method is available in the Appendix A.

According to the Gauss-Markov theorem, the OLS is a best linear unbiased estimator when the assumptions MLR.1 through MLR.5 are satisfied. If the heteroskedasticity is detected, i.e. the assumption MLR.5 is violated, the estimates are still unbiased. The most important assumption is MLR.4 regarding zero conditional mean which states that the expected value of the error term u, given all the explanatory variables, is zero:

$$\mathbb{E}(u|x_1, \ldots, x_k) = 0. \tag{5.2}$$

This assumption is violated if there is a correlation between the error term and any of the explanatory variables.

In order to find $\hat{\beta}_0, \ldots, \hat{\beta}_k$, the weighted sum of squares is minimized:

$$WSS(\boldsymbol{\beta}, \mathbf{w}) = \sum_{i=1}^{n} w_i (y_i - x_i \cdot \boldsymbol{\beta})^2. \tag{5.3}$$

This process is similar as in the case of the OLS, the only difference is that the OLS uses $w = 1$, which might suggest, that it is in fact the special case of the weighted least squares method. The estimation process provides the equation:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k. \tag{5.4}$$

The estimates and standard errors obtained by using the WLS are different from the ones from the OLS, nevertheless, their interpretation is the same, i.e. the ceteris paribus effect. From the previous equation, it is possible to express Δy (change in y) as function of $\Delta x_1, \ldots, \Delta x_k$ (change in $x_1, \ldots, x_k$):

$$\Delta\hat{y} = \hat{\beta}_1 \Delta x_1 + \hat{\beta}_2 \Delta x_2 + \ldots + \hat{\beta}_k \Delta x_k. \tag{5.5}$$

Holding $x_2, \ldots, x_k$ fixed, the β_1 is interpreted as:

$$\hat{\beta}_1 = \frac{\Delta\hat{y}}{\Delta x_1}. \tag{5.6}$$

Spatial Error Model

After using the weighted least squares method for estimation of the results, statistical tests for the presence of spatial autocorrelation are performed. Based on these tests that are available in Chapter 6, the spatial error model, which incorporates spatial effects of the error terms, is chosen for the analysis. The test statistics for the second option, the spatial lag model, are not significant. Another method that was considered for this purpose is called the Spatial Durbin Model, which incorporates the influence of independent and dependent variables from nearby spatial units. However, Glass et al. (2012) describe this method as an extension of the spatial lag model and therefore it is not used for the analysis. In addition, it provides coefficients that cannot be interpreted in the same way as in the case of WLS, which would make a comparison between these methods impossible.

Since the inverse-distance matrix is being used in this model, it is convenient to switch to matrix notation. The model is similar to the previous case, the only difference is in the error term u:

$$\begin{aligned} y &= \mathbf{X}\beta + u \\ u &= \lambda \mathbf{W} u + \epsilon, \end{aligned}$$

where y denotes an $n \times 1$ vector of the dependent variable, $\mathbf{X}$ stands for a $n \times (k+1)$ vector of independent variables, β denotes a $(k+1) \times 1$ vector of

parameters, and u is an $n \times 1$ vector representing the error terms. $\mathbf{W}$ denotes the inverse-distance matrix introduced in Chapter 4, λ denotes vector expressing the correlation of residuals and ϵ is the error term satisying $\epsilon \sim N(0, \sigma^2\mathbf{I})$, where $\mathbf{I}$ is an identity matrix.

The error term u can be expressed as:

$$\begin{aligned} u &= \lambda\mathbf{W}u + \epsilon \\ u - \lambda\mathbf{W}u &= \epsilon \\ u(\mathbf{I} - \lambda\mathbf{W}) &= \epsilon \\ u &= (\mathbf{I} - \lambda\mathbf{W})^{-1}\epsilon, \end{aligned}$$

which enables to describe the data generating process as:

$$y = \mathbf{X}\beta + (\mathbf{I} - \lambda\mathbf{W})^{-1}\epsilon. \tag{5.7}$$

There are various methods that can be used for estimation of the parameters. For this analysis, the maximum likelihood estimation is used. LeSage & Pace (2014) describe the log-likelihood function as:

$$ln\ L = -\frac{n}{2}ln(\pi\sigma^2) + ln|\mathbf{I} - \lambda\mathbf{W}| - \frac{e^T e}{2\sigma^2}, \tag{5.8}$$

where $e = (\mathbf{I} - \lambda\mathbf{W})(y - \mathbf{X}\beta)$. It is possible to rewrite (5.8) as:

$$ln\ L = -\frac{n}{2}ln(\pi\sigma^2) + ln|\mathbf{I} - \lambda\mathbf{W}| - \frac{1}{2\sigma^2}(\mathbf{I} - \lambda\mathbf{W})^T(y - \mathbf{X}\beta)^T(\mathbf{I} - \lambda\mathbf{W})(y - \mathbf{X}\beta). \tag{5.9}$$

Since the function contains the expression $|\mathbf{I}-\lambda\mathbf{W}|$, finding maximum likelihood estimators from this equation would require advanced computational methods. Therefore, a simplification provided by Ord (1975) is used in the estimation process. If it is possible to find eigenvalues $\omega_1, \ldots, \omega_n$ of matrix $\mathbf{W}$, then:

$$|\omega\mathbf{I} - \lambda\mathbf{W}| = \prod_{i=1}^{n}(\omega - \lambda\omega_i). \tag{5.10}$$

If the ω is set to 1, then it is possible to express the determinant in terms of eigenvalues:

$$|\mathbf{I} - \lambda\mathbf{W}| = \prod_{i=1}^{n}(1 - \lambda\omega_i), \tag{5.11}$$

substitute it in (5.9), and maximize the log-likelihood function in order to obtain the ML estimators.

The interpretation of β coefficients is the same as in the case of weighted least squares, which enables a good comparison of the methods.

Chapter 6

Results

This chapter introduces particular econometric models and interprets their results with respect to the hypotheses provided by the existing literature in this field. The main goal of the thesis is to describe the voting behaviour as a whole and therefore both the analysis of voter turnout and the analysis of voting support of political parties are included. The list of explanatory variables is the same for both categories of the models. After the interpretation of WLS results, the data are subject to a test for the presence of spatial autocorrelation. Subsequently, a spatial model is introduced and used for the same variables. The results of the model are interpreted and compared to the WLS. All the operations performed in this thesis such as processing of data, estimating the models, exporting figures, etc., are performed in a statistical software R.

Weighted Least Squares Results

The model related to voter turnout examines whether there are certain connections between the variables concerning socio-economic characteristics of people and their decision to participate in the elections. According to Gallego (2010), in the Czech Republic, there should be observed a significant impact of education on the voter turnout, compared to the majority of other countries where the effect is ambiguous. Regarding other variables, the unemployment rate and the share of people facing distraints are expected to have a negative impact on the voter turnout since they might be associated rather to people from a lower social class who probably do not vote so often. The migration and the population growth are assumed to be positively connected to the overall standard of living within the municipality. People tend to move into more perspective

areas and have more children if they are financially secure, and therefore a positive effect on the voter turnout is anticipated. The potential influence of the share of men and the average age is ambiguous because the vast majority of observations are concentrated around the mean value and do not vary a lot.

Regarding the models which use the results of political parties as the explained variables, the expected effects of certain variables are suggested in Chapter 2. The main focus of the thesis is to analyse particular variables and determine which of them play a significant role in terms of the voting results. Concerning specific parties, the support of extremists and populists is the primary area of interest since there are many academic papers dedicated to the recent rise of these political subjects across Western Europe. Municipalities with greater unemployment rate, higher share of people facing distraints and negative balance of migration are assumed to be the main supporters of extremist parties since Rodríguez-Pose (2018) argues that the inhabitants of worse-performing areas might feel isolated and tend to revenge the politicians through the voting ballot.

Among the variables, there is no perfect collinearity and the correlation coefficients mostly fluctuate between 0 and .2, the highest correlation being found between the share of university educated people and shares of people with primary or high school education.[1] In order to determine whether any correlation coefficient is sufficiently large to cause problems in the model, it is possible to use the method called variance inflation factor which determines the increase in variance because of the collinearity. Using all variables concerning the level of education, the value of the VIF for the share of people with university education significantly exceeds the threshold which identifies the suitability of particular variables for the model. This problem might be solved by excluding the share of people with university education from the model, however, this variable provides an important piece of information in terms of the interpretation of voting behaviour, and therefore the other option is used, i.e. excluding shares of people with primary or high school education from the model.

Concerning the coefficient of determination, R-squared, it is necessary to mention Willet & Singer (1988) who describe a pitfall that appears in the case of the weighted least squares method. The R-squared may be significantly higher than in the situation when the OLS is used. Nevertheless, the increase does

[1]The correlation table of all variables is available in the Appendix C.

Table 6.1: Correlation Table of Variables Related to Education

	no education	primary school	high school	university
no education	1	0.098 ***	-0.128 ***	-0.118 ***
primary education	0.098 ***	1	0.159 ***	-0.577 ***
high school	-0.128 ***	0.159 ***	1	-0.616 ***
university	-0.118 ***	-0.577 ***	-0.616 ***	1

Note: *p<0.1; **p<0.05; ***p<0.01

not mean a considerable increase in the goodness of fit since it is rather caused by solving the problem with the heteroskedasticity. This fact needs to be taken into account while interpreting the output from the estimation processes.

Regarding the heteroscedasticity, all the models are examined by the Breusch-Pagan test, whose null hypothesis states that there is a constant variance of errors for all values of the independent variables. The p-values of all the tests are sufficiently small to reject the null hypothesis that assumes homoskedasticity which implies that there is a certain form of heteroskedasticity in the models. During the interpretation of results, a possible problem with endogeneity needs to be taken into account. The variables used in the analyses might not be exogenous and therefore the estimated coefficients should be interpreted in the form of a correlation rather than as a causal relationship between the variables.

Table 6.2 and Table 6.3 provide results of the estimation processes that were computed using the weighted least squares method. The description of particular variables is available in Appendix B.

Table 6.2: Weighted Least Squares Results

	Dependent variable:				
	turnout	ANO	ODS	Pirati	SPD
	(1)	(2)	(3)	(4)	(5)
unemployment	−0.533***	0.155***	−0.162***	−0.286***	0.329***
	(0.033)	(0.033)	(0.023)	(0.018)	(0.020)
distraint	−0.613***	−0.023*	−0.052***	0.006	0.038***
	(0.014)	(0.014)	(0.009)	(0.008)	(0.008)
average age	−0.0002	0.004***	−0.001***	−0.001***	−0.001***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
men share	0.353***	0.079**	−0.126***	−0.144***	−0.009
	(0.036)	(0.036)	(0.024)	(0.020)	(0.022)
migration	0.292***	0.058	0.043	−0.014	−0.009
	(0.043)	(0.042)	(0.029)	(0.024)	(0.026)
population growth	0.351***	0.258**	−0.095	0.038	0.044
	(0.113)	(0.112)	(0.077)	(0.062)	(0.070)
no education	−0.663***	−0.120**	−0.082**	−0.036	−0.006
	(0.057)	(0.057)	(0.039)	(0.032)	(0.035)
university	0.184***	−0.248***	0.263***	0.113***	−0.146***
	(0.014)	(0.014)	(0.010)	(0.008)	(0.009)
believers	0.046***	−0.149***	−0.039***	−0.066***	−0.039***
	(0.006)	(0.006)	(0.004)	(0.003)	(0.004)
businessmen	0.265***	−0.526***	0.203***	0.390***	−0.213***
	(0.018)	(0.018)	(0.013)	(0.010)	(0.011)
central bohemia	0.018***	−0.007***	0.015***	0.008***	−0.017***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
southwest	−0.007***	−0.003*	0.012***	−0.004***	−0.004***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
central moravia	−0.012***	0.011***	−0.010***	−0.009***	0.027***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
eastern moravia	−0.024***	0.051***	−0.023***	0.0003	0.022***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
northwest	−0.007***	0.041***	0.003**	−0.004***	−0.004***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
constant	0.459***	0.204***	0.193***	0.181***	0.194***
	(0.025)	(0.025)	(0.017)	(0.014)	(0.015)
Observations	6,248	6,248	6,248	6,248	6,248
R^2	0.729	0.696	0.651	0.721	0.605
Adjusted R^2	0.728	0.695	0.651	0.720	0.604
Residual Std. Error (df = 6232)	1.403	1.394	0.956	0.777	0.869
F Statistic (df = 15; 6232)	1,118.312***	951.716***	776.602***	1,074.544***	635.515***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6.3: Weighted Least Squares Results - Other Political Parties

	Dependent variable:				
	KSCM	CSSD	KDU.CSL	TOP.09	STAN
	(1)	(2)	(3)	(4)	(5)
unemployment	0.296***	0.077***	−0.088***	−0.150***	−0.143***
	(0.022)	(0.018)	(0.020)	(0.014)	(0.030)
distraint	−0.011	0.003	−0.032***	0.037***	0.014
	(0.009)	(0.008)	(0.008)	(0.006)	(0.013)
average age	0.002***	0.003***	−0.001***	−0.001***	−0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0003)
men share	0.229***	−0.081***	0.024	−0.003	0.020
	(0.023)	(0.020)	(0.022)	(0.015)	(0.033)
migration	−0.025	−0.079***	−0.099***	0.023	0.082**
	(0.028)	(0.024)	(0.026)	(0.018)	(0.039)
population growth	−0.159**	−0.017	0.030	−0.019	−0.068
	(0.074)	(0.063)	(0.068)	(0.047)	(0.104)
no education	0.066*	−0.060*	0.118***	0.037	0.033
	(0.037)	(0.032)	(0.034)	(0.024)	(0.052)
university	−0.186***	0.015*	0.071***	0.204***	−0.144***
	(0.009)	(0.008)	(0.009)	(0.006)	(0.013)
believers	−0.051***	0.043***	0.374***	−0.015***	−0.038***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.006)
businessmen	−0.121***	−0.184***	−0.094***	0.382***	0.135***
	(0.012)	(0.010)	(0.011)	(0.008)	(0.017)
central bohemia	−0.007***	−0.003***	−0.013***	0.008***	0.022***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
southwest	0.008***	0.001	−0.008***	0.004***	−0.010***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
central moravia	−0.001	−0.012***	0.002*	−0.010***	0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
eastern moravia	−0.001	0.003***	−0.016***	−0.009***	−0.020***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
northwest	−0.007***	−0.010***	−0.008***	0.004***	−0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
constant	−0.072***	0.021	0.047***	0.035***	0.104***
	(0.016)	(0.014)	(0.015)	(0.010)	(0.023)
Observations	6,248	6,248	6,248	6,248	6,248
R^2	0.460	0.273	0.765	0.834	0.195
Adjusted R^2	0.459	0.271	0.765	0.834	0.193
Residual Std. Error (df = 6232)	0.917	0.783	0.847	0.590	1.290
F Statistic (df = 15; 6232)	354.227***	155.805***	1,355.140***	2,087.720***	100.633***

Note: *p<0.1; **p<0.05; ***p<0.01

The results of Model (1) mostly correspond to the previous expectations. The unemployment rate and the share of people facing distraints are both negatively connected to the voter turnout since a 1% higher unemployment rate is associated with a .53% lower turnout and in the case of the distraints it is even .61%. The variables concerning the proportional change of population are related to the explained variable in a positive way, as well as the share of entrepreneurs. The hypothesis given by Gallego (2010) is in accordance with the estimated coefficients since 1% greater share of people with no education is related to .66% lower turnout and, on the other hand, the turnout has a moderate positive relationship with the share of people with university education.

Dummy variables expressing a fixed effect of particular regions almost follow the assumption that they should be ordered according to their degree of socioeconomic development, the only exception is the Northwest which has just a small negative coefficient despite the fact that it is the least developed region. Negligible effects can be observed also by the average age and the share of religious people. Surprisingly, the municipalities with a 1% greater share of men evince .35% higher voter turnout which was not anticipated since this variable does not vary a lot across the municipalities.

Results of other models naturally vary in the coefficients since particular parties are assumed to attract different groups of people. Nevertheless, an unexpected gap is observed in the coefficient of determination. Despite the fact that it might be artificially increased by the use of weighted least squares, the R-squared in the case of the parties CSSD and STAN are quite low, compared to other cases, which signifies that the variables used in the analysis do not explain the voting support of these parties very well. A significant contrast between Model (1) and other models is observed in the share of people facing distraints. This variable was strongly related to the turnout, however, in the case of political parties, the coefficients are close to zero and some of them are not statistically significant even on a 90% level. This finding might imply that people in difficult financial situation, including for instance the distraint(s), do not participate in the voting procedure. Negligible effects appear also by the men share with the exceptions of moderate negative relationships found by the ODS and the Pirati which suggest that they receive slightly greater support in municipalities with greater shares of women.

The coefficients on the unemployment rate suggest an anticipated differentiation of political spectrum since the parties ODS, TOP 09, and Pirati, which are considered to be liberal or right-wing oriented, evince a negative relationship with the unemployment. On the other hand, the opposite effect is observed by the parties ANO, SPD, and KSCM, which, incidentally, negotiated in the process of establishing a coalition and the government. This division can be noticed also in terms of the education since the right-wing or liberal parties are associated with greater support of people with university education and the situation is opposite for the other group. The share of entrepreneurs is also an important determinant since its coefficients are relatively high, for instance, a 1% higher share is associated with .52% lower voting result for the winner of the election, the party ANO. Moreover, the signs and magnitudes of the coefficients tend to be reversed for the above-mentioned groups of parties and this fact supports the argument related to the differentiation of the political scene.

Unfortunately, there are no available data regarding, for instance, the number of manually-working people or the number of people working in agriculture or other fields. This data would probably provide an interesting insight into the problem. Another shortage of this analysis is reflected in the average age since its coefficients are relatively close to zero while other analyses working with individual data usually consider the age to be a significant factor. The share of believers has an anticipated positive connection to the voting result of the KDU-CSL and ambiguous effects in other cases, which indicates a lower importance of the religion in the 21^{st} century.

Testing the Data

The academic literature related to a similar topic as this thesis provides many analyses that work with spatial econometrics. There might be a spatial effect also in the results of political parties or in voter turnout and therefore the data are subject to a test for spatial autocorrelation. Chen (2015) summarizes various approaches to this testing and suitability of the methods for a specific type of data. He describes the Durbin-Watson test that is used in order to detect serial correlation and states that it might be used only for ordered time series or ordered spatial series since the sequences of data points are defined. This implies that in the case of this thesis, which works with cross-sectional data, the procedure is ineffective and another method has to be used, in par-

ticular, the Moran's index. Table 6.2 displays the results of this test which provides general information about the presence of spatial autocorrelation. For this purpose, there are used the models regarding voter turnout and the political parties ANO and SPD along with the inverse-distance matrix that is introduced in Chapter 5.

Table 6.4: Statistical Tests for Presence of Spatial Autocorrelation and for Use of Spatial Error Model or Spatial Lag Model

turnout model			
TEST	statistic	df	p-value
Moran's index	86.28	1	0.000
LM-ERR	28 003	1	0.000
RLM-ERR	27 949	1	0.000
LM-LAG	692.340	1	0.000
RLM-LAG	638.010	1	0.000
ANO model			
TEST	statistic	df	p-value
Moran's index	49.94	1	0.000
LM-ERR	902.32	1	0.000
RLM-ERR	903.17	1	0.000
LM-LAG	0.890	1	0.345
RLM-LAG	1.736	1	0.188
SPD model			
TEST	statistic	df	p-value
Moran's index	39.397	1	0.000
LM-ERR	573.350	1	0.000
RLM-ERR	574.040	1	0.000
LM-LAG	0.159	1	0.691
RLM-LAG	0.8531	1	0.356

The Moran's indices provide enough evidence to reject the null hypothesis assuming zero spatial autocorrelation in all the models. This means that the data should be subsequently analysed also by means of spatial econometrics. In order to choose an appropriate model, Lagrange multiplier tests LM-ERR, LM-LAG, and their robust alternatives, RLM-ERR, and RLM-LAG are performed in order to decide for the use of spatial error model or spatial lag model, respectively. Regarding the results, the spatial error model will be used in further analysis since the p-values of the spatial lag tests are statistically insignificant in the case of political parties. Comparing the methods, Glass et al. (2012) mention

that the spatial error model has an advantage over the spatial lag model. The coefficients estimated by the SEM express direct effects and therefore enable the same interpretation as in cases of OLS or WLS estimations.

Spatial Error Model Results

The spatial error model is applied for the same set of independent variables as in the case of weighted least squares. Since the aim of this Chapter is to compare results of different methods, the same weights as in the case of WLS are used, i.e. the weights for particular observations are set according to the number of inhabitants within a municipality. The results of estimation processes are available in Table 6.5 and Table 6.6 on the following pages. Regarding the estimated coefficients, it is possible to say that the results do not significantly differ from the results of the WLS. This is in accordance with academic literature, which states that an appropriately-chosen spatial error model should not provide any estimates that are considerably different from the least squares method.

In the model concerning voter turnout, magnitudes of the coefficients are similar to the WLS results. The argument regarding the positive effect of education on the voter turnout in the Czech Republic, provided by Gallego (2010), is thus confirmed also in this model. As it was previously explained in this Chapter, other variables related to education cannot be included because of multicollinearity. Therefore, it is not possible to observe whether the coefficients on these variables are ordered according to a level of education. However, the available results evince the expected trend that educated people more likely take part in the voting procedure and it is possible to conclude that there is a significant connection between the education and the voter turnout. The coefficients for the university degree and no education are .196, and -.643, respectively.

The models regarding voting support of political parties also evince similar results as the WLS. The imaginary barrier between ANO, SPD, KSCM, and the other group including ODS, Pirati, and TOP 09, persists also in the SEM. The former group receives greater support in municipalities with lower share of university-educated people, greater unemployment, and lower share of businessmen. The latter group is characterized by the exact opposite results and this

Table 6.5: Spatial Error Model Results

	Dependent variable:				
	turnout	ANO	ODS	Pirati	SPD
	(1)	(2)	(3)	(4)	(5)
unemployment	−0.523***	0.173***	−0.159***	−0.244***	0.279***
	(0.034)	(0.033)	(0.023)	(0.018)	(0.020)
distraint	−0.597***	−0.021	−0.054***	0.007	0.035***
	(0.014)	(0.013)	(0.009)	(0.007)	(0.008)
average age	−0.0003	0.004***	−0.001***	−0.001***	−0.001***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
men share	0.350***	0.079**	−0.111***	−0.152***	0.005
	(0.035)	(0.035)	(0.024)	(0.019)	(0.021)
migration	0.313***	0.065	0.048*	−0.017	−0.017
	(0.042)	(0.041)	(0.029)	(0.023)	(0.025)
population growth	0.341***	0.236**	−0.094	0.045	0.051
	(0.110)	(0.109)	(0.076)	(0.060)	(0.067)
no education	−0.643***	−0.123**	−0.071*	−0.035	−0.013
	(0.056)	(0.055)	(0.038)	(0.030)	(0.034)
university	0.196***	−0.238***	0.262***	0.132***	−0.176***
	(0.015)	(0.015)	(0.010)	(0.008)	(0.009)
believers	0.069***	−0.140***	−0.046***	−0.044***	−0.068***
	(0.007)	(0.007)	(0.005)	(0.004)	(0.004)
businessmen	0.267***	−0.515***	0.193***	0.333***	−0.158***
	(0.020)	(0.020)	(0.014)	(0.011)	(0.012)
central bohemia	0.020***	−0.005**	0.018***	0.001	−0.011***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
southwest	−0.005**	−0.0003	0.015***	−0.007***	−0.003**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
central moravia	−0.005***	0.014***	−0.009***	−0.009***	0.023***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
eastern moravia	−0.018***	0.050***	−0.022***	−0.002**	0.021***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
northwest	−0.004*	0.044***	0.004***	−0.007***	−0.002
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Constant	−1.099***	0.170***	0.266***	0.820***	0.371***
	(0.050)	(0.066)	(0.020)	(0.095)	(0.115)
Observations	6,248	6,248	6,248	6,248	6,248
Log Likelihood	8,467.511	8,542.751	10,811.520	12,249.150	11,592.580
σ^2	1.889	1.843	0.892	0.563	0.694
Akaike Inf. Crit.	−16,899.020	−17,049.500	−21,587.040	−24,462.310	−23,149.150
Wald Test (df = 1)	12,367.760***	22,447.020***	2,434.225***	154,288.400***	183,197.600***
LR Test (df = 1)	223.492***	291.486***	115.649***	396.582***	488.226***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6.6: Spatial Error Model Results - Other Political Parties

	Dependent variable:				
	KSCM	CSSD	KDU.CSL	TOP.09	STAN
	(1)	(2)	(3)	(4)	(5)
unemployment	0.203***	0.050***	−0.081***	−0.132***	−0.067**
	(0.021)	(0.019)	(0.020)	(0.014)	(0.029)
distraint	−0.004	0.006	−0.024***	0.038***	−0.003
	(0.009)	(0.007)	(0.008)	(0.006)	(0.012)
average age	0.002***	0.002***	−0.001***	−0.001***	−0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0003)
men share	0.219***	−0.088***	0.024	−0.014	0.022
	(0.022)	(0.019)	(0.021)	(0.015)	(0.030)
migration	−0.032	−0.082***	−0.089***	0.021	0.080**
	(0.026)	(0.023)	(0.025)	(0.017)	(0.036)
population growth	−0.156**	−0.011	0.027	−0.016	−0.067
	(0.069)	(0.061)	(0.066)	(0.046)	(0.096)
no education	0.049	−0.060**	0.123***	0.036	0.042
	(0.035)	(0.031)	(0.033)	(0.023)	(0.048)
university	−0.216***	−0.005	0.063***	0.217***	−0.093***
	(0.009)	(0.008)	(0.009)	(0.006)	(0.013)
believers	−0.082***	0.022***	0.377***	0.006**	−0.001
	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)
businessmen	−0.091***	−0.155***	−0.061***	0.339***	0.081***
	(0.012)	(0.011)	(0.012)	(0.008)	(0.017)
central bohemia	−0.009***	−0.003***	−0.010***	0.002**	0.021***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
southwest	0.005***	−0.002	−0.004***	0.001	−0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
central moravia	−0.002	−0.010***	−0.0003	−0.010***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
eastern moravia	0.0005	0.006***	−0.015***	−0.012***	−0.020***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
northwest	−0.009***	−0.010***	−0.005***	0.001	−0.018***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Constant	−1.326***	−0.160**	0.050	0.333***	−0.223
	(0.119)	(0.064)	(0.048)	(0.071)	(0.259)
Observations	6,248	6,248	6,248	6,248	6,248
Log Likelihood	11,351.440	12,202.920	11,665.360	13,908.860	9,345.036
σ^2	0.750	0.571	0.678	0.331	1.425
Akaike Inf. Crit.	−22,666.880	−24,369.840	−23,294.720	−27,781.720	−18,654.070
Wald Test (df = 1)	182,724.800***	69,920.760***	33,488.680***	146,153.300***	449,414.100***
LR Test (df = 1)	672.857***	399.644***	312.410***	280.581***	926.260***

Note: *p<0.1; **p<0.05; ***p<0.01

differentiation can be observed also in the fixed effects of specific regions. Central Bohemia, which is denoted as a more developed region in socio-economic terms, indicates increased support of right-wing or liberal parties, and vice versa for the parties ANO, SPD, and KSCM. On the other hand, Eastern Moravia is considered to be a less developed region and it evinces the opposite trends than Central Bohemia. In this case, the coefficients are relatively considerable since the support of parties ANO and SPD in Eastern Moravia are greater by 5%, and 2.1%, respectively.

The region of Eastern Moravia might be associated with lower quality of life due to heavy industry, air pollution, lower wages and other factors that indicate a worse-performing area. Rodríguez-Pose (2018) argues that inhabitants of these places tend to cast more votes for populist and anti-establishment parties. Even though there are definitions of populism, it is not easy to determine whether a political party satisfies the criteria. Nevertheless, both ANO and SPD might be sometimes considered to belong among the populists since they are led by distinctive leaders who often speak out against traditional politicians and promise a significant improvement in the political situation as Algan et al. (2017) describe in their paper. With respect to the fact that ANO and SPD receive greater voting support in less developed regions, the hypothesis provided by the literature seems to be in accordance with the results of this analysis.

The parties CSSD and STAN do not seem to be well explained by the chosen set of explanatory variables since their coefficients are negligible in terms of real-life interpretation. Similarly, the KDU-CSL party evinces undistinguished coefficients with the exception of the share of believers. As the summary of variables from Chapter 4 suggests, the religion slowly dwindles in importance in society. Especially if it is compared to the situation in the 1950s in the United States that is analysed by Lazarsfeld (1944) who considers the religion to be one of the significant factors influencing the electoral choice.

Chapter 7

Conclusion

The thesis provides analyses of the voter turnout and the voting support of political parties in the Czech political scene. The most significant variables in the case of voter turnout are unemployment, the share of people facing distraints and the variables related to education. The results of particular parties evince greater instability, and among the important variables are: unemployment, the number of people with a university degree and the number of businessmen. In general, the variables related to professions seem to be an important factor in this type of analysis since they significantly influence the social group of individuals. Unfortunately, this kind of data is mostly not available on the municipal level.

Regarding the hypotheses provided by the literature, it is possible to state that they are mostly in accordance with the results of this analysis. There was found a strong connection between the education and the voter turnout which corresponds to the statement of Gallego (2010). In the results, there was found a voting pattern since the parties ANO, SPD, and KSCM mostly evince the opposite trends than the parties ODS, Pirati, and TOP 09. The former group might be considered to be anti-establishment or populist and the results show that it is supported more in places with greater unemployment, which is the same relationship as in the case of Algan et al. (2017). Also, the argument of Rodríguez-Pose (2018), who says that people from worse-performing areas tend to support extremists and populists, is supported by the results. Concerning the methods used in the thesis, the weighted least squares provide similar results as the spatial error model which corresponds to the assumptions of the related literature.

In the Czech Republic, there was performed a spatial analysis of voting results by Kouba (2007), who worked with aggregated data on a regional level. This thesis provides an actual and more detailed analysis since the data are available on the municipal level. Spatial econometrics seems to be an appropriate tool for the analysis of voting results and further work in this field is significantly related to the availability of data. As it was previously mentioned, primarily the data defining the social status of people might be very useful. In 2021, there should be both the parliamentary elections and the national census which is a convenient situation in terms of further studying of the voting results.

Bibliography

AGNEW, J. (1996): "Mapping politics: how context counts in electoral geography." *Political Geography* 15(2): pp. 129-146.

ALGAN, Y., S. GURIEV, E. PAPAIOANNOU & E. PASSARI, (2017): "The European Trust Crisis and the Rise of Populism." BPEA Fall Conference.

ANTUNES, R. (2010): "Theoretical models of voting behaviour." *Escola Superior de Educacão - Instituto Politécnico de Coimbra* 4: pp. 145-170.

EZROW, L., M. TAVITS & J. HOMOLA (2014): "Voter Polarization, Strength of Partisanship, and Support for Extremist Parties." *Comparative Political Studies* 47: pp. 1558-1583.

GALLEGO, A. (2010): "Understanding unequal turnout: education and voting in comparative perspective." *Electoral Studies* 29(2): pp. 239-247.

GLASS, A. J., K. KENJEGALIEVA & R. SICKLES (2012): "The Economic Case for the Spatial Error Model with an Application to State Vehicle Usage in the U.S."*Department of Economics, Rice University, Houston.*

GUISO, L., H. HERRERA, M. MORELLI & T. SONNO (2017): "The spread of populism in Western countries." *Einaudi Institute for Economics and Finance.*

HAYEK, F. A. (1945): "The Use of Knowledge in Society." *The American Economic Review* 35(4): pp. 519-530.

HORKÁ, L. (2013): "Factors of Socioeconomic Development of Regions: An Alternative Approach to an Assessment of Regions" Departement of Regional Economics and Administration, Faculty of Economics and Administration.

CHEN, Y. (2015): "Spatial Autocorrelation Approaches to Testing Residuals from Least Squares Regression." Department of Geography, College of Urban and Environmental Sciences, Peking University.

KOTRČ, M. (2017): "The Rise of Extreme Right: Evidence from the Slovak Parliamentary Election." Bachelor thesis. Charles University.

KOUBA, K. (2007): "Spatial Analysis of the Czech Party System. Institutionalization and Spatial Regimes." *Czech Sociological Review* 43: pp. 1017-1037.

KUBÁT, M. (2006): "Teorie antisystémovosti a Komunistická strana Čech a Moravy." pp. 548-556.

LAZARSFELD, P. (1944): "The Election Is Over." *The Public Opinion Quarterly* 8(3): pp. 317-330.

LESAGE, J. P. & R. K. PACE (2014): "Introduction to Spatial Econometrics." *Chapman & Hall/CRC.*

MEDIAN (2017): "Výzkum pro volební studio ČT." *MEDIAN.*

MEDIAN (2017): "Analýza výsledků voleb podle typu obcí." *MEDIAN.*

NOVÁK, M. (1996): "Volby do Poslanecké sněmovny, vládní nestabilita a perspektivy demokracie v ČR. *Sociologický časopis* 32(4): pp. 407-422.

ORD, K. (1975): "Estimation Methods for Models of Spatial Interaction." *Journal of the American Statistical Association* 70(349): pp. 120-126.

PECKOVÁ, B. (2012): "Charakter a budoucnost KSČM v kontextu jejího vztahu se spřízněnou politickou mládežnickou organizací KSM/ SMKČ." Faculty of Social Sciences, Charles University.

POOLE, K. T. & H. ROSENTAHL (1984): "U.S. Presidential Elections 1968-80: A Spatial Analysis." *American Journal of Political Science* 28(2): pp. 282-312.

QUEYRANNE, M., J. HALLAERT, & I. YACKOVLEV (2018): "Inequality and Poverty Across Generations in the European Union." *IMF Staff Discussion Note.*

RODRÍGUEZ-POSE, A. (2018): "The revenge of the places that don't matter."

ŠÁROVEC, D. (2018): "Nástup nových politických stran v ČR od roku 2013: Hnutí ANO 2011 a Úsvit přímé demokracie pohledem konceptu novosti." *Sociológia 2018* 50(1): pp. 78-113.

WOOLDRIDGE, J. M. (2012): "Introductory Econometrics: A Modern Approach." *Cengage Learning*.

Appendix A

OLS Assumptions

Assumption MLR.1 (Linear in Parameters)
The model in the population can be written as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u,$$

where $\beta_0, \beta_1, \beta_2, \ldots, \beta_k$ are the unknown parameters (constants) of interest and u is an unobserved random error or disturbance term.
Assumption MLR.1 describes the population relationship we hope to estimate, and explicitly sets out the β_j-the ceteris paribus population effects of the x_j on y-as the parameters of interest.

Assumption MLR.2 (Random Sampling)
We have a random sample of n observations, $(x_{i1}, x_{i2}, \hat{a}\pounds\c{e}, x_{ik}, y_i) : i = 1, \hat{a}\pounds\c{e}, n$, following the population model in Assumption MLR.1.
This random sampling assumption means that we have data that can be used to estimate the β_j, and that the data have been chosen to be representative of the population described in Assumption MLR.1.

Assumption MLR.3 (No Perfect Collinearity)
In the sample (and therefore in the population), none of the independent variables is constant, and there are no exact linear relationships among the independent variables. Once we have a sample of data, we need to know that we can use the data to compute the OLS estimates, the $\hat{\beta}_j$. This is the role of Assumption MLR.3: if we have sample variation in each independent variable and no exact linear relationships among the independent variables, we can compute the $\hat{\beta}_j$.

Assumption MLR.4 (Zero Conditional Mean)
The error u has an expected value of zero given any values of the explanatory variables. In other words, $\mathbb{E}(u|x_1, x_2, \ldots, x_k) = 0$.
As we discussed in the text, assuming that the unobserved factors are, on average, unrelated to the explanatory variables is key to deriving the first statistical property of each OLS estimator: its unbiasedness for the corresponding population parameter. Of course, all of the previous assumptions are used to show unbiasedness.

Assumption MLR.5 (Homoskedasticity)
The error u has the same variance given any values of the explanatory variables. In other words,

$$Var(u|x_1, x_2, \ldots, x_k) = \sigma^2$$

.
Compared with Assumption MLR.4, the homoskedasticity assumption is of secondary importance; in particular, Assumption MLR.5 has no bearing on the unbiasedness of the $\hat{\beta}_j$. Still, homoskedasticity has two important implications: (1) We can derive formulas for the sampling variances whose components are easy to characterize; (2) We can conclude, under the GaussMarkov assumptions MLR.1 to MLR.5, that the OLS estimators have smallest variance among all linear, unbiased estimators.

Assumption MLR.6 (Normality)
The population error u is independent of the explanatory variables $x_1, x_2, \ldots, x_k$ and is normally distributed with zero mean and variance $\sigma^2 : u \sim (0, \sigma^2)$.

Source: WOOLDRIDGE, J. M. (2012): "Introductory Econometrics."

Appendix B

Descriprion of Variables

Table B.1: Description of Variables

Variable	Description	Date	Source
turnout	voting turnout	2017	Czech Statistical Office, volby.cz
ANO, ODS, etc.	share of votes for ANO, ODS, etc. within a municipality	2017	Czech Statistical Office, volby.cz
unemployment	share of unemployed people	2017	Ministry of Labour and Social Affairs
distraint	share of people in enforcement proceedings	2017	Mapa exekucí
average age	average age	2018	Czech Statistical Office
men share	share of male population	2018	Czech Statistical Office
migration	change in population (people who moved in or away)	2017	Czech Statistical Office
population growth	natural population growth	2017	Czech Statistical Office
no education	share of people with no education	2011	Czech Statistical Office, census
primary education	share of people with only a primary education	2011	Czech Statistical Office, census
high school	share of people with a completed high school education	2011	Czech Statistical Office, census
university	share of people with a university degree	2011	Czech Statistical Office, census
believers	share of religious people	2011	Czech Statistical Office, census
businessmen	share of enterpreneurs	2017	Czech Statistical Office
population	number of inhabitants within a municipality	2018	Czech Statistical Office

Appendix C

Correlation Table

Table C.1: Correlation Table, Part 1

	unempl.	distraint	average age	men sh.	migration	pop. gr.	no educ.
unempl	1	0.302	0.055	0.025	-0.031	-0.055	0.040
distraint	0.302	1	-0.052	0.092	-0.056	-0.045	0.103
average age	0.055	-0.052	1	0.065	-0.113	-0.374	0.041
men share	0.025	0.092	0.065	1	-0.027	-0.043	0.001
migration	-0.031	-0.056	-0.113	-0.027	1	0.002	-0.002
population growth	-0.055	-0.045	-0.374	-0.043	0.002	1	-0.055
no education	0.040	0.103	0.041	0.001	-0.002	-0.055	1
primary school	0.298	0.360	0.204	0.082	-0.056	-0.181	0.098
high school	0.072	-0.017	0.168	0.108	-0.071	-0.059	-0.128
university	-0.194	-0.252	-0.207	-0.156	0.086	0.113	-0.118
believers	0.015	-0.338	0.098	-0.044	-0.066	-0.038	-0.077
businessmen	-0.169	-0.142	0.004	0.052	0.010	0.026	-0.078
ANO	0.089	0.159	0.093	0.042	-0.007	-0.032	0.021
ODS	-0.174	-0.072	-0.117	-0.060	0.085	0.061	-0.061
Pirati	-0.171	-0.101	-0.132	-0.013	0.057	0.077	-0.035
SPD	0.165	0.180	-0.057	0.037	-0.031	-0.006	0.037
KSCM	0.214	0.197	0.190	0.079	-0.061	-0.117	0.067
CSSD	0.035	-0.039	0.162	-0.033	-0.051	-0.048	-0.003
KDU.CSL	-0.021	-0.289	0.010	-0.049	-0.061	-0.006	-0.049
TOP.09	-0.146	-0.067	-0.165	-0.047	0.077	0.084	-0.062
STAN	-0.093	-0.026	-0.045	-0.011	0.024	0.023	0.031

Table C.2: Correlation Table, Part 2

	primary sch.	high sch.	university	believers	businessmen	ANO	ODS
unemployment	0.298	0.072	-0.194	0.015	-0.169	0.089	-0.174
distraint	0.360	-0.017	-0.252	-0.338	-0.142	0.159	-0.072
average age	0.204	0.168	-0.207	0.098	0.004	0.093	-0.117
men share	0.082	0.108	-0.156	-0.044	0.052	0.042	-0.060
migration	-0.056	-0.071	0.086	-0.066	0.010	-0.007	0.085
population growth	-0.181	-0.059	0.113	-0.038	0.026	-0.032	0.061
no education	0.098	-0.128	-0.118	-0.077	-0.078	0.021	-0.061
primary education	1	0.159	-0.577	0.111	-0.279	0.134	-0.343
high school	0.159	1	-0.616	0.091	-0.271	0.182	-0.351
university	-0.577	-0.616	1	0.019	0.350	-0.245	0.409
believers	0.111	0.091	0.019	1	-0.036	-0.246	-0.200
businessmen	-0.279	-0.271	0.350	-0.036	1	-0.282	0.441
ANO	0.134	0.182	-0.245	-0.246	-0.282	1	-0.325
ODS	-0.343	-0.351	0.409	-0.200	0.441	-0.325	1
Pirati	-0.282	-0.243	0.292	-0.172	0.302	-0.287	0.242
SPD	0.197	0.139	-0.201	-0.066	-0.155	-0.030	-0.246
KSCM	0.325	0.266	-0.350	-0.120	-0.276	-0.016	-0.313
CSSD	0.096	0.130	-0.123	0.156	-0.152	-0.140	-0.175
KDU.CSL	0.067	0.044	0.035	0.795	-0.054	-0.321	-0.210
TOP.09	-0.312	-0.330	0.430	-0.151	0.354	-0.250	0.352
STAN	-0.121	-0.068	0.057	-0.144	0.131	-0.187	0.033

Table C.3: Correlation Table, Part 3

	Pirati	SPD	KSCM	CSSD	KDU.CSL	TOP.09	STAN
unemployment	-0.171	0.165	0.214	0.035	-0.021	-0.146	-0.093
distraint	-0.101	0.180	0.197	-0.039	-0.289	-0.067	-0.026
average age	-0.132	-0.057	0.190	0.162	0.010	-0.165	-0.045
men share	-0.013	0.037	0.079	-0.033	-0.049	-0.047	-0.011
migration	0.057	-0.031	-0.061	-0.051	-0.061	0.077	0.024
population growth	0.077	-0.006	-0.117	-0.048	-0.006	0.084	0.023
no education	-0.035	0.037	0.067	-0.003	-0.049	-0.062	0.031
primary school	-0.282	0.197	0.325	0.096	0.067	-0.312	-0.121
high school	-0.243	0.139	0.266	0.130	0.044	-0.330	-0.068
university	0.292	-0.201	-0.350	-0.123	0.035	0.430	0.057
believers	-0.172	-0.066	-0.120	0.156	0.795	-0.151	-0.144
businessmen	0.302	-0.155	-0.276	-0.152	-0.054	0.354	0.131
ANO	-0.287	-0.030	-0.016	-0.140	-0.321	-0.250	-0.187
ODS	0.242	-0.246	-0.313	-0.175	-0.210	0.352	0.033
Pirati	1	-0.264	-0.258	-0.165	-0.162	0.278	0.062
SPD	-0.264	1	0.089	-0.066	-0.098	-0.242	-0.197
KSCM	-0.258	0.089	1	0.053	-0.142	-0.297	-0.170
CSSD	-0.165	-0.066	0.053	1	0.046	-0.174	-0.169
KDU.CSL	-0.162	-0.098	-0.142	0.046	1	-0.133	-0.170
TOP.09	0.278	-0.242	-0.297	-0.174	-0.133	1	0.056
STAN	0.062	-0.197	-0.170	-0.169	-0.170	0.056	1