

CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies



**Economic Effects of Germans' Expulsion
from Czechoslovakia Revisited**

Master's Thesis

Author: Bc. Kateřina Kafková

Study program: Finance, Financial Markets and Banking

Supervisor: PhDr. Jaromír Baxa, Ph.D.

Year of defense: 2022

Declaration of Authorship

The author hereby declares that she 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, July 25, 2022

Kateřina Kafková

Abstract

This thesis studies long-run effects that expulsion of Germans from Czechoslovakia after World War II had on the Czech borderlands. Specifically, the thesis tests sensitivity of prior literature's findings by investigating previously not addressed potentially confounding factors – presence of military areas and border zone in the data; heavy air pollution in the North-West Bohemia during the communist regime; and impact of Wild expulsion in the Soviet liberation zone right after the end of World War II. Minor to no effects of these factors are revealed. Therefore, the findings of this thesis are in line with the findings of the existing literature - the expulsion of Germans caused higher unemployment, lower population density, lower employment in high-skilled sectors, and a lower share of population with secondary and tertiary education in the borderlands, when compared to the Czech interior, that lasts until today.

JEL Classification D74, I25, N34, N94, O15, R12

Keywords Expulsion, Czechoslovakia, postwar, local development

Title Economic Effects of Germans' Expulsion from Czechoslovakia Revisited

Author's e-mail

Supervisor's e-mail

Abstrakt

Tato práce se zabývá dlouhodobými dopady, které měl odsun Němců z Československa po druhé světové válce na české pohraničí. Konkrétně práce testuje citlivost závěrů předchozí literatury zkoumáním dříve nezkoumaných potenciálně matoucích faktorů - přítomnost vojenských prostorů a pohraničního pásma v datech; silného znečištění ovzduší v severozápadních Čechách během komunistického režimu; a dopad divokého odsunu v sovětské osvobozené zóně těsně po skončení druhé světové války. Vlivy těchto faktorů se ukázaly jako nepatrné až nulové. Zjištění této práce jsou tedy v souladu se závěry dosavadní literatury - odsun Němců způsobil v pohraničí vyšší nezaměstnanost, nižší hustotu zalidnění, nižší zaměstnanost v odvětvích s vysokou kvalifikací a nižší podíl obyvatelstva se středoškolským a vysokoškolským vzděláním ve srovnání s vnitrozemím Čech, které trvají dodnes.

Klasifikace JEL	D74, I25, N34, N94, O15, R12
Klíčová slova	Odsun, Československo, poválečný, lokální rozvoj
Název práce	Ekonomické dopady vyhnání Němců z Československa: Nový pohled
E-mail autora	
E-mail vedoucího práce	

Acknowledgments

This thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 870245.

The author would like to express her gratitude to the primary supervisor, PhDr. Jaromír Baxa, Ph.D., professor at the Institute of Economic Studies, Charles University in Prague, who contributed with his valuable assistance throughout the preparation and completion of this thesis.

Typeset in FSV L^AT_EX template with great thanks to prof. Zuzana Havránková and prof. Tomáš Havránek of Institute of Economic Studies, Faculty of Social Sciences, Charles University.

Bibliographic Record

Kafková, Kateřina: *Economic Effects of Germans' Expulsion from Czechoslovakia Revisited*. Master's Thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 101. Advisor: PhDr. Jaromír Baxa, Ph.D.

Contents

List of Tables	viii
List of Figures	x
Acronyms	xi
Thesis Proposal	xii
1 Introduction	1
2 Historical Background	3
3 Literature Review	12
4 Data and Methodology	16
4.1 Data	16
4.2 Methodology	20
4.2.1 Geographic RDD Using MAL	20
4.2.2 Score variable	23
4.2.3 Exclusion of Military Areas and Border Zone	25
4.2.4 Spatial Heterogeneity and Clustering Standard Errors . .	28
4.3 Balance Testing and Other Concerns	30
4.3.1 Economic Differences between Sudetenland and Interior prior to the Expulsion	30
4.3.2 Pre-Treatment Sorting and Czechs in the Borderlands . .	34
4.3.3 Pollution in North-West Bohemia	37
4.3.4 1945 US-Soviet Demarcation Line	38
5 Empirical Results	45
5.1 Results	45
5.1.1 RDD Using the Munich Agreement Line	45

5.1.2 RDD Using the 1945 Demarcation Line	50
5.2 Robustness	55
6 Discussion	60
7 Conclusion	70
References	I
A Appendix	VI

List of Tables

4.1	German Population in the Czech Lands in 1930 and 1950. . . .	22
4.2	Economic Differences, 1930.	32
4.3	Sectoral Differences, 1930.	33
4.4	Economic Differences with Restricted Sample, 1930.	35
4.5	Sectoral Differences with Restricted Sample, 1930.	36
4.6	Mean and Standard Deviation of Economic Factors in Sample, 1930.	42
4.7	Mean and Standard Deviation of Employment in Sectors in Sam- ple, 1930.	43
5.1	Long-run Differences in Labour Market at MAL, No Military Areas and Border Zone, 2011.	47
5.2	Long-run Differences in Sectoral Composition at MAL, No Mil- itary Areas and Border Zone, 2011.	48
5.3	Long-run Differences in Educational Attainment at MAL, No Military Areas and Border Zone, 2011.	49
5.4	Long-run Differences in Labour Market at 45DL, 2011.	51
5.5	Long-run Differences in Sectoral Composition at 45DL, 2011. . .	52
5.6	Long-run Differences in Educational Attainment at 45DL, 2011.	54
5.7	Long-run Differences at MAL, 10 km bandwidth, 2011.	56
5.8	Long-run Differences at MAL, 50 km bandwidth, 2011.	57
5.9	Long-run Differences at MAL, 8 border segments, 2011.	58
5.10	Long-run Differences at MAL, No Pollution, 2011.	59
6.1	Long-run Differences in Labour Market at MAL, Comparison, 2011.	62
6.2	Long-run Differences in Sectoral Composition at MAL, Compar- ison, 2011.	64

6.3	Long-run Differences in Educational Attainment at MAL, Comparison, 2011.	66
6.4	Long-run Differences at MAL, No Pollution, Comparison, 2011.	68
A.1	Summary Statistics, 2011 (full sample).	VII
A.2	Summary Statistics, 1930 (full sample).	VIII
A.3	Long-run Differences in Labour Market at MAL, 12 km Border Zone, 2011.	X
A.4	Long-run Differences in Labour Market at MAL, 15 km Border Zone, 2011.	XI
A.5	Comparison of non-military and military areas, 1930.	XII
A.6	Comparison of non-military and military areas, 2011.	XIII
A.7	Long-run Differences in Labour Market at MAL, no military areas, 2011.	XIV
A.8	Long-run Differences in Educational Attainment at MAL, no military areas, 2011.	XV

List of Figures

2.1	1945 US-Soviet Demarcation Line in the Czech Lands, municipalities as of 2011	8
2.2	German population in US and Soviet zones (in % of 1930 population)	9
4.1	Czech Lands after the Munich Agreement, 1939	18
4.2	Germans and Poles in Czech lands, pre-expulsion	21
4.3	Military areas in the Czech Lands, municipalities as of 2011	26
4.4	Military areas and the Border zone in the Czech Lands, municipalities as of 2011	27
4.5	Heatmaps, districts as of 2011	29
4.6	Sample municipalities, 2011	40

Acronyms

45DL 1945 Demarcation Line

MAL Munich Agreement Line

RDD Regression Discontinuity Design

Master's Thesis Proposal

Author	Bc. Kateřina Kafková
Supervisor	PhDr. Jaromír Baxa, Ph.D.
Proposed topic	Economic Effects of Germans' Expulsion from Czechoslovakia Revisited

Motivation Forced migration is an event that is occurring quite often, especially in areas plagued by war or also increasingly in areas affected by climate change. Most papers, however, concentrate on analysing the impacts on the migrants or on the host economies, but not many papers focus on the migrants' original areas. For that reason, in this thesis, we will investigate the long-run effects of Germans' expulsion from Czechoslovakia on the Sudetenland, from where the Germans were expelled after World War II.

The expulsion of Germans is an ideal event of forced migration for analysing its impacts on the original country. This is because until the Munich agreement, which became effective in 1938, Czech and German inhabitants lived in the Sudetenland in harmony. In the context of our analysis this means that the treatment group (i.e. German expellees) is very similar to the control group (i.e. Czechs and some Germans that were allowed to stay). The results could be then applied to similar cases of forced migration.

The economic effects of Germans' expulsion from Czechoslovakia have already been studied by Patrick A. Testa in 2020 in his paper "The Economic Legacy of Expulsion: Lessons from Postwar Czechoslovakia". He found dramatic effects of this forced migration – significant negative effects on population density, employment, educational attainment, and on the sector composition (fewer high-skilled jobs, human capital). However, his results might be overestimated by inclusion of regions that were systematically affected by other processes that complemented the Germans' expulsion.

Thus, this thesis will provide broad replication of Testa's paper and will test the robustness of his results.

Methodology We will estimate the differences between regions at the language border between Czechs and Germans using data since the expulsion of Germans in 1945. The dataset we will be using is available at the personal website of P. A. Testa, and it comprises of digitized data from historical censuses, demographic yearbooks, and statistical journals.

Similar to Testa's paper, we shall estimate the long-term effects using the regression discontinuity around the language border. However, before the estimation we will exclude some data from the dataset according to our stated hypotheses, in order to test the hypotheses. We might also add data from more recent years.

Expected Contribution To our knowledge there is only one study analysing the economic effects of Germans' expulsion on the Sudetenland – the one written by P. A. Testa (2020). We shall provide challenging of this study's results by excluding military areas from the sample as well as areas nearby. This act largely reduces the dataset, but it might reveal a lot of valuable information. Also omitting data from areas at the South and Western border that were deserted during the Cold War might yield less significant results. Another matter that may be affecting the Testa's results is that the incentives to settle in North-West Bohemia have then been offset by a heavy environmental pollution in 1990 in this area. For that reason, it should also be interesting to exclude this area from the analysis. Lastly, we may inspect whether the industrial policy of the Council for Mutual Economic Assistance (Comecon) during the Cold War have also affected the economic development of the Borderland.

In general, results of this thesis should provide new information about the effects of forced migration on the migrants' original area of residence. This knowledge of what will happen to such areas, where the population suddenly decreases in a significant way, will also be useful in the future when the cases of forced migration due to climate change might be increasing, or in determining of what will happen in current or future warzone areas.

Outline First, the thesis will provide the review of economic literature that studies similar topics. Second, we will describe the data and the used methodology, together with explanations of the differences from the methods used in Testa's paper. Third, we will provide the results and their interpretation, followed by the discussion of what the findings suggest, which economic impacts did the Germans' expulsion bring, and how does it differ from the Testa's findings. Finally, the thesis will conclude with summary of the contribution of our findings.

Core bibliography

Becker, S. O. and A. Ferrara (2019). "Consequences of forced migration: A survey of recent Findings." *Labour Economics*. 59. Available at: <https://www.sciencedirect.com/science/article/pii/S0927537119300132>.

Gerlach, D. (2017). "The Economy of Ethnic Cleansing: The Transformation of the German- Czech Borderlands after WWII." New York: Cambridge University Press.

Grossmann, J., Jurajda, S. and Roesel, F (2021). "Forced Migration, Staying Minorities, and New Societies: Evidence from Post-War Czechoslovakia." CE-Sifo Working Paper No. 8950, Available at SSRN: <https://ssrn.com/abstract=3812067>.

Hamermesh, D. (2007). "Replication in Economics." IZA Discussion Paper No. 2760. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2966.2007.00428.x>.

Testa, P. A. (2020). "The Economic Legacy of Expulsion: Lessons from Post-war Czechoslovakia." Available at SSRN: <https://ssrn.com/abstract=3251541>.

Ústav pro soudobé dějiny AV ČR (2005). "Komplex odsunu: Vysídlení Němců a české pohraničí po roce 1945." *Soudobé dějiny XII / 3-4*. Available at: http://www.usd.cas.cz/wp-content/uploads/sd_3_4_2005.pdf.

Chapter 1

Introduction

Forced migration is not something that occurs rarely in the history or even today. In the past, minorities such as Jews, Greeks, and Turks have experienced forced migration first-hand. Nowadays, some cases of forced migration can be seen in countries such as Bosnia, Uganda, Syria, or Myanmar, where it is mainly driven by civil and ethnic conflicts. Most recently, we have also seen a forced migration of Ukrainian people caused by the Russian invasion of Ukraine.

A quite large amount of literature studies the impacts of such forced migrations on the migrants and on the economies to which they migrate. However, the topic that is barely discussed in the existing literature is what are effects on the home economies. In other words, how are the economies from which the migrants originate affected by such cases of forced migration. This is precisely what this thesis explores, specifically on the case of the expulsion of German population from Czechoslovakia, which was one of the aftermaths of the World War II.

In this historical event, the citizenship and property rights of almost all Germans living in the Czech lands (this area now corresponds to the Czech Republic) were taken from them, and they were permanently expelled. Nevertheless, this Germans' expulsion is an ideal event of forced migration for analysing its impacts on the original country. This is because until the Munich Agreement, which became effective in 1938, Czech and German inhabitants lived in the Sudetenland next to each other for several centuries with a lot of interactions between both nationalities (including marriages). In the context of our analysis this means that the treatment group (i.e. German expellees) is very similar to the control group (i.e. Czechs and some Germans that were allowed to stay). The results could be then applied to similar cases of forced

migration.

The method that is used to study the effects of the Germans' expulsion is the regression discontinuity design, specifically using the discontinuity in ethnic composition at the Munich Agreement Line (i.e. the border of the Sudetenland) where Germans used to live prior to the expulsion. The value added is that this thesis investigates potentially confounding factors that are not addressed by the prior literature and are suspected of biasing its findings. First one being a presence of military areas and border zone in the data because the resettlement of Czechs has not been allowed there. Second one is a heavy air pollution in the North-West Bohemia during the communist regime that led many people to move out of these polluted areas after the fall of the communist regime when labour force was allowed to move freely and the incentives to live particularly in Ústí nad Labem were eliminated. Third and last studied factor is an impact of Wild expulsion in the Soviet liberation zone right after the end of the World War II. All these factors are expected to have negative effect on population density and, perhaps, on employment, and potentially also some other effects unrelated to the impact of the Germans' expulsion.

The results of this thesis do not confirm any impact of the above listed factors which, in other words, test the sensitivity of findings of the existing literature that say the expulsion of Germans from Sudetenland had significant negative effect on its economic development. Therefore, this thesis provided evidence that the presence of the confounding factors (the military areas, the border zone, pollution in the North-West Bohemia, and the Wild expulsion in the Soviet liberation zone) does not offset the main results of the prior literature.

The thesis is organized as follows. Chapter 2 presents historical background of the Sudetenland and of the Germans' expulsion. Chapter 3 reviews the existing literature that relates to the main topic of this thesis – the literature on the effects of forced migration (or expulsion) in general as well as on the effects of the Germans' expulsion itself. Chapter 4 contains description of data and the used methodology, and also addresses concerns about factors that might have an influence on the effects of the Germans' expulsion as well as the explanation of how they are controlled for in the analysis. Chapter 5 then presents the results of the estimations, and also provides details about testing the robustness of the results. Then, Chapter 6 is dedicated to discussing the findings. And finally, Chapter 7 concludes the thesis.

Chapter 2

Historical Background

Habsburg Rule (1526-1918)

Germans first occurred in the Czech and Moravian borderlands in the 12th century after German-speaking craftsmen were given a permission to enter (Zayas 1994). Then even more Germans came to the borderlands during the Habsburg rule (1526-1918) contributing to the matter of a 'language border' between Czechs and Germans (Daněk 1995). Nevertheless, as the two nationalities coexisted in peace at the local level, even the economic, cultural as well as the genetic factors were largely independent of the language by the beginning of the 19th century (King 2005; Tampke 2003; Zahra 2008; Testa 2020). By the end of the 19th century, the language barrier manifested to be increasingly more problematic. Czech societies lobbying for limitation of bilingual education were established, which then led to Germans wanting to 'Germanize' the mixed areas and appropriating the whole borderlands. In order to prevent Germanization of the borderlands, special institutions, whose objective was to protect the Czech language and the Lands of the Bohemian Crown, were built. As a result, there was a large decline in the German speaking population in the interior of the Czech lands (Zahra 2008).

Interwar Period (1918-1939)

The establishment of the Czechoslovak state in 1918 was praised by its people but not as much by the Czech and Moravian borderlands, which was inhabited

predominantly by the German population. German politicians were invited to join the newly prepared government, but they refused this offer. Instead, they opposed belonging to the Czechoslovak state and established four provinces in the borderlands, the largest of which was the province of Deutsch Böhmen. The Czech and Moravian Germans requested to join the state of German Austria and within it to the entire German Empire. In the second half of November 1918, the Czechoslovak army began occupying the German territory, and at the end of December 1918, the sovereignty of the Czechoslovak state was enforced in all border areas. Following these events, a cooperation between German and Czech populations was established, thus restoring the historical integrity of the Czech lands. This was further strengthened when the Allies – the winning nations of the World War I (WWI), mainly the United Kingdom (UK), France, Russia, Italy, and the United States (US) – supported the Czech efforts to preserve the historic borders (Olivová 2008).

Minority rights were enshrined in the Constitutional Charter of 29 February 1920. The Language Act of 29 February 1920, in accordance with international treaties, guaranteed the existence of minority education and the communication of minority language authorities wherever the minority accounted for at least 20% of the population. But there were complaints from the Germans that these rights were disregarded by many authorities, and in places where both Czech and German languages were always used, officials suddenly pretended that German was an unknown language to them. For some time, the Germans refused any integration into the newly formed republic, and overall, the minorities in the Czech lands held negative attitudes towards the new state (Arburg and Staněk 2011).

In the interwar period, education was one of the closely monitored areas in which national and minority interests traditionally clashed. The language of instruction was derived from the ratio of nationalities in the municipalities in which the school was established. Thus, the need to learn Czech increased for the German population and the Czechs theoretically lost the obligation to learn German. Even the number of Czech schools began to increase, while the number of German primary schools began to decline. Nevertheless, in this interwar period, the German's network of national education was still developed and relatively dense - it had all levels and various specializations, including a university in Prague and two technical schools in Prague and Brno (Arburg and Staněk 2011).

The approach to minorities applied by the Czechoslovak authorities has

gradually improved over the years. From the second half of the 1920s, German politicians began to participate in building the new state, which resulted in a considerable stabilization of conditions in the nationality area. Moreover, from 1926 to 1938, German ministers finally sat even in the Czechoslovak governments (Arburg and Staněk 2011).

This positive development was disrupted by the Great Depression, which was caused by the collapse of the New York Stock Exchange on October 24, 1929, and then another disruption was also caused by financial crisis in Germany in 1931. In the Czech lands, it caused the greatest economic and social problems in the Czech and Moravian borderlands, and thus largely impacted the German population living there. The borderlands were hit by the crisis more severely, in contrast to the interior of the Czech lands, because of an internal weakness of the borderlands' economy. This weakness was related to the borderlands' uneven development and the resulting historical burden, including structural differences – in the borderlands there was a predominance of light and consumer, export-oriented industries such as glass-making and textile weaving. In 1936, at the height of the crisis, the unemployment rate in the majority German areas was around 15%, while in the majority Czech or Slovak districts it was between 6-9%. Many Germans blamed the Czechoslovak government for the crisis (Arburg and Staněk 2011).

In October 1933, the separatist Sudetendeutsche Partei (SdP) headed by Konrad Henlein was founded. In May 1935, before the parliamentary elections, SdP launched a large-scale political campaign, raising funds from Hitler's Germany. The rise of Hitler's power in Germany began to arouse growing admiration among the German population of Czechoslovakia, and so the inhabitants increasingly supported SdP. In the parliamentary elections in May 1935, Henlein's party won 63% of the votes of the German population in Czechoslovakia. The party's success was largely due to the dissatisfaction of the German population with political parties that under-represented their interests and rights (Olivová 2008).

Munich Agreement (1938-1939)

In the autumn of 1937, Nazi Germany was ready to invade Czechoslovakia. In an attempt to avert the impending war, the four great powers (Germany, Italy, France, and the UK) signed the Munich Agreement on September 30, 1938. The result of this agreement was a cession of the Sudetenland to the

Hitler's German Empire, but Czechoslovakia itself was not invited to the negotiation. The agreement confirmed that Czechoslovakia should also meet the territorial requirements of Poland, which requested part of Těšín area (former political districts Český Těšín and Fryštát), and the requirements of Hungary for southern Slovakia and Subcarpathian Russia (Olivová 2008).

Since the borderlands had natural resources, fertile land, and important industry in north-western Bohemia, the newly created borders disrupted the trade linkages and weakened Czechoslovakia's industrial capacities. The most important lines of military fortifications were also located in the annexed areas, so Czechoslovakia was prevented from defending itself (Kural and Radvanovský 2002). The immediate consequence of the Munich Agreement was the gradual territorial disintegration of the republic. On March 14, 1939, the independent Slovak state was created and a day later, on March 15, 1939, the German military occupation of the Czechoslovak territory began. On March 16, 1939, Adolf Hitler arrived in Prague to officially declare the formation of the Protectorate of Bohemia and Moravia and, thus, connecting it to the German Empire. This also meant that all Protectorate's activities were subject to Hitler's decisions. Half a year later, the WWII began when on September 1, 1939, Germany invaded Poland (Olivová 2008).

Expulsion of Germans from Czechoslovakia (1945-1947)

The Czechoslovak domestic resistance as well as the Czechoslovak resistance in exile demanded for tough and uncompromising action against the minority German population. Though this demand began to intensify in 1941, Edvard Beneš, the Czechoslovak president in exile, and the politicians who surrounded him did not finalise the specific forms of the expulsion until 1943. The legal basis for the expulsions were contained in presidential decrees, referred to as the Beneš Decrees. These documents were outlining the expulsion plan for approximately two million Germans, and they also said that all property of the German expellees would be confiscated without compensation and allocated to incoming settlers. According to Beneš, this was also a plan of social transformation of Czechoslovakia. Stalin, the then governor of the Soviet Union, supported all demands against the defeated Germany, so he fully agreed with the Beneš' plans. On the contrary, the Western powers had a more distant ap-

proach, and demanded that Czechoslovakia does not take any unilateral action on this issue and waits for the Allies' agreement (Staněk 1991).

Period of Wild Expulsion (May 1945 – August 1945)

After the end of WWII in Europe, in May 1945, Czechoslovakia was liberated by the Allied forces and divided into two zones – one was an area liberated by the US forces (US was one of the Western Allies, second one was the UK) and the other was an area which the Soviet forces liberated. Figure 2.1 shows the 1945 US-Soviet demarcation line, where the two forces met. Then, a period of wild expulsions of Germans followed and lasted until the beginning of August 1945 (at least officially). During this time, the expulsion actions lacked a sufficient preparedness and organization, and thus there were very harsh conditions for the German population. There was no unified control, and the wild expulsions were even accompanied by a series of violence and brutality on the German population, especially in the first few weeks after the liberation of Czechoslovakia. These actions were the result of the rise of radical nationalism, the desire for revenge and the strong anti-German thinking of the Czech population. By end of July 1945, about 800,000 Germans had been expelled (Staněk 1992).

Even though the war ended in Europe in May 1945, the US and Soviet forces did not leave Czechoslovakia until December 1945. Some literature notes that the treatment of the Sudeten Germans that were in the Soviet liberation zone differed from the treatment of those in the US liberation zone during this early post-war period. According to Brandes (2002), following the capitulation of Germany, wild and violent expulsion of Germans performed by Czech mobs began immediately in the part of Czechoslovakia which was controlled by the Soviet forces. But since this act was not officially approved by the Allies until the Potsdam Conference in August 1945, the US army prevented the violent expulsion from happening in the area that they controlled (Slapnicka 2000). Suppan (2006) also states that during this wild period in the Soviet zone at least 700 000 Germans were expelled, and thousands were killed.

Grossmann, Jurajda, and Roesel (2021) supported this also with historical data on German population in Karlovy Vary district which they collected from local archives in Sokolov and Karlovy Vary, and also from archives of the Czech Ministry of Foreign Affairs, the Czech Statistical Office, and the German Statistical Office. Figure 2.2 is from their paper, and it shows separately the

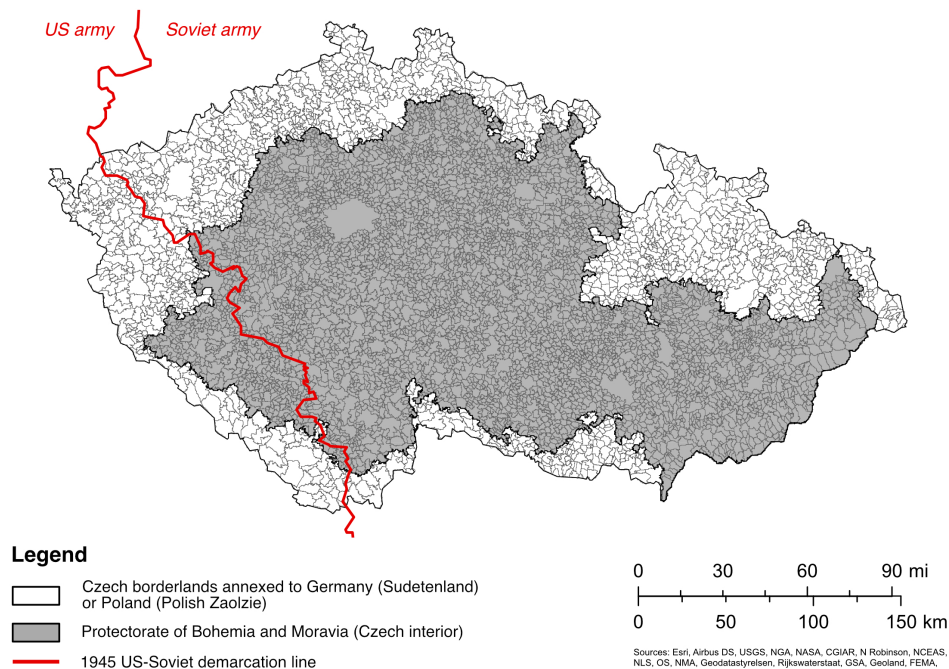


Figure 2.1: 1945 US-Soviet Demarcation Line in the Czech Lands, municipalities as of 2011

development of the share of staying German population in the 1930 total population in the US and in the Soviet zones during years 1930 to 1947. There is a significant difference in the rate of decline of the German population between the two liberation zones. Specifically, the decrease of the German population in the Soviet-liberated area during the first three months (May-July 1945) is much more significant compared to the decrease in the US army-liberated area.

Period of Organized Expulsion (August 1945 – 1947)

At the Potsdam Conference, which was held from July 17 to August 2, 1945, the Allies officially approved the expulsion of the German population, stripping of the citizenship of the remaining Germans. The voluntary departure of Germans from the Sudetenland as well as from the interior automatically intensified, and local authorities were ordered not to obstruct this movement (Kučera 1992). The plan to expel the German population in accordance with the principles agreed in Potsdam was approved in November 1945 in Berlin (Staněk 1991). The exact number of Germans to be transferred to the individual occupation zones in Germany was also determined. A period of more

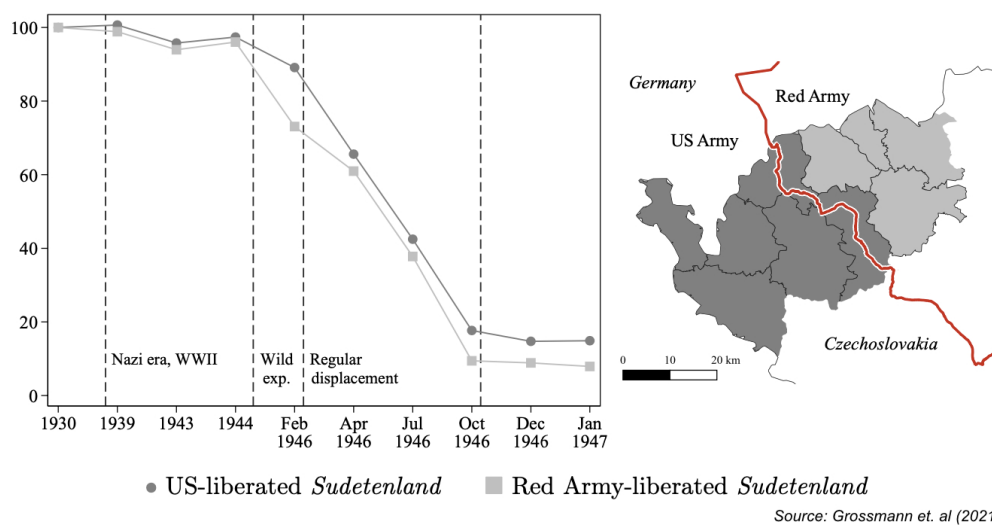


Figure 2.2: German population in US and Soviet zones (in % of 1930 population)

organized and systematic transports of Sudeten Germans followed. In case of ambiguity, the authorities decided whether the person should be deported or whether he/she could stay, and these decisions were not completely unbiased. For example, some Germans who became Czechs before the war were allowed to stay, while some non-German citizens who became Germans after annexation and other citizens of unclear or mixed nationality were expelled. The stay was guaranteed only to those Germans who were Czechs because of marriage, to those demonstrably loyal to the Czechoslovak state, or to those who were considered to be economically vital (Spurný 2013). According to Kučera (1992), during the organized expulsion in 1946, 1,222,628 people of German nationality were displaced from Czechoslovakia to the US occupation zone in Germany and 633,984 to the Soviet occupation zone in Germany (Kučera 1992). Then in mid-1947, the number of expelled Germans reach 2.2 million, and by 1950, it was 3 million Germans (Gerlach 2017).

Resettlement of Czech Borderlands (1948-1960)

The Germans' expulsion as well as the resettlement of the borderlands were realized for the most part by mid-1947. In both cases, their implementation was directed by the Communist Party, which put their specific ideas into the plans and pursued their own political goals. While the expulsion was a result of a nationwide consensus, the settlement was enforced by the Communists,

regardless of criticism from other political parties (Topinka 2005). In order to motivate interior Czechs to move to the borderlands, the government introduced various incentive systems. For example, the government announced that a new settler could gain a post of an administrator of a farm or some large firm to temporarily secure the operation of the enterprise, and eventually he/she would become the owner. In order to become an administrator, the settler had to apply to a national committee and prove his/her knowledge or training relevant to the field of the particular enterprise. Nevertheless, in practice, the verification of the settlers' experiences prior to being awarded a property or title was not so thorough. Other incentives were aimed at attracting workers – low prices of lands and houses, bonuses to workers in critical industries (for example in mining), lower rents than in the interior areas (lower by 25%), low-interest loans to farmers (Testa 2020).

From the beginning of May 1945 to the end of May 1947, 1.4 million people moved to the borderlands. So, after the end of this main settlement phase, a total of 2.2 million people lived in the Czech and Moravian borderlands, a third of whom were the original population, including about 160 thousand staying Germans. This was still less than the Czechoslovak government had planned in their program, which spoke of populating the borderlands by 2.5 million people of Slavic descent, so that the pre-war output could be maintained. Also, the resettlement was uneven because, while cities were populated quite quickly and easily (especially those with appealing property, and those close to the interior), the countryside usually did not attract new settlers due to the more difficult living conditions. Thus, even though the settlers were planned to consist exclusively of Czechs from the near interior, politicians were forced to recruit some Slovaks and other people from abroad due to labour shortages (Topinka 2005).

Military Areas and Border Zone

Since military areas and the border zone that are (or were) located on the territory of the present-day Czech Republic are addressed in this analysis, a brief summary of their history is included here. In total 9 military areas were established in the Czech lands during the 20th century. The first one was Brdy (also known as Jince) and it was created prior to the WWII in 1927 west of Příbram. In that year, Dědice (also known as Březina) was established west of Vyškov. Then, during the WWII, military districts were expanded

and used by the armed forces of the German Empire. After the war, the army's interest in training areas waned for a short time, but then the emerging communist regime set aside areas comprising cadastral for almost two hundred municipalities for military purposes. The Government of the Czech Republic decided on the creation of these new military areas in 1946, which included the Libavá Military Training Area (northeast of Olomouc), Boletice (west of Český Krumlov), Ralsko (east of Mimoň), Prameny (in Slavkovský les), Dobrá Voda (west of Hartmanice), Doupov (also known as Hradiště, south of Kadaň) and Milovice (north of Lysá nad Labem) (Kubisa 2010).

Then, in 1950, also an area with significantly limited access called 'border zone' was established along the borders with Austria and the West Germany (the Federal Republic of Germany) and operated with certain changes and modifications until November 1989 (Etnologický ústav AV n.d.). This zone was guarded by the army, but it was not a military training area (Kubisa 2010). Its purpose was to create an impenetrable border between the Western and the Eastern bloc. The border zone extended about 12-15 km into the Czech territory from the state border, and then, after 1964, it was reduced to a strip of territory about 2-3 km wide. During the creation of the border zone, a number of villages and solitudes near the borders have been displaced, much of which has been demolished, either completely or for the most part (Etnologický ústav AV n.d.).

Chapter 3

Literature Review

Many papers studying forced migration have been published, especially in the last two decades, signifying this is a current and quite popular topic among the academics. The contribution of this thesis to the already existing literature is as follows. Firstly, the majority of the papers that study the topic of forced migration focus not on the expellees' country of origin, but on the host countries. The papers that studied forced migration caused by the World War II (WWII) usually found positive long-term effects on the host economies due to effective resettlement policies (Braun and Mahmoud 2011; Schumann 2014; Braun, Kramer, et al. 2021). On the contrary, the studies that explored developing host economies mostly found evidence for adverse effects of forced migration such as loss of productive workers (Tobin 1974; Calderón and Ibáñez 2005; Alix-Garcia and Saah 2010; Hornung 2014).

Then, another large part of the existing literature focuses on studying the impacts on the migrants themselves that are forced to migrate (Bauer, Braun, and Kvasnicka 2013; Becker, Grosfeld, et al. 2020). Serious consequences are found on migrants from developing countries as they then generally have worse labour market outcomes and lower income, while the impacts in cases of forced migration caused by WWII are positive due to the effective resettlement policies. A specific example of this is the expulsion from Poland as it resulted in the migrants' resettlement in largely empty territory left by Germans, and due to their experiences of loss of physical assets and unlimited access to schooling, these migrants were positively affected in the long-term (Becker, Grosfeld, et al. 2020).

On the contrary, a specific example of forced migration that is the subject of this thesis – the expulsion of Germans from Czechoslovakia – has only been

investigated by a few papers. To my knowledge, there are 3 papers that focus exactly on this topic. First, a paper that studies how this expulsion affected the integration of Germans that were allowed to stay in Czechoslovakia into the resettled communities is by Grossmann, Jurajda, and Roesel (2021). Many of these Germans which were not expelled had left-leaning anti-fascists attitudes. In the paper, the authors studied quasi-experimental local variation in the number of anti-fascist Germans that stayed in Czechoslovakia. They found a long-lasting effect in the places where the anti-fascist Germans stayed in larger numbers. In these areas, the left-leaning attitudes (including support of the Communist Party and far-left social policies) are stronger even today. Grossmann, Jurajda, and Roesel (2021) also found that this political identity even replaced German ethnic identity among staying Germans who were faced with new local ethnic majorities. Second, another paper on the Germans' expulsion is from Guzi, Huber, and M. Mikula (2019), who study its effects on migration and by using the nearest neighbour matching and regression discontinuity design (RDD). They find that even today there is a higher emigration from the areas that were resettled after the expulsion.

And third, a paper that is the most important one for this thesis is by Testa (2020). It studies the long-run effects of the Germans' expulsion from Czechoslovakia after WWII on the economic development in the Czech borderlands (so, it analyses effects the expulsion has on the origin economy). To detect it, Testa (2020) uses the discontinuity in ethnic composition at the border of the Sudetenland (the Czech borderlands) where Germans lived. In his analysis, he uses his observation that Germans and Czechs had very similar characteristics. According to Testa's results, following the expulsion there were persistent differences between the Sudetenland and the rest of Czech lands. He found dramatic effects of this forced migration on the former-German areas – significant negative effects on population density, employment, educational attainment, and on the sector composition (fewer high-skilled jobs, human capital). Testa (2020) also identified the channels that caused the urban decay and decline in the human capital, and these channels were a selective initial resettlement and capital extraction after the expulsion.

Nevertheless, there remains a suspicion that the results from Testa (2020) are possibly exaggerated. That is because there may be some other aspects playing a role in the identification of long-run effects of Germans' expulsion, which Testa (2020) has not considered in his paper. And this is exactly the aim of this thesis – to identify other factors that might be driving the long-run

impact of the expulsion of the Sudeten Germans.

First possible effect may be the presence of military areas and border zone in the analysed sample as Czechs were not able to resettle there. To my knowledge, no literature on the Germans' expulsion addresses this. Then, another factor that may be biasing the results is not excluding areas that had high levels of air pollution in the second half of the 20th century because of production of large amounts of SO₂ by lignite-burning power plants. No literature on Germans' expulsion controls for it even though, for example, Š. Mikula and Pytliková (2021), who study the effects of pollution in North-West Bohemia that was there during the communist era, find that after implementing desulfurization technologies in 1990s the emigration from the highly polluted areas largely decreased, thus suggesting that the pollution was indeed affecting the economic development in these areas.

And lastly, the 1945 US-Soviet demarcation line in the Czech lands may also influence long-run developments in the borderlands. This is addressed only in Grossmann, Jurajda, and Roesel (2021), who find variations in the anti-fascist German population between the West and East sides of the demarcation line. Similar topic is also studied by Eder and Halla (2016) who study effects of migration in Austria that was part of the immediate aftermath of the WWII. At that time, Austria was divided into four occupation zones, one of them being the occupation zone of the Union of Soviet Socialist Republics (USSR). It was this Soviet zone which suffered a loss of 11% of its population, as these people decided to migrate into other parts of Austria (behind the demarcation line). Eder and Halla (2016) then examined how this internal migration shock affected the distribution of economic activity across Austria. Although this is not a proper case of forced migration, it is relevant to this thesis as they use spatial method in their analysis, specifically the spatial equilibrium method. Their results confirm that the distortion on the population distribution has fully persisted, and on top of that they found there is an even higher concentration of population in the former non-Soviet zones. According to them, this unevenness in population distribution, and thus this gap in economic activity, is long-lasting and it is even growing over time. This shows that even a temporary shock as such is able to shift economy's spatial equilibrium. A very similar paper from Ochsner (2017) also finds economic backwardness in South Austrian regions liberated by the Soviet army that lasts until today.

Another paper that has a slightly different topic but is interesting for us mainly because of its methodology is Wyrwich (2020). In this paper, Wyr-

wich (2020) studies long-run effects that a migration barrier has on regional development. Specifically, his point of interest is the Germans' expulsion after WWII, but he studies it in a much larger scale because he analyses the whole Central and Eastern Europe. During this event, the expelled Germans were initially prohibited from entering and settling in the French occupation zone in Germany. Difference-in-Differences (DiD) analysis is used to assess the impacts. Wyrwich (2020) confirms that the migration barrier had long-lasting consequences because, according to his results, in the former French occupation zone there were lower population growth and lower population density that lasted even until today (60 years after the barrier's removal). The article also finds that a trend in regional development which was common prior to the barrier was then disturbed by the changes in population dynamics in agglomerated areas.

Chapter 4

Data and Methodology

This chapter describes the data that are used in the analysis. Specifically, it presents sources and a description of the principal variables (along with their derivation). This chapter also contains information about the empirical approach of this thesis. Namely, it depicts the econometric methods that are applied and why they are appropriate to use in this case. The details of tests performed are also included in this chapter.

4.1 Data

The data used in this thesis are statistical data for the Czech lands that include various indicators such as population density, share of different nationalities (Czechoslovak, German, Russian & Ukrainian, Polish, Hungarian, Romani, Jewish), labour force density, unemployment, roads, railways, percentage of labour force in each working sector (for example agricultural sector, industry, service sector).

The dataset covers the period from 1930 to 2011 on a roughly ten-yearly basis, so that it captures all phases to help reveal the effects of the Germans' expulsion from Czechoslovakia. The data can be divided into four phases. The first phase, when data from 1930 are used, is the situation before the expulsion (also before the WWII), which contains information on how the Germans and Czechs lived prior to the conflicts. This first phase is therefore crucial to clarify the differences among the two nationalities. This identification of differences is needed to decide on the empirical methods that are suitable for the analysis of effects. Then the second phase, for which data from 1945-1950 are used, is the expulsion itself and the period shortly after it, which reveals the primary

effects of the expulsion such as changes in the population and in the population composition. The third phase is the situation few more years after the expulsion, which assists in identifying the short-run effects. And finally, the fourth phase is the period long after the expulsion of Germans, and the data that are used for this period are from 2011. This last phase is important for studying the long-run effects, which is the main aim of this thesis.

The dataset is compiled from several sources. First and main data source is the website of Testa (2021), where the dataset he used in his paper is published. Testa compiled this dataset himself and it contains data of district- and municipal-level for the period from 1930 to 2011. He gathered the data from statistical journals, demographic yearbooks, and also from historical censuses. According to his paper, some of the data was already digitized, but some he was compelled to digitize himself.

Data Source for the Treatment Variable ‘In borderlands’

First, the treatment variable in this thesis is a dummy variable indicating whether a district is located in the borderlands (formerly known as the Sudetenland) or in the interior of the Czech lands. The border that divides the borderlands from the interior is the Munich Agreement line (MAL), which was concluded in Munich in 1938 by representatives of four countries – Germany, Italy, France and the UK.

This treatment variable was used by Testa (2020) as well, so I apply his method of obtaining it. There are two data sources for the creation of this dummy variable. The first source is the Official German Local Register for the Protectorate of Bohemia and Moravia (1940)¹, which lists all villages in the Protectorate of Bohemia and Moravia (i.e. the interior villages) by German and Czech name, regional council district, and subdistrict. The second source for the dummy variable is rather supplemental, and it is the Sudeten German Directory of Place Names (1987)². This directory lists villages in the annexed borderlands, where the German population prevailed, also by German and Czech name, and government district.

Finally, as in the modern times new villages have emerged, there is a need to determine whether they belong to the borderlands (i.e. the treatment group) or to the interior (i.e. the control group). Similarly to Testa (2020), to sort these new villages and by that precisely define the MAL also in the modern

¹“Amtliches Deutsches Ortsbuch für das Protektorat Böhmen und Mähren” (1940)

²“Sudetendeutsches Ortsnamenverzeichnis” (1987)

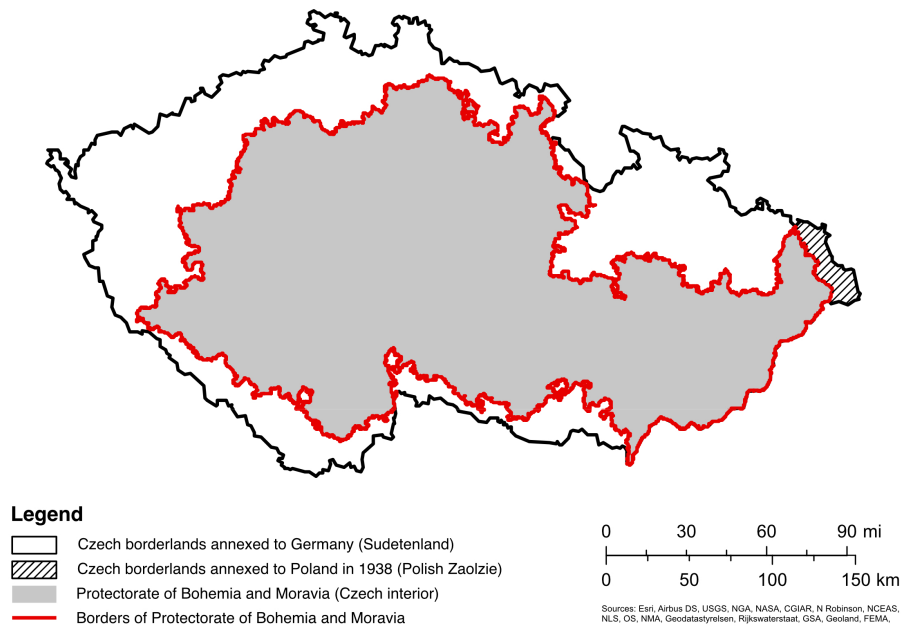


Figure 4.1: Czech Lands after the Munich Agreement, 1939

years, I use the GIS maps of the Protectorate. The creator of these maps is Jelínek (2011). Then, second source of data for the MAL creation are the Czech Statistical Office and the Czech Land Survey Office (ARCDATA PRAHA) that list modern sub-municipal villages. The final form of the MAL is shown in Figure 4.1.

Data Source for the Treatment Variable ‘In US zone’

The same dataset with municipal-level data described earlier in this chapter is used to study the differences between the US and the Soviet liberation zones. It covers part of the interwar period (1930-1939) and the post-WWII period (1947-2011). The data contain information on indicators such as population density, share of different nationalities, labour force density, unemployment, and percentage of labour force in each working sector. The sources for this dataset were historical censuses, demographic yearbooks, and statistical journals.

In order to distinguish whether a municipality is located in the US or in the Soviet liberation zone, the 1945 US-Soviet demarcation line (45DL) is constructed, and the data are combined with the information on its location. The construction of this 45DL is done by creating a shapefile in Google Earth follow-

ing a detailed description from Pecka (1995) (see the Appendix for an excerpt from the Pecka's report describing the exact location of the 45DL in the Czech lands). Grossmann, Jurajda, and Roesel (2021) also used this description of the demarcation line in their analysis. The final form of the 45DL is shown later in this chapter.

Census Data

Second, data from censuses are used in the analysis. In the Czech lands, the censuses are held each ten years from 1921. One exception is the WWII, during which no census was held. Nevertheless, a population index was compiled in 1947, which is sufficient to use as a substitute for the gap in the censuses.

The census from 1930 lists 330 judicial and 151 political districts, and contains data on employment by sector, ethnic composition, and literacy. Then, the number of judicial districts increased to 163 in the index from 1947. But later, in the census from 1950, the judicial and political districts have been merged into 182 districts. Both, the 1947 index and the 1950 census are missing information on literacy and education. Nevertheless, the 1950 census is important as it is the first set of data that provides information on the situation after the WWII. Next, in censuses from 1961-1991 the districts are aggregated to only 76 districts and contain data on employment by sector, ethnic composition, and educational attainment. And lastly, the number of districts in censuses from 2001 and 2011 increased to 77 and they also contain data on individual municipalities. The census data were compiled by Testa (2020), and they were obtained from the Czech Statistical Office in either digital or scanned form. The other sources for census data are also the ARCDATA PRAHA (2017) and the Urban and Regional Laboratory at Charles University in collaboration with the Czech Ministry of Culture (2017), who also created and published corresponding district shapefiles for each census year.

Non-Census Data

Finally, the non-census outcome data are obtained from various sources. The data sources for the years prior to the expulsion are state social insurance reports and taxation reports, which contain a share of eligible taxpayers, a number of unemployed, and income. From these data, the unemployment rate and income per capita are then derived. The pre-expulsion data on railways are from the Historical Geographical Information System for Europe (HGISE), and

data on roads are derived from a map from Autoclub of the Czechoslovak Republic. The crime data are obtained from the Urban and Regional Laboratory (2018). Next, the post-expulsion non-census outcome data are from several statistical reports. Then, migration data are provided by CSO in their demographic yearbooks, and data on abandoned and destroyed villages or places are from the website zanikleobce.cz (2018). Moreover, municipal unemployment data for 2011 are available on the website of the Czech Ministry of Labour and Social Affairs.

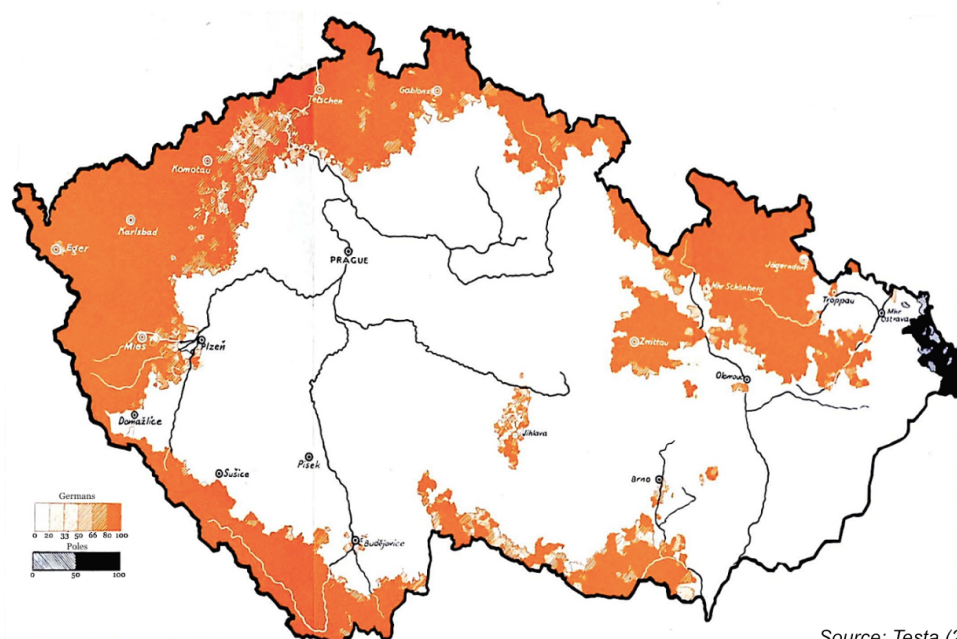
And lastly, details about the Czech military areas are from the webpage of Ministry of the Interior of the Czech Republic (Kubisa 2010), where there is a list of all military areas that existed in the Czech lands, their founding dates, and their approximate location. Their exact location is illustrated in a map from the Institute of Ethnology of the Czech Academy of Sciences (Etnologický ústav AV n.d.), which also shows the approximate location of the former border zone. This map is then used to determine, which modern municipalities are located in areas, where military districts or the border zone used to be. The same approach is used to identify also the pre-expulsion districts with military areas or border zone.

4.2 Methodology

4.2.1 Geographic RDD Using MAL

This section of the thesis explains the empirical approach that is used to identify how the expulsion of Germans from Czechoslovakia affected the former-German places over the long-run. Studying the effect of this event is possible because there was a clear variation in where in the Czech lands people identifying as being German and people identifying as being Czech (or other nationality) before the expulsion happened. People that were identifying as Germans lived almost exclusively in the borderlands (also known as the Sudetenland). This is illustrated in Figure 4.2 as it shows a spatial discontinuity in the German identification (Testa 2020).

This German identification was a key characteristic during the expulsion because it concerned the people that were identified as Germans. Therefore, it can be observed how the exposure to the expulsion differs across Czech municipalities or districts (generally said, across administrative units). Moreover, this provides a way to detect impacts of the forced migration on the places



Source: Testa (2020)

Figure 4.2: Germans and Poles in Czech lands, pre-expulsion

from which the Germans were expelled. Also, in Table 4.1, the data from the sample that is used in this analysis – the districts (or later municipalities) that are no more than 25 km away from the MAL – show that prior to the expulsion a majority of the population living in the borderlands was German (over 80%), compared to less than 2% in the interior. Table 4.1 also shows how dramatically the expulsion changed the composition of the borderlands – German population decreased to about 4.5%, and in the interior districts to roughly 0.5%. The percentages are roughly the same even if the bandwidth around the MAL is increased to 50 km as well as when the whole sample (without any bandwidth) is used. This discontinuity in the German population is important as it allows us to apply the empirical approach that is described in the following paragraphs.

Empirical approach in this thesis is similar to a form of geographic Regression Discontinuity Design (RDD) from Lee and Lemieux (2010) as it analyses discontinuities in average treatment intensity across space. Nevertheless, instead of using present-day borders of some geographical area, historical political borders of an area that no longer exists are used as the cut-off point. Another aspect that distinguishes the approach in this thesis from the standard methods is the very form of the treatment assignment. This is because the districts (or municipalities) identified as majority-German were also inhabited

Table 4.1: German Population in the Czech Lands in 1930 and 1950.

Region (subsample)	Pre-expulsion	Post-expulsion
	% German, 1930	% German, 1950
Borderlands (25 km bandwidth, no overlap)	80.342 (2.443)	4.534 (0.614)
Interior (25 km bandwidth, no overlap)	1.601 (0.414)	0.481 (0.048)
Borderlands (50 km bandwidth, no overlap)	85.317 (1.837)	6.903 (1.326)
Interior (50 km bandwidth, no overlap)	1.781 (0.476)	0.469 (0.041)
Borderlands (full sample)	79.514 (2.046)	5.560 (0.899)
Interior (full sample)	3.573 (0.566)	0.565 (0.044)

Notes: A district is classified as ‘within 25 km’ if its centroid lies within 25 km of the MAL. Military areas are excluded (explained later in this chapter).

by few Czechs, and the same holds for majority-Czech districts (or municipalities) where there were also some German inhabitants. Thus, our empirical method is also close to the fuzzy RDD as the variation between the treatment and the control group is not binary but it is discontinuous (Testa 2020). In order to use this method, assumptions similar to the ones of the geographic RDD must be met and they are to be discussed later.

So, the treatment variable is a dummy variable that signifies whether a district or a municipality is located in the borderlands or in the interior. Thus, it indicates how a district (or a municipality) is exposed to the effects of the expulsion. Specifically, when a district is in the borderlands (or on the side of the MAL, where most of the population is German) then it is assigned a value of 1. On the other hand, if a district is located in the interior of the Czech lands (so, on the side of the MAL, where the population of Germans was low) then it is assigned a value of 0.

When creating this treatment variable, conditions of what a district (or municipality) must meet to be identified as being located in the borderlands were set in the following way. When 50% or more of a district’s (or municipality’s) area lies inside the lands that were annexed by Germany in 1938 in accordance with the MAL, then it is classified as borderlands. On the other hand, when 50% or more of a district’s (or municipality’s) area lies outside these annexed lands, then it is classified as interior. Moreover, ‘no-overlap’ districts (or municipalities) are also distinguished in the data, and it is the area that has more than 95% of its land on one side of the MAL or the other. In other words,

almost the whole territory of the no-overlap area lies either in the borderlands or in the interior (this area nearly does not overlap the two groups). So, this treatment variable allows us to compare the areas on both sides of the MAL, or stated differently, how important economic indicators (such as unemployment or human capital) change when we cross the MAL. Certainly, other potentially confounding factors such as other historical shocks must be considered in the analysis, in order to obtain unbiased results.

To begin with, the equation for municipal-level data is as follows

$$y_{mdb} = \alpha + \beta InBorderlands_m + f(geographic\ location_m) + \gamma G'_m + B_b + D_d + \epsilon_{mdb},$$

where y_{mdb} is the dependent variable for municipality m in district d along segment b of the MAL. Then, α is the intercept; β is the coefficient of interest and it is showing the slope of the treatment variable, and $InBorderlands_m$ is the binary treatment variable for a municipality m with value of 1 for municipalities located in the borderlands and value of 0 for municipalities located in the interior. This β measures the (local average) treatment effect at the MAL (border between the Sudetenland and the interior). Next, $f(geographic\ location_m)$ is a score variable for municipality m . The following term, G'_m , represents a set of geographic characteristics of a municipality m . B_b is a set of border segment fixed effects for border b (explained later in this chapter), D_d is a set of district fixed effects for district d , and ϵ_{mdb} is the error term for municipality m in district d along segment b of the MAL.

4.2.2 Score variable

The score variable, $f(geographic\ location_m)$, is included in the model so that relevant factors which vary across space (except for the treatment) are accounted for in the analysis. For example, such factors are differences in geography and in access to the market, or it may also be differences in the historical experiences. But these factors must be evolving smoothly at the boundary, which is explained in more detail later in this thesis. The ambition is to have β in the above equation produce estimates that are only related to the treatment, while other factors are held fixed. According to Gelman and Imbens (2018), a suitable score variable is a linear polynomial as it serves as a good approximation while having low noise in the final estimates. In our

case, a suitable score variable is a linear polynomial of geographic location, specifically two types of this variable are used, and they are discussed in the following paragraphs.

But first, it is necessary to limit our sample to only contain a narrow bandwidth around the border of interest, which is a standard approach in RDD. In the spatial RDD estimations, a choice of bandwidth tends to influence the estimates. Thus, it is important to choose one that is reasonable for the analysed area and the available data. In the existing literature that studies similar topics a standard choice of bandwidth is 25 km. Since the area that is analysed here is rather small and having a very wide bandwidth may unbalance the sample as some segments of the MAL would contain much more municipalities than others, a bandwidth of 25 km from the MAL is adopted as a baseline bandwidth choice in this analysis as well. Nevertheless, 10 km and 50 km bandwidths are used as well, but they serve as robustness checks.

So, as mentioned above, two score variables are chosen. According to Becker, Boeckh, et al. (2014), for studying the long-run effects of such historical border polynomials in distance to the border should be used. So, the first score variable in this analysis is a polynomial in distance to the MAL. Specifically, the linear polynomial interacted with the treatment is used, which allows the analysis to focus only on the areas in the borderlands (i.e. treated areas) that are closest to the interior areas (i.e. control areas). This is beneficial to the analysis because in the areas that are located further from the MAL the differences from the interior areas grow (for example differences in pre-expulsion labour markets). Also, the further an area is from the MAL, the more expensive it is for the Czech inhabitants to resettle there, meaning that significant long-run differences after the expulsion in those areas would not be so surprising. The baseline equation with the polynomial in distance to the MAL is then as follows:

$$y_{mdb} = \beta_0 + \beta_1 \cdot InBorderlands_m + \beta_2 \cdot DTB_m + \beta_3 \cdot DTB_m \cdot InBorderlands_m + \beta_4 \cdot G'_m + B_b + D_d + \epsilon_{mdb},$$

where DTB_m denotes the distance to border of the Sudetenland for municipality m . This distance is positive for the borderland municipalities and negative for the interior municipalities.

Then, according to Dell (2010), we should not rely only on the one-dimensi-

onal distance score variable. Thus, a second score variable that is used is a two-dimensional linear polynomial in latitude and longitude (Dell 2010). The baseline equation in this case is as follows:

$$y_{mdb} = \beta_0 + \beta_1 \cdot InBorderlands_m + \beta_2 \cdot Latitude_m + \beta_3 \cdot Longitude_m + \beta_4 \cdot G'_m + B_b + D_d + \epsilon_{mdb}.$$

But according to Lee and Lemieux (2010), this approach is not capable of abstracting from the relevant factors, other than the treatment, which might also influence the results. Therefore, there is a high chance that using this second score variable will produce biased estimates. This problem cannot be solved by having this variable interact with the treatment because then this would not estimate the treatment effects at the MAL. Nonetheless, this latitude and longitude score variable is useful to have and produce secondary estimates to serve as a completeness check. Then, to further verify the robustness of the results, different degrees of polynomials in both distance and latitude-longitude are used for some of the dependent variables.

4.2.3 Exclusion of Military Areas and Border Zone

The presence of military areas in the analysed groups is not addressed by the existing literature. It might be important to exclude them from the analysis because many of the military areas in the Czech lands were established after the WWII (Kubisa 2010), i.e. during the period of our interest. Many villages ceased to exist due to the creation of the military areas and its inhabitants were forced to resettle somewhere else (zanikleobce.cz). Also, places that are in the vicinity of the military areas might not be very appealing to the Czech inhabitants, so not many of them may settle there. For these reasons, large economic differences between such areas and the rest of the areas in the borderlands or in the interior would not be surprising. As described in Chapter 2, in total 9 military areas were established in the Czech lands during the 20th century – Brdy (also known as Jince), Dědice (also known as Březina), Libavá, Boletice, Ralsko, Prameny, Dobrá Voda, Doupov (also known as Hradiště), and Milovice. Figure 4.3 shows a map of municipalities in the modern Czech Republic, and dark grey colour highlights the municipalities that are also military areas. These highlighted municipalities are then excluded from the analysis.

Then, the border zone, which was created in 1950 along the borders with

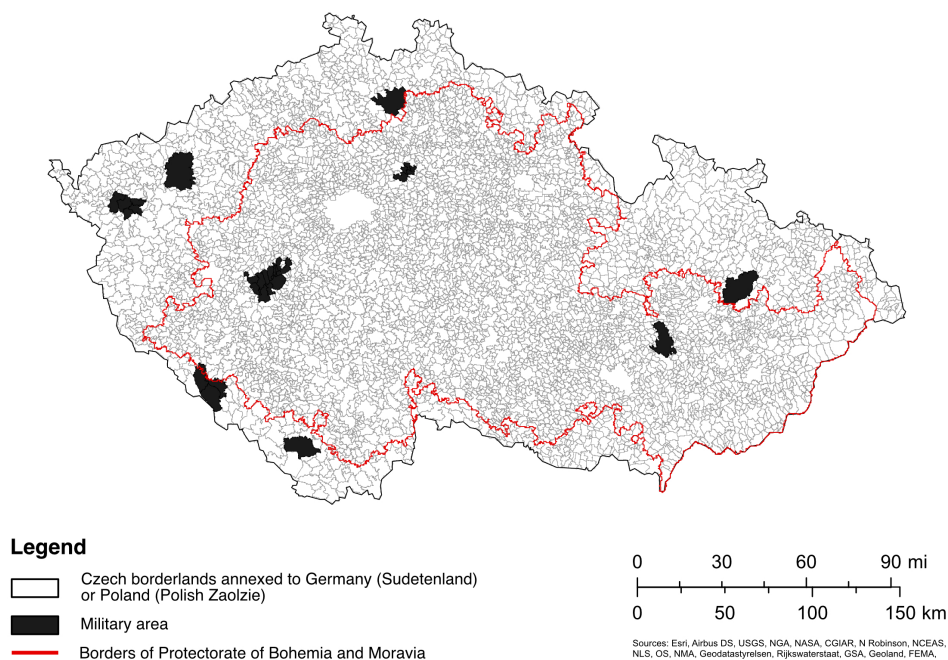


Figure 4.3: Military areas in the Czech Lands, municipalities as of 2011

Austria and the West Germany is also not a concern in the existing literature. But during its creation, a number of villages and solitudes existing near the borders have been displaced, much of which has been demolished, either completely or for the most part (Etnologický ústav AV n.d.). Since this area was guarded by the army, there was also a very limited access (Kubisa 2010). Therefore, the forced abandoning of these areas is expected to have local long-run effects. Thus, the presence of the former border zone in the data could also bias the results of the Germans' expulsion. The border zone extended about 12-15 km into the Czech territory from the state border, and then, after 1964 it was reduced to a strip of territory about 2-3 km wide. Therefore, in the data the border zone is represented by an area that extends 10 km into the Czech territory from the state border as Figure 4.4 illustrates – in light grey are the municipalities that are located in the border zone and, thus, excluded from the sample. Different widths of the border zone are also used to check if the estimates change significantly. So, results of regressions with the border zone that extends 12 km and then also 15 km into the Czech territory from the state border are in the Appendix in Table A.3 and Table A.4, respectively, and the estimates are consistent with the ones when the 10 km wide border zone is

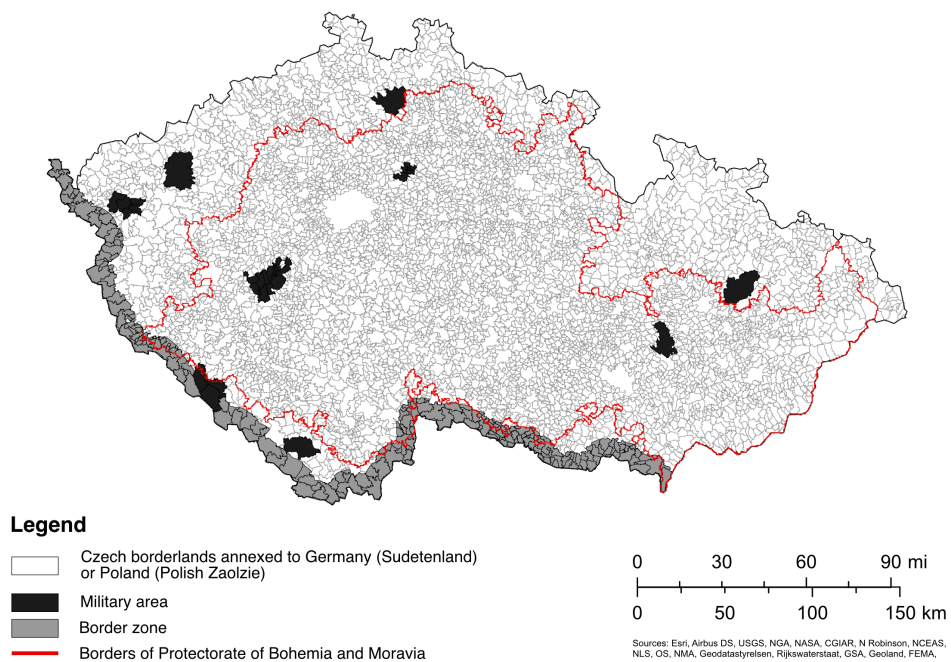


Figure 4.4: Military areas and the Border zone in the Czech Lands, municipalities as of 2011

used.

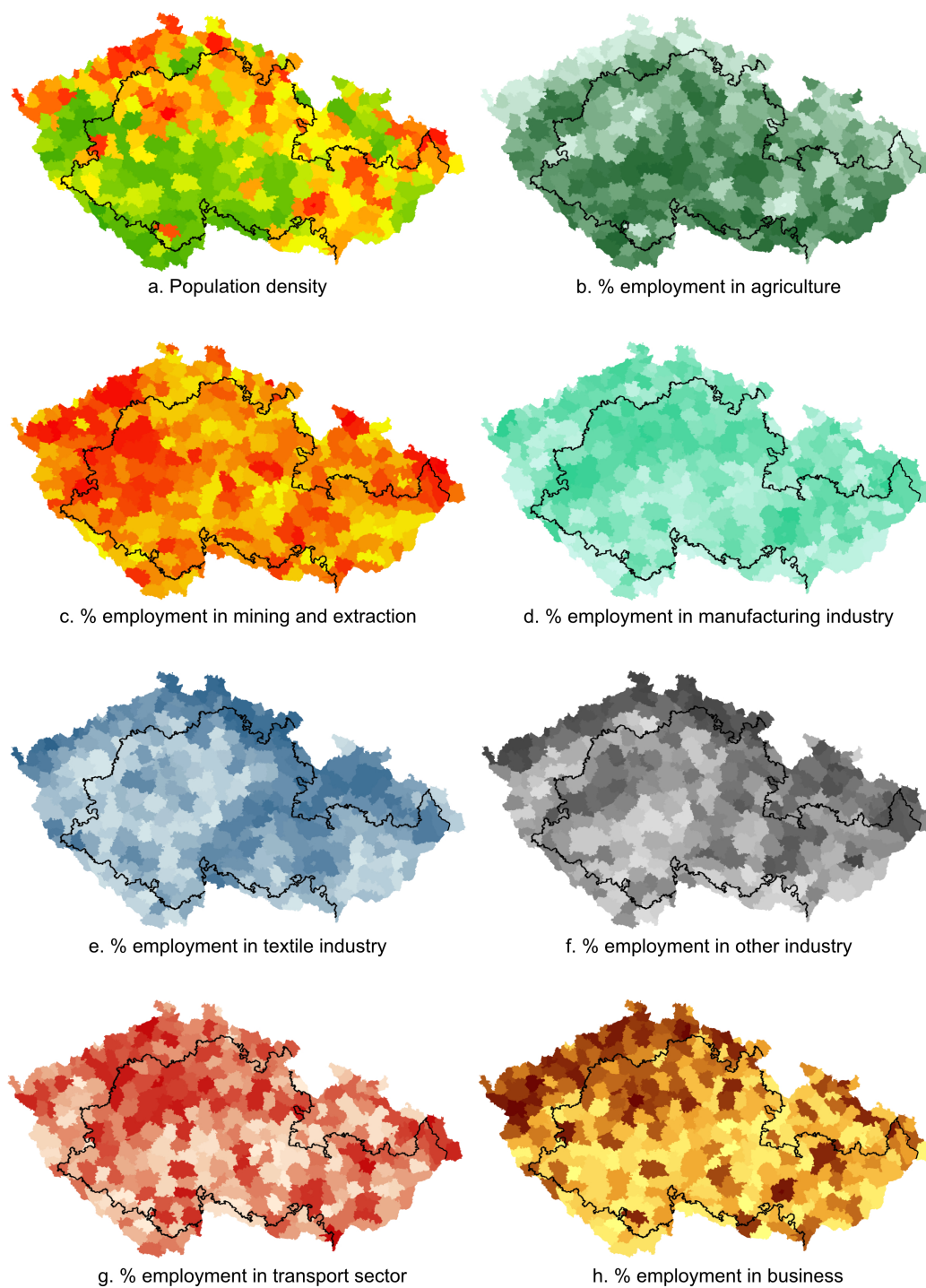
4.2.4 Spatial Heterogeneity and Clustering Standard Errors

To account for spatial heterogeneities that may be present in the data and to prevent it from having an effect on the final results, other covariates are also used in the analysis as the score variable alone might not be sufficient enough to overcome the spatial heterogeneities. According to Testa (2020), in the Czech lands, attributes of geographic characteristics that do not vary over time (i.e. they are fixed) of some regions change sizably or frequently as one moves to other regions. To name them, these characteristics include elevation, ruggedness, precipitation, temperature, and rivers (measured as km per km²). So, since they vary across space, there is a need to control for them in the econometric models. This is done by including a set of geographic characteristics listed above into the regression models in this thesis.

Then, a set of border segment fixed effects is also included in the models to capture unobservable characteristics that vary within regions around the MAL. Testa (2020) provided evidence for the importance of inclusion of such set into the models by showing heatmaps of population density and % employment in various sectors (such as agriculture, business, industry) in 1930. By gathering the individual heatmaps from Testa (2020) I create Figure 4.5, where all the portrayed economic indicators show at least some degree of variation in its intensity across the MAL.

Since some variables have observations on municipal-level and some are available only on more aggregated levels (judicial or political), different segments are used for different estimations, as in Testa (2020). Municipalities are divided into 50 segments that are approximately 50 km long, judicial districts into 24 segments, each being about 100 km long, and political districts into 16 segments, each approximately 150 km long. But the more segments are used the lower is the variation, therefore, 8 long border segments are then used for all regressions as a robustness check.

In regressions that use the municipal-level data (for example the main model described earlier in this chapter), there is a need to also include fixed effects for districts. That is because districts had governing power (control of public goods) over the municipalities in their territory until the end of 2002, when these districts were officially abolished and replaced by new administrative districts of municipalities with extended powers (NUV 2010). Thus, municipalities that belong to one district may have some characteristics that correlate with each other as the district's decisions were applied equally to all public goods in



Source: Testa (2020)

Figure 4.5: Heatmaps, districts as of 2011

these municipalities. The set of district fixed effects controls for these correlations.

The standard errors in this analysis need to be clustered because the observations are at the municipal-level, but each municipality belongs to a district which has the governing power. Thus, as mentioned above, there is a high possibility that characteristics of municipalities are correlated within a district, and if not addressed then the standard errors may not be independent (i.e. violation of the assumption) and also a downward bias in the estimated standard errors might occur (Moulton 1990). For these reasons, the standard errors are clustered at the district-level, as in Testa (2020).

4.3 Balance Testing and Other Concerns

The aim of this subchapter is to address concerns about which factors that have not been addressed by the literature might have an influence on the long-run effects of the Germans' expulsion, and thus may be influencing the final results of the prior literature. Also, this chapter then describes how each of these factors is controlled for in our analysis, so that its effects on the results are mitigated.

4.3.1 Economic Differences between Sudetenland and Interior prior to the Expulsion

This section provides analysis of pre-expulsion differences in economic indicators between the Czech borderlands and the interior around the MAL for which there is available data. This is useful because finding significant differences would suggest there were some other shocks affecting the economic indicators of interest before the Germans' expulsion happened. It would also mean that the estimation of long-run effects of the expulsion would produce unreliable results due to the presence of these historical shocks other than the expulsion. Or, according to Testa (2020), the pre-expulsion differences might interact with the treatment. The re-settlement of the borderlands in the short-run after the expulsion may for example be limited if the sector in which the Sudeten German population worked differed from the sector in which the Czech population worked, or if for example the majority-German areas had very different natural factors. Then, there could be significant long-run differences,

which might, however, be the results of these pre-expulsion differences rather than of the expulsion.

Thus, pre-expulsion factors the data are available for, such as population density, unemployment, or labour force participation, as well as employment in various sectors, are examined to reveal if there were some differences. Table 4.2 shows estimated pre-expulsion differences in economic factors between the borderland and the interior municipalities using RDD with polynomials in distance and latitude-longitude. The only statistically significant difference is in the German population. Specifically, the estimated coefficient shows an increase of more than 60 percentage points (pp) in the 1930 German population as one crosses the MAL from the interior to the borderlands. This again shows that the majority of the then German population lived in the borderlands. But there are no statistically significant differences in the other factors. Even the literacy rate does not yield any difference and is around 98% on average in the borderland as well as in the interior areas. This can be viewed as evidence that, prior to the expulsion, the availability of education was very similar for all children, despite the differences in their nationality. Table 4.3 then shows estimates of the pre-expulsion differences in sectoral composition but none of them is statistically significant. Therefore, because no factors other than the share of the German population is estimated to be different when crossing the MAL, it can be concluded that both analysed sides of the MAL were very similar in terms of economic conditions.

Table 4.2: Economic Differences, 1930.

	% German	Literacy rate	ln Population density	Labour force participation
Linear polynomial in distance to border	65.643*** (7.467)	-0.032 (0.232)	-0.273 (0.255)	-0.645 (1.496)
Linear polynomial in latitude and longitude	68.840*** (4.737)	0.289 (0.175)	-0.165 (0.177)	0.479 (1.100)
Mean dep. var. in interior	1.584 (3.819)	98.37 (0.684)	4.776 (0.750)	45.46 (4.687)
Observations	143	143	143	143
Clusters	89	89	89	89
Border segments	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930
	Unemployment rate	Income per capita	% Taxpayers	Convicts per capita
Linear polynomial in distance to border	-3.581 (2.798)	-0.034 (2.058)	0.737 (0.748)	-0.053 (0.831)
Linear polynomial in latitude and longitude	1.137 (3.029)	-0.127 (1.774)	0.626 (0.629)	0.503 (0.565)
Mean dep. var. in interior	10.09 (10.55)	9.368 (4.880)	5.588 (2.008)	7.262 (2.098)
Observations	83	83	84	142
Clusters	-	-	-	89
Border segments	16	16	16	24
Bandwidth	50 km	50 km	50 km	25 km
Year	1933	1933	1933	1923-1927
	% Roma	% Jewish	km Roads per km ²	km Railways per km ²
Linear polynomial in distance to border	-0.000 (0.001)	-0.068 (0.159)	0.020 (0.016)	0.004 (0.010)
Linear polynomial in latitude and longitude	-0.000 (0.001)	0.104 (0.102)	0.012 (0.011)	0.008 (0.008)
Mean dep. var. in interior	0.00116 (0.00644)	0.190 (0.489)	0.217 (0.0738)	0.0992 (0.0650)
Observations	143	143	227	227
Clusters	89	89	100	100
Border segments	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

Table 4.3: Sectoral Differences, 1930.

	Agricultural sector	Mining and extraction	Metals	Machinery and auto
Linear polynomial in distance to border	3.505 (4.721)	-1.532 (1.988)	0.413 (1.745)	-0.646 (0.632)
Linear polynomial in latitude and longitude	-1.158 (4.104)	-0.950 (1.532)	0.106 (1.084)	-0.723 (0.446)
Mean dep. var. in interior	30.99 (13.03)	3.563 (4.680)	4.533 (3.286)	2.530 (2.116)
	Glass	Textiles	Other industry	Construction
Linear polynomial in distance to border	1.286 (1.814)	-4.729 (2.916)	1.497 (1.288)	-0.331 (0.631)
Linear polynomial in latitude and longitude	0.414 (0.694)	-0.426 (2.138)	0.087 (1.130)	-0.280 (0.525)
Mean dep. var. in interior	0.703 (1.941)	5.865 (9.854)	14.80 (5.108)	6.709 (1.860)
	Transport sector	Finance and insurance	Trade	Other service
Linear polynomial in distance to border	-0.205 (0.829)	-0.091 (0.119)	0.092 (0.850)	-0.356 (1.051)
Linear polynomial in latitude and longitude	0.028 (0.606)	-0.030 (0.097)	0.552 (0.585)	0.968 (0.902)
Mean dep. var. in interior	3.593 (2.200)	0.433 (0.358)	5.179 (1.761)	6.252 (3.330)
Observations	143	143	143	143
Clusters	89	89	89	89
Border segments	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

4.3.2 Pre-Treatment Sorting and Czechs in the Borderlands

When studying the long-run effects of the Germans' expulsion, it is also important to address that, while the Czech interior was populated by Czechs (it was homogenous), some areas in the borderlands were mixed (both Germans and Czechs lived there). This is especially true for borderland areas near the MAL. Having these mixed areas in the analysed sample may cause the results of regressions to be biased. Moreover, this can also affect the results of the pre-expulsion differences discussed in the previous section. It may mean that the similarity of the interior and the borderlands cannot be taken as the similarity of Czechs and Germans.

Therefore, following Testa (2020), economic and sectoral differences are analysed once more but this time using a sample with only homogenous interior areas (municipalities from the interior where only Czechs lived) and majority-German borderland areas where only small number of Czechs lived. Naturally, as it can be seen in Table 4.4, there is an even larger increase in the percentage of German population as one crosses the MAL from the interior to the borderlands. On the contrary, the estimation now reveals some differences income per capita and eligible taxpayers. The small increase in eligible taxpayers discovered Testa (2020) as well, but since it means that in the borderlands there is only 1.6 eligible taxpayers more per 100 persons compared to the interior, this is not very important. On the contrary, the difference in income per capita was not detected by Testa (2020), and thus it arises when the military areas and the border zone are excluded from the regression. Since the data for these regressions are quite aggregated (they are on the district-level) the categorization of districts as the military areas or the border zone is not as precise as in the very analysis of the long-run effects, in which municipal-level data are used. Thus, the reason for this statistically significant difference in income per capita may be due to this lower precision. And finally, Table 4.5 presents results for the sector composition, and it shows no differences between the borderlands and the interior.

So, considering the results of the comparison of the pre-expulsion homogenous districts from the borderlands, where Germans lived, with the interior, where Czechs lived, the two nationalities (and their places) appear very similar as well.

Table 4.4: Economic Differences with Restricted Sample, 1930.

	% German	Literacy rate	ln Population density	Labour force participation
Linear polynomial in distance to border	85.704*** (1.776)	0.097 (0.302)	-0.062 (0.229)	1.375 (1.695)
Linear polynomial in latitude and longitude	88.377*** (1.337)	0.450 (0.217)	0.044 (0.163)	1.375 (1.695)
Mean dep. var. in interior	0.990 (0.951)	98.32 (0.738)	4.669 (0.445)	45.90 (4.691)
Observations	87	87	87	87
Clusters	56	56	56	56
Border segments	8	8	8	8
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930
	Unemployment rate	Income per capita	% Taxpayers	Convicts per capita
Linear polynomial in distance to border	-10.675 (6.567)	8.048*** (2.402)	3.374*** (1.010)	-0.105 (0.811)
Linear polynomial in latitude and longitude	4.489 (4.813)	5.445*** (1.590)	2.455*** (0.761)	1.134* (0.621)
Mean dep. var. in interior	10.97 (12.81)	7.637 (2.718)	4.813 (1.628)	6.779 (1.677)
Observations	48	48	48	87
Clusters	-	-	-	56
Border segments	4	4	4	8
Bandwidth	50 km	50 km	50 km	25 km
Year	1933	1933	1933	1923-1927
	% Roma	% Jewish	km Roads per km ²	km Railways per km ²
Linear polynomial in distance to border	-0.001 (0.002)	0.011 (0.096)	0.007 (0.018)	0.001 (0.010)
Linear polynomial in latitude and longitude	0.000 (0.001)	0.136 (0.089)	0.003 (0.012)	0.008 (0.008)
Mean dep. var. in interior	0.00176 (0.00797)	0.129 (0.309)	0.215 (0.0696)	0.0935 (0.0556)
Observations	87	87	149	149
Clusters	56	56	68	68
Border segments	8	8	8	8
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

Table 4.5: Sectoral Differences with Restricted Sample, 1930.

	Agricultural sector	Mining and extraction	Metals	Machinery and auto
Linear polynomial in distance to border	0.757 (4.066)	-1.008 (0.936)	0.018 (1.634)	0.035 (0.651)
Linear polynomial in latitude and longitude	-7.288** (3.606)	0.272 (1.095)	0.565 (0.797)	0.308 (0.611)
Mean dep. var. in interior	31.20 (11.61)	2.559 (2.257)	4.028 (2.735)	2.199 (2.136)
	Glass	Textiles	Other industry	Construction
Linear polynomial in distance to border	0.631 (2.048)	-8.638* (4.579)	2.817 (2.370)	-0.052 (0.671)
Linear polynomial in latitude and longitude	0.375 (1.591)	-0.732 (4.019)	1.794 (1.819)	0.697 (0.595)
Mean dep. var. in interior	0.761 (2.295)	8.058 (11.53)	15.36 (5.454)	6.119 (1.546)
	Transport sector	Finance and insurance	Trade	Business sector
Linear polynomial in distance to border	0.288 (0.673)	0.144 (0.110)	1.154 (1.179)	1.298 (1.261)
Linear polynomial in latitude and longitude	1.136** (0.531)	0.074 (0.078)	1.089 (0.742)	1.163 (0.807)
Mean dep. var. in interior	3.303 (2.065)	0.387 (0.158)	4.915 (1.185)	5.303 (1.317)
Observations	87	87	87	87
Clusters	56	56	56	56
Border segments	8	8	8	8
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

4.3.3 Pollution in North-West Bohemia

During the communist regime in Czechoslovakia (1948-1989), great emphasis was put on industrialization, and heavy industry with great demands on the energy supply became the centre of the economy. To meet the energy demand, new power plants had to be created. As there were large deposits of lignite in the North-West Bohemia, a total of six lignite-burning power plants were built within this quite small area, near to the mines, to reduce transportation costs. There were also large cities close to the newly built power plants, but the quality of life of the inhabitants was not considered by the government, and even the historical centre of the town Most was relocated in order to expand the mines (Š. Mikula and Pytliková 2021). The government decided to compensate the local inhabitants by only introducing benefits such as free healing stays in the mountains for children and some financial contributions (Glassheim 2005). The burning of lignite was accompanied by production of large amounts of SO₂ emissions, and because of mountains on one side of the North-West Bohemia the air pollution concentrated in this region (Š. Mikula and Pytliková 2021). According to Š. Mikula and Pytliková (2021), from 1970 the districts that had the highest pollution were Teplice, Most, and Chomutov, but districts Ústí nad Labem, Louny, and Litoměřice also had very high levels of pollution.

This pollution problem is relevant to this analysis because it had detrimental effects on the environment and on the local people. From 1960 to 1980, the North-West region had many depopulated villages with empty houses, and the smog situation there was so bad that it complicated the traffic (Glassheim 2005). Furthermore, between 1982 and 1998 the life-expectancy was lower in the polluted North-West Bohemia compared to cleaner South Moravia (Kotěšovec et al. 2000), and there was even a higher risk of intrauterine growth retardation in the polluted areas (Dejmek et al. 1999). Nevertheless, the communist regime tried to fight the depopulation by introducing benefits that aimed at attracting new workers (especially high-skilled workers) to these areas. It was not until the fall of the communist regime in Czechoslovakia in 1989 that the new government committed itself to improving the environment and the lives of the people of northern Bohemia (Š. Mikula and Pytliková 2021).

Given all this information, it is highly possible that the economic development in the North-West Bohemian region might have been disrupted by the consequences of the pollution. If it is so, then having this region in the analysis of long-run effects of the Germans' expulsion could bias the results. Thus,

as a robustness check, the municipalities from Teplice, Most, Chomutov, Ústí nad Labem, Louny, and Litoměřice districts are then also excluded to reveal whether they do influence the findings or not.

4.3.4 1945 US-Soviet Demarcation Line

As described in more detail in Chapter 2, early after the end of WWII there was the period of Wild expulsion, during which the treatment of Germans in the area that was liberated by the Soviet army differed from how Germans were treated by the US forces on the western side of the demarcation line. While the US army prevented violent acts on the German population, the Soviet army allowed the situation to ‘run its course’. For that reason, the economic development could have been disrupted by this Wild expulsion in the Soviet army zone, but not in the US army zone. Thus, whether a municipality lies in the area of the former US- or the Soviet-liberation zone can also have some effect on the differences in the long-run development in those areas. If there is a significant difference when crossing the 45DL in the borderlands compared to the interior, then it may bias the results of analysing the Germans’ expulsion. Thus, this section describes a method that is used to inspect whether there are some variations in the long-run effects of the Germans’ expulsion in areas which used to be part of the US army zone, and in those that used to be part of the Soviet army zone. To do so, a very similar form of the equation described earlier in this chapter is used. Only the treatment variable is different because now we are interested in how the dependent variable changes when we cross from the Soviet army liberation zone to the US army liberation zone.

As outlined above, the treatment variable is now a dummy variable that has a value of 1 for municipalities (or districts) located in the US army zone and a value of 0 for municipalities (or districts) located in the Soviet army zone. The condition of assigning these values is the same as in the case of the borderlands dummy variable – if 50% or more of a municipality’s (or district’s) area lies on the West side of the 45DL then it is identified as being located in the US zone, and if 50% or more of a municipality’s (or district’s) area lies on the East side then it is identified as being located in the Soviet zone. Then a variable ‘dl cross’ is constructed, which is similar to the ‘no-overlap’ variable as it deals with the municipalities (or districts) that the 45DL crosses. Specifically, ‘dl cross’ variable equals 1 if less than 95% of a municipality’s (or district’s) area

lies in either of the two forces (i.e. if the 45DL overlaps the municipality's area).

So, the municipal-level equation now has the following form:

$$y_{mdb} = \alpha + \beta InUSZone_m + f(\text{geographic location}_m) + \gamma G'_m + D_d + \epsilon_{mdb},$$

where y_{mdb} is the dependent variable for municipality m in district d . Then, α is the intercept, β is the coefficient of interest and it is showing the slope of the treatment variable, and $InUSZone_m$ is the binary treatment variable for a municipality m with value of 1 for municipalities located in the US army liberation zone and value of 0 for municipalities located in the Soviet army liberation zone. This β measures the (local average) treatment effect at the 45DL. Next, $f(\text{geographic location}_m)$ is a score variable for municipality m . The following term, G_m , represents a set of geographic characteristics of a municipality m . D_d is a set of district fixed effects for district d , and ϵ_{mdb} is the error term for municipality m in district d .

The score variables remain the same – first one being the polynomial in distance to border and the second one being the polynomial in latitude and longitude. But the border is now the 45DL, so the one-dimensional score variable here is a polynomial in distance to the 45DL. For consistency, the baseline bandwidth around the 45DL is 25 km, so that it is the same as the one used around the MAL. Nevertheless, 10 km and 50 km bandwidths are again used as robustness checks as well as other degrees of polynomials for some dependent variables. Then, municipalities in which there were military areas and the border zone are also excluded from the analysis to avoid potential bias of the estimates (for reasons previously described). Excluded are also municipalities that are south of Žernovice because there ethnic divides correspond with the 45DL according to Grossmann, Jurajda, and Roesel (2021). The final sample of the municipalities included in the analysis is shown in Figure 4.6.

The set of geographic characteristics (elevation, ruggedness, precipitation, temperature, and rivers measured as km per km²) is included in this model as well to deal with spatial heterogeneities. Border segment fixed effects are not used here as the sample around 45DL is quite small, so there may not be as much variation to begin with and dividing the sample into segments would lower the variation even more. Also, in the heatmaps in Figure 4.5 it can be

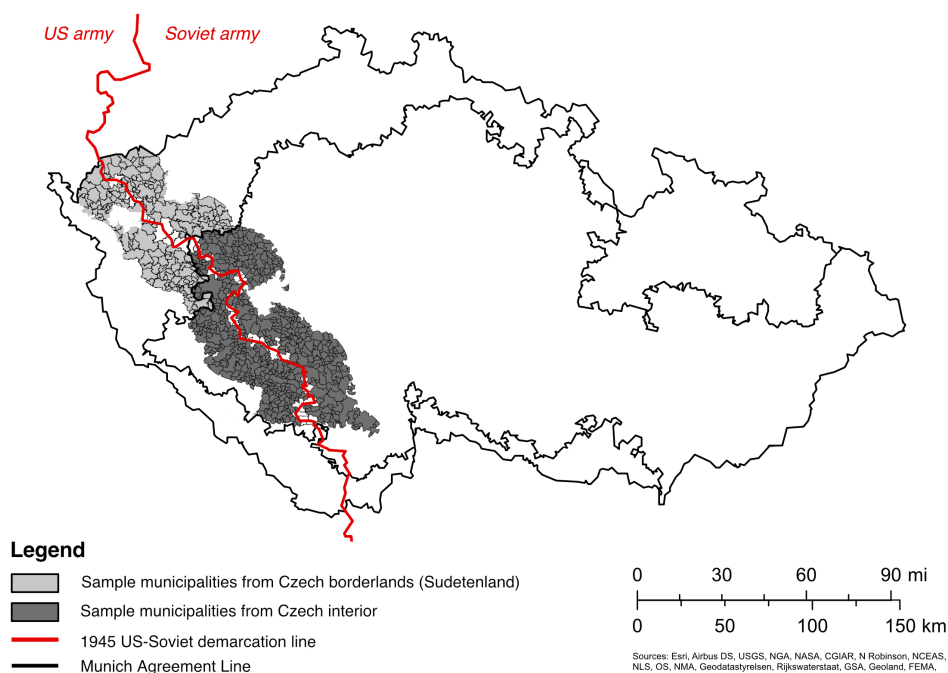


Figure 4.6: Sample municipalities, 2011

seen that on both sides of the 45DL the economic indicators had roughly the same intensity (especially when we consider the areas that are no more than 25 km far from the 45DL). On the other hand, the district fixed effects are included in this model to control for the governing power they had over the municipalities until 2002. The need for clustering the standard errors prevails as the same municipal-level data are used. So, the standard errors are again clustered at the district-level.

Finally, as the new sample of municipalities around the 45DL now contains municipalities that are not included in the sample of municipalities around the MAL, the pre-expulsion differences need to be re-examined. Since the municipalities around 45DL are now divided into the borderland group, and the interior group and these two groups will be used in separate regressions, the differences between the US and the Soviet zone in the borderlands and in the interior are examined separately. If there would be some significant differences between the US and the Soviet zone in the sample that is used here, then the estimated effects would not be reliable as the driver of those effects may not be the Germans' expulsion.

The most granular data that are available for the pre-expulsion period are

only at the district-level (judicial or political, depending on the variable, with 330 or 151 districts, respectively), and the sample of 25 km from the 45DL and division of the sample into the borderland and the interior group decreases the number of observations rapidly. Therefore, an estimation of the pre-expulsion differences using RDD is not suitable here. So instead, means of the dependent variables available are compared. Table 4.6 shows means and standard deviations of economic factors in 1930 in all four groups. As explained, we focus only on the differences within the borderlands and within the interior, and then compare them. The table shows that in the borderlands the means for most of the pre-expulsion factors are almost the same except for labour force participation, income per capita, and convicts per capita. Nevertheless, the difference between the mean values of these three factors is very small. Then in the interior, the two zones also appear to not have much difference in almost all means of the pre-expulsion economic factors, only labour force participation and income per capita do but the difference among them is even smaller than the one in the borderlands.

Next, Table 4.7 presents means and standard deviations of variables indicating a 1930 percentage employment in individual sectors in both zones in the borderlands and in the interior. The only considerable difference in the borderlands is in the agricultural sector which shows a higher mean percentage employment in the US zone. But this difference is most probably driven by variations in geographical conditions, which are controlled for in the econometric model. In the interior areas, there are no substantial differences between the US and the Soviet zones in any of the sectors. Thus, given all this information from the examination of the pre-expulsion differences, we can conclude that the US and the Soviet zone in the borderlands as well as in the interior were very similar, and we can, thus, continue with the analysis.

Finally, when using the municipal-level model to analyse the long-run effects the regressions produce estimates of how a dependent variable changes if one crosses the 45DL from the former Soviet-liberation zone to the former US-liberation zone. Nevertheless, the aim of this thesis is not to study the long-run impacts of the existence of the 45DL, but only to control for the possible effects the division of the Czech lands into two liberation zones after the WWII might have had on the studied areas (and thus, might be causing some bias in the results of long-run effects of the Germans' expulsion when the full sample of municipalities around the MAL is used). Therefore, the sample is divided into two groups – one is a group of municipalities that are located in the borderlands

Table 4.6: Mean and Standard Deviation of Economic Factors in Sample, 1930.

	Borderlands		Interior	
	US zone (1)	Soviet zone (2)	US zone (3)	Soviet zone (4)
% German	92.190 (9.586)	92.939 (7.819)	0.389 (0.250)	0.653 (0.489)
Literacy rate	99.024 (0.450)	98.905 (0.353)	98.288 (0.354)	98.540 (0.210)
ln Population density	4.233 (0.389)	4.759 (0.730)	4.374 (0.120)	4.126 (0.157)
Labour force particip.	46.005 (4.435)	50.729 (5.914)	42.835 (4.526)	44.963 (4.333)
Unemployment rate	25.238 (21.985)	24.879 (14.784)	4.572 (0.999)	4.795 (2.523)
Income per capita	8.809 (1.982)	11.360 (4.865)	6.927 (0.708)	8.032 (3.156)
% Taxpayers	5.460 (0.721)	6.408 (1.879)	6.237 (1.785)	5.119 (1.630)
Convicts per capita	8.672 (1.624)	6.574 (1.893)	6.458 (0.758)	6.282 (1.432)
% Roma	0.000 (0.000)	0.002 (0.006)	0.000 (0.000)	0.001 (0.003)
% Jewish	0.057 (0.064)	0.175 (0.265)	0.035 (0.041)	0.053 (0.087)
km Roads per km ²	0.211 (0.018)	0.244 (0.040)	0.229 (0.062)	0.195 (0.029)
km Railways per km ²	0.067 (0.054)	0.109 (0.040)	0.080 (0.036)	0.073 (0.031)
Max. observations	11	9	12	12
Bandwidth	25-50 km	25-50 km	25-50 km	25-50 km
Year	1930	1930	1930	1930

Notes: Standard deviations are in parenthesis. In all calculations military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

Table 4.7: Mean and Standard Deviation of Employment in Sectors in Sample, 1930.

	Borderlands		Interior	
	US zone (1)	Soviet zone (2)	US zone (3)	Soviet zone (4)
Agricultural sector	29.776 (12.157)	20.004 (20.095)	39.766 (5.786)	41.022 (7.829)
Mining and extraction	3.789 (2.949)	6.240 (5.851)	3.421 (2.874)	3.669 (2.443)
Metals	3.624 (2.420)	4.748 (4.117)	3.260 (1.082)	2.906 (0.976)
Machinery and auto	4.654 (6.409)	1.453 (0.689)	2.011 (1.595)	1.163 (0.314)
Glass	0.834 (1.831)	0.176 (0.291)	0.069 (0.139)	0.000 (0.000)
Textiles	8.127 (12.305)	10.634 (16.364)	1.508 (2.477)	0.230 (0.214)
Other industry	12.250 (2.009)	19.775 (17.481)	11.977 (3.319)	13.796 (5.826)
Construction	8.528 (2.590)	6.564 (1.527)	8.142 (2.635)	6.268 (0.910)
Transport sector	2.359 (0.529)	3.008 (1.345)	2.723 (1.239)	2.625 (0.797)
Finance and insurance	0.275 (0.127)	0.477 (0.355)	0.348 (0.087)	0.380 (0.141)
Trade	6.021 (1.102)	7.201 (4.205)	3.823 (0.526)	3.888 (0.544)
Other service	5.943 (1.606)	5.992 (2.622)	5.054 (0.693)	6.827 (4.116)
Max. observations	11	9	12	12
Bandwidth	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930

Notes: Standard deviations are in parenthesis. In all calculations military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are districts.

and the other is a group of municipalities from the interior (thus, there are four studied groups in total – US army zone in the borderlands, Soviet army zone in the borderlands, US army zone in the interior, and Soviet army zone in the interior). So, all regressions are run twice, once with the borderland municipalities and the second time with the interior municipalities.

Chapter 5

Empirical Results

5.1 Results

This chapter presents results of analysing long-run effects of Germans' expulsion from Czechoslovakia. First, results of estimations using RDD with the MAL as the cut-off point are shown. And then, this chapter also provides results of the analysis of the former US and Soviet army zones, for which RDD is used as well but this time using the 45DL as the threshold.

5.1.1 RDD Using the Munich Agreement Line

The baseline results of this thesis are in Table 5.1, Table 5.2, and Table 5.3, and they show evidence of long-run effects that the Germans' expulsion from Czechoslovakia has on the labour market, on some sectors, and also on the educational attainment, respectively, with military areas and the border zone being excluded.

In Table 5.1 in column (1) are the estimated coefficients of the change in population density as one crosses the MAL. First row shows a result of a regression with linear polynomial in distance, and it is statistically significant. Specifically, it shows that the population in the borderland municipalities is 39.8% lower on average compared to the nearby municipalities in the interior (this is based on $e^{0.335} - 1 = 0.398$). Then to check the completeness a regression with linear polynomial in latitude and longitude is run, and the resulting coefficient in fourth row confirms the negative effect with a slightly lower result of a 28.2% decrease of population density on average as one crosses the MAL from the interior to the borderlands (this is based on $e^{0.248} - 1 = 0.282$). Second row then replaces the linear polynomials with quadratic polynomials. Even in these

specifications, the coefficient of population density stays statistically significant at the conventional levels, and the resulting effect is still a drop of about 30% on average when crossing the MAL. But then in the third row, where cubic polynomials are used, the effect is ambiguous as the distance specification does not produce any statistically significant effect, but the latitude-longitude specification does. Then, Column (2) of Table 5.1 shows results for labour force density and they are very similar to those of population density, which is not surprising as these two variables are interconnected to some degree.

The last column of Table 5.1, column (3), presents estimated differences in the unemployment rate. In the distance specification when using the linear polynomial, a statistically significant increase of 2.6 percentage points (pp) on average is estimated when crossing the MAL from the interior municipalities to the borderland ones. This is also supported by a result of the regression with the two-dimensional latitude and longitude polynomial. Also, the statistically significant increase in the unemployment rate is highly robust when the quadratic polynomial is used in both specifications, and this is true even for the use of the cubic polynomial, which is already a relatively demanding order of polynomial.

Table 5.2 shows estimates of differences in sectoral composition when crossing the MAL, which reveal a lower employment in skill-intensive sectors in the borderland areas compared to the nearby interior areas. The largest difference is in the employment in financial and insurance sector, where it is on average 28.3% lower in the borderlands compared to the interior. Then, the employment in communications and in healthcare decrease on average by 24.8% and 21.4%, respectively, as one crosses the MAL from the interior to the borderlands. Then also the employment in education is different in the borderland municipalities when compared with the nearby interior ones, specifically, there is a drop of 23% on average. And lastly, the sector of auto trade and repair also experiences a change in the employment when crossing the MAL in the borderlands the employment in this sector is on average 10.9% lower than in the interior.

And finally, Table 5.3 provides the results of analysing differences in educational attainment between the treatment and the control groups (i.e. between the borderlands and the interior). The estimates show that people in the borderland municipalities are less educated than those in the interior municipalities. Specifically, the percentage of people with primary education (or less) increases by 5.4 pp on average when crossing the MAL from the interior

Table 5.1: Long-run Differences in Labour Market at MAL, No Military Areas and Border Zone, 2011.

	ln Population density	ln Labour force density	Unemployment rate
	(1)	(2)	(3)
<i>Polynomials in distance to border</i>			
Linear polynomial	-0.335*** (0.105)	-0.343*** (0.108)	2.595*** (0.616)
Quadratic polynomial	-0.270** (0.118)	-0.277** (0.119)	1.498** (0.568)
Cubic polynomial	-0.132 (0.124)	-0.148 (0.126)	2.277*** (0.534)
<i>Polynomials in latitude and longitude</i>			
Linear polynomial	-0.248*** (0.093)	-0.251** (0.096)	3.343*** (0.587)
Quadratic polynomial	-0.329*** (0.091)	-0.331*** (0.094)	3.303*** (0.629)
Cubic polynomial	-0.328*** (0.094)	-0.335*** (0.097)	3.223*** (0.650)
Mean dep. var. in interior	4.042 (0.873)	3.302 (0.898)	10.50 (4.815)
Observations	3845	3845	3845
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and polynomials in distance to the MAL interacted with the treatment or polynomials in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table 5.2: Long-run Differences in Sectoral Composition at MAL, No Military Areas and Border Zone, 2011.

	Agricultural sector	Industry	Construction	Transport sector
	(1)	(2)	(3)	(4)
Linear polynomial in distance to border	-0.398 (0.457)	-0.453 (0.933)	-0.416 (0.282)	0.065 (0.266)
Linear polynomial in latitude and longitude	-0.721 (0.467)	-1.015 (0.800)	-0.043 (0.225)	0.146 (0.280)
Mean dep. var. in interior	7.582 (6.339)	26.66 (7.476)	7.218 (3.183)	5.304 (2.660)
	Finance and insurance	Hospitality	Auto trade and repair	Public
	(5)	(6)	(7)	(8)
Linear polynomial in distance to border	-0.399*** (0.078)	0.159 (0.259)	-0.866*** (0.304)	-0.056 (0.294)
Linear polynomial in latitude and longitude	-0.386*** (0.070)	0.255 (0.225)	-0.858*** (0.267)	-0.320 (0.249)
Mean dep. var. in interior	1.412 (1.277)	2.113 (1.580)	7.971 (3.298)	4.464 (2.367)
	Communi- cations	Education	Healthcare	Other service
	(9)	(10)	(11)	(12)
Linear polynomial in distance to border	-0.322*** (0.089)	-0.972*** (0.207)	-1.002*** (0.258)	-0.208 (0.195)
Linear polynomial in latitude and longitude	-0.3949*** (0.087)	-0.862*** (0.173)	-0.718*** (0.242)	-0.345** (0.158)
Mean dep. var. in interior	1.299 (1.275)	4.219 (2.297)	4.686 (2.801)	4.530 (2.653)
Observations	3845	3845	3845	3845
Clusters	70	70	70	70
Border segments	50	50	50	50
Bandwidth	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011
Military areas and border zone	Excluded	Excluded	Excluded	Excluded

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

to the borderlands, while the percentage of people with secondary and tertiary education decreases by 4.2 and 2.2 pp on average, respectively.

Table 5.3: Long-run Differences in Educational Attainment at MAL, No Military Areas and Border Zone, 2011.

	% Primary education or less	% Secondary education	% Tertiary education
	(1)	(2)	(3)
Linear polynomial in distance to border	5.384*** (0.612)	-4.145*** (0.537)	-2.187*** (0.404)
Linear polynomial in latitude and longitude	5.269*** (0.508)	-3.822** (0.456)	-2.410*** (0.420)
Mean dep. var. in interior	20.72 (4.935)	66.96 (4.795)	8.741 (3.907)
Observations	3845	3845	3845
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Military areas and border zone	Excluded	Excluded	Excluded

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

5.1.2 RDD Using the 1945 Demarcation Line

As it is described in Chapter 4, the results from the previous section could be biased due to the situation during the period of Wild expulsion, when in the areas that were liberated by the US army the unorganized and violent acts on German population have been prevented, while in the areas controlled by the Soviet army this was not the case. The results in this section, thus, show whether there are some long-run effects of this historical event by estimating differences between municipalities which lie in the areas of former US and Soviet zones (i.e. on the West and East sides of the 45DL). If there is an effect of the presence of US forces in the borderlands during the expulsion on the present-day economic situation, then there should not be any effect when crossing the 45DL in the Czech interior.

Table 5.4 shows estimated differences for population density, labour force density, and unemployment. In the borderlands, the US army zone is estimated to have higher population density as well as labour force density compared to the Soviet army zone. But almost the same increase in population and labour force density is present also when one crosses the 45DL in the interior. Then, the coefficient of unemployment rate is statistically insignificant in the borderlands and the same holds also for the interior. Thus, since the results are practically the same in the borderlands as well as in the Czech interior, the presence of the US forces in the South-West borderlands does not seem to have had any effect on these three factors.

Then, in Table 5.5, the only statistically significant effect is in the employment in public administration, defence, and social services and it is positive. Nevertheless, this coefficient is statistically significant and positive in the interior as well. In the interior there are differences also in some other sectors. One of them is the agricultural sector which shows a decrease of 35.5% on average in the employment as one crosses the 45DL from the Soviet zone into the US zone (this is based on $\frac{3.821}{10.77} \cdot 100 = 35.5\%$). Next, in the interior the employment in transportation and post drops by 15% on average when we go from the Soviet to the US zone. Then in the interior the employment in education and healthcare is higher in the US zone by 32% and 45% on average, respectively. And finally, the employment in other services which include real estate, legal and accounting services, architecture, research & development, advertising, and others, also increases when crossing the 45DL from the Soviet into the US zone in the interior municipalities.

Table 5.4: Long-run Differences in Labour Market at 45DL, 2011.

	Borderlands		
	ln Population density	ln Labour force density	Unemployment rate
	(1)	(2)	(3)
Linear polynomial in distance to border	0.564** (0.169)	0.545*** (0.155)	3.591 (3.478)
Linear polynomial in latitude and longitude	0.579*** (0.066)	0.545*** (0.063)	2.095 (3.478)
Mean dep. var. in interior	3.643 (1.248)	2.914 (1.267)	15.06 (6.542)
Observations	159	159	159
Clusters	9	9	9
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
	Interior		
	ln Population density	ln Labour force density	Unemployment rate
	(4)	(5)	(6)
Linear polynomial in distance to border	0.473** (0.166)	0.493** (0.165)	-0.279 (0.596)
Linear polynomial in latitude and longitude	0.495*** (0.158)	0.526*** (0.162)	0.022 (0.629)
Mean dep. var. in interior	3.489 (0.890)	2.739 (0.926)	9.165 (4.983)
Observations	532	532	532
Clusters	14	14	14
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In US zone'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the 45DL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Municipality Vřesová is also excluded as it is an outlier in the unemployment rate. Observed units are municipalities. Link to discussion: [Section 6](#).

Table 5.5: Long-run Differences in Sectoral Composition at 45DL, 2011.

	Borderlands				Interior			
	Agricultural sector	Industry	Construction	Transport sector	Agricultural sector	Industry	Construction	Transport sector
	(1)	(2)	(3)	(4)	(13)	(14)	(15)	(16)
Linear polynomial in distance to border	2.173 (2.296)	-4.294* (2.082)	0.031 (0.360)	-1.178 (0.796)	-3.821*** (1.130)	-0.713 (1.588)	-0.020 (0.506)	-0.808** (0.343)
Linear polynomial in latitude and longitude	1.345 (1.857)	-3.378 (2.031)	-0.340 (0.239)	-1.258 (1.398)	-4.573*** (0.735)	0.778 (1.934)	0.140 (0.529)	-0.850*** (0.260)
Mean dep. var. in interior	5.429 (5.712)	19.57 (8.276)	6.664 (3.252)	4.653 (2.955)	10.77 (8.167)	25.14 (8.454)	6.993 (3.404)	5.357 (2.943)
	Finance and insurance	Hospitality	Auto trade and repair	Public	Finance and insurance	Hospitality	Auto trade and repair	Public
	(5)	(6)	(7)	(8)	(17)	(18)	(19)	(20)
Linear polynomial in distance to border	-0.282 (0.464)	-0.717 (1.238)	0.039 (1.143)	1.420** (0.564)	0.422** (0.159)	0.459 (0.318)	0.103 (0.744)	1.099*** (0.309)
Linear polynomial in latitude and longitude	-0.207 (0.457)	-0.898 (1.349)	0.450 (1.464)	1.897*** (0.522)	0.230 (0.212)	0.581* (0.307)	-0.061 (0.571)	0.825* (0.456)
Mean dep. var. in interior	1.136 (1.035)	4.345 (3.536)	7.203 (3.143)	4.820 (2.275)	1.211 (1.195)	2.302 (1.960)	8.169 (3.985)	4.520 (2.456)
	Communica-tions	Education	Healthcare	Other service	Communica-tions	Education	Healthcare	Other service
	(9)	(10)	(11)	(12)	(21)	(22)	(23)	(24)
Linear polynomial in distance to border	0.133 (0.109)	-0.119 (0.684)	0.576 (0.938)	0.760 (0.712)	0.154 (0.288)	1.147** (0.458)	2.060** (0.699)	1.483* (0.706)
Linear polynomial in latitude and longitude	0.103 (0.237)	-0.014 (0.757)	0.560 (0.759)	0.833 (0.647)	0.120 (0.249)	1.000** (0.392)	1.648* (0.848)	1.37** (0.595)
Mean dep. var. in interior	0.680 (0.775)	4.165 (2.140)	4.927 (3.205)	4.246 (2.218)	1.318 (1.573)	3.569 (2.525)	4.572 (4.817)	4.429 (2.797)
Observations	159	159	159	159	532	532	532	532
Clusters	9	9	9	9	14	14	14	14
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In US zone'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the 45DL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Municipality Vřesová is also excluded as it is an outlier in the unemployment rate. Observed units are municipalities. Link to discussion: [Section 6](#).

Lastly, changes in the educational attainment are shown in Table 5.6, and it can be seen that in the borderlands there are some differences between the two zones. First, there are no differences in the percentage of people that have no more than primary education and the same is true for the percentage of people with tertiary education. On the contrary, the percentage of people with secondary education drops by 3.7 pp on average when crossing from the Soviet zone to the US zone in the borderlands. Then in the interior, there is also no statistically significant effect on the percentage of people with primary or less education. Then, though only with a small statistical significance, the percentage of people with secondary education is 3.3 pp lower in the US zone compared to the Soviet zone, which is almost the same effect as when crossing the 45DL in the borderlands. Lastly, in the interior, there is an increase of almost 2 pp on average in the highly educated population (people with tertiary education).

Table 5.6: Long-run Differences in Educational Attainment at 45DL, 2011.

	Borderlands		
	% Primary education or less	% Secondary education	% Tertiary education
	(1)	(2)	(3)
Linear polynomial in distance to border	2.717 (1.875)	-3.646** (1.163)	-0.080 (1.319)
Linear polynomial in latitude and longitude	2.253 (2.209)	-3.344** (1.146)	-0.284 (1.335)
Mean dep. var. in interior	24.68 (5.735)	62.18 (4.386)	6.836 (4.075)
Observations	149	149	149
Clusters	9	9	9
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
	Interior		
	% Primary education or less	% Secondary education	% Tertiary education
	(4)	(5)	(6)
Linear polynomial in distance to border	0.807 (0.985)	-3.303* (1.575)	1.853** (0.798)
Linear polynomial in latitude and longitude	0.601 (0.993)	-2.606* (1.270)	1.628** (0.661)
Mean dep. var. in interior	20.93 (5.531)	67.48 (5.946)	7.879 (3.585)
Observations	532	532	532
Clusters	14	14	14
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Notes: All reported estimates are coefficients of the variable ‘In US zone’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the 45DL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague, Polish Zaolzie and municipalities with % Secondary education higher than 70 or lower than 50 are excluded. Observed units are municipalities. Link to discussion: [Section 6](#).

5.2 Robustness

Before any conclusions are made, robustness of the results needs to be checked by using alternative specifications. First, different bandwidths around the MAL are used while applying the same empirical approach, i.e. the specific form of the geographic RDD described in Chapter 4. Table 5.7 shows results of a regression with 10 km wide bandwidth around the MAL. This small bandwidth decreases the number of observations significantly, and the results are slightly lower but very consistent with those of using the 25 km bandwidth. Only, the employment in auto trade and repair sector lost its statistical significance, and the percentage of people with tertiary education is even higher now when one crosses the MAL from the interior into the borderlands. Next, in Table 5.8 a wider bandwidth of 50 km around the MAL is employed. The results in this table are almost the same or slightly lower than the original ones (when the 25 km bandwidth is used). The only exception is the unemployment rate, which now shows an even higher increase of 3.2 pp on average when crossing the MAL from the interior to the borderlands. Overall, the described changes in the results when the different bandwidths are used do not reveal anything concerning as there are only very minor variations in the estimates.

Then, Table 5.9 shows regression results when only 8 long border segments are used instead of the 50 that are used in the main regressions of this thesis. As it is explained in Chapter 4, the more segments are used in the analysis the lower is the variation in the sample. For this reason, there is a need to check whether the results change if only low number of border segments is used, and in our case, we use the 8 long border segments in the regressions. So, when we look at the results, the estimates in Table 5.9 are nearly identical to the ones produced by the regressions with 50 border segments from Section 5.2. Most of the estimated coefficients in Table 5.9 are only slightly lower compared to the ones in Table 5.1, Table 5.2, and Table 5.3, but their direction and statistical significance are unchanged.

Final robustness check is concerned with possible effects of high levels of air pollution that occurred in the North-West Bohemia during the communist regime (1948-1989). As described in Chapter 4, whether the presence of these previously highly polluted areas affects the results of our analysis of long-term effects of expulsion of Germans can be explored by excluding those areas from the regressions. Specifically, the municipalities from Teplice, Most, Chomutov, Ústí nad Labem, Louny, and Litoměřice districts are excluded. Table 5.10

Table 5.7: Long-run Differences at MAL, 10 km bandwidth, 2011.

	ln Pop. density (1)	ln Labour force den. (2)	Unempl. rate (3)	Fin. and insurance (4)
Linear polynomial in distance to border	-0.217* (0.117)	-0.225* (0.120)	2.084*** (0.526)	-0.267*** (0.076)
Linear polynomial in latitude and longitude	-0.274*** (0.092)	-0.278*** (0.095)	3.106*** (0.569)	-0.313*** (0.064)
Mean dep. var. in interior	4.047 (0.839)	3.308 (0.865)	10.51 (4.797)	1.363 (1.210)
	Auto trade and repair (5)	Communic. (6)	Education (7)	Healthcare (8)
Linear polynomial in distance to border	-0.270 (0.281)	-0.233*** (0.083)	-0.824*** (0.262)	-0.739** (0.294)
Linear polynomial in latitude and longitude	-0.758*** (0.271)	-0.347*** (0.072)	-0.799*** (0.192)	-0.784*** (0.260)
Mean dep. var. in interior	7.858 (3.179)	1.211 (1.163)	4.349 (2.312)	4.794 (2.665)
	Other service (9)	% Primary educ. or less (10)	% Secondary education (11)	% Tertiary education (12)
Linear polynomial in distance to border	0.129 (0.175)	4.240*** (0.679)	-3.470*** (0.609)	-1.604*** (0.361)
Linear polynomial in latitude and longitude	-0.243* (0.143)	4.908*** (0.549)	-3.623*** (0.438)	-2.165*** (0.367)
Mean dep. var. in interior	4.291 (2.393)	20.63 (4.881)	67.19 (4.676)	8.628 (3.682)
Observations	1947	1947	1947	1947
Clusters	53	53	53	53
Border segments	50	50	50	50
Bandwidth	10 km	10 km	10 km	10 km
Year	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table 5.8: Long-run Differences at MAL, 50 km bandwidth, 2011.

	ln Pop. density (1)	ln Labour force den. (2)	Unempl. rate (3)	Fin. and insurance (4)
Linear polynomial in distance to border	-0.292*** (0.105)	-0.294*** (0.107)	3.185*** (0.647)	-0.419*** (0.072)
Linear polynomial in latitude and longitude	-0.240*** (0.089)	-0.243*** (0.091)	3.420*** (0.594)	-0.371*** (0.065)
Mean dep. var. in interior	4.051 (0.879)	3.314 (0.906)	10.44 (4.792)	1.468 (1.326)
	Auto trade and repair (5)	Communic. (6)	Education (7)	Healthcare (8)
Linear polynomial in distance to border	-0.929*** (0.276)	-0.382*** (0.091)	-0.972*** (0.197)	-1.0609*** (0.255)
Linear polynomial in latitude and longitude	-0.922*** (0.238)	-0.399*** (0.089)	-0.847*** (0.175)	-0.771*** (0.232)
Mean dep. var. in interior	8.174 (3.367)	1.448 (1.470)	4.087 (2.226)	4.644 (2.858)
	Other service (9)	% Primary educ. or less (10)	% Secondary education (11)	% Tertiary education (12)
Linear polynomial in distance to border	-0.268 (0.180)	5.479*** (0.568)	-3.999*** (0.486)	-2.332*** (0.433)
Linear polynomial in latitude and longitude	-0.330** (0.165)	5.312*** (0.507)	-3.831*** (0.449)	-2.446*** (0.433)
Mean dep. var. in interior	4.291 (2.393)	20.77 (5.108)	66.81 (4.962)	8.879 (4.243)
Observations	5577	5577	5577	5577
Clusters	76	76	76	76
Border segments	50	50	50	50
Bandwidth	50 km	50 km	50 km	50 km
Year	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table 5.9: Long-run Differences at MAL, 8 border segments, 2011.

	ln Pop. density (1)	ln Labour force den. (2)	Unempl. rate (3)	Fin. and insurance (4)
Linear polynomial in distance to border	-0.301*** (0.100)	-0.309*** (0.102)	2.524*** (0.647)	-0.349*** (0.082)
Linear polynomial in latitude and longitude	-0.232*** (0.083)	-0.236*** (0.085)	3.425*** (0.626)	-0.359*** (0.070)
Mean dep. var. in interior	4.042 (0.873)	3.302 (0.898)	10.50 (4.815)	1.412 (1.277)
	Auto trade and repair (5)	Communic. (6)	Education (7)	Healthcare (8)
Linear polynomial in distance to border	-0.895*** (0.294)	-0.330*** (0.090)	-0.947*** (0.191)	-0.956*** (0.258)
Linear polynomial in latitude and longitude	-0.884*** (0.235)	-0.403*** (0.074)	-0.827*** (0.152)	-0.663*** (0.242)
Mean dep. var. in interior	7.971 (3.298)	1.299 (1.275)	4.219 (2.297)	4.686 (2.801)
	Other service (9)	% Primary educ. or less (10)	% Secondary education (11)	% Tertiary education (12)
Linear polynomial in distance to border	-0.198 (0.199)	5.234*** (0.596)	-4.062*** (0.535)	-2.120*** (0.392)
Linear polynomial in latitude and longitude	-0.370** (0.140)	5.205*** (0.505)	-3.836*** (0.453)	-2.366*** (0.373)
Mean dep. var. in interior	4.530 (2.653)	20.72 (4.935)	66.96 (4.795)	8.741 (3.907)
Observations	3845	3845	3845	3845
Clusters	70	70	70	70
Border segments	8	8	8	8
Bandwidth	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

then presents the results of such regressions. Also in this case, the estimated coefficients do not differ significantly from the main estimates in Table 5.1, Table 5.2, and Table 5.3, most of them are again only subtly lower. Thus, these results do not reveal any disturbing effects of these North-West-Bohemian areas that were formerly highly polluted on the results of the analysis.

Table 5.10: Long-run Differences at MAL, No Pollution, 2011.

	ln Pop. density (1)	ln Labour force den. (2)	Unempl. rate (3)	Fin. and insurance (4)
Linear polynomial in distance to border	-0.325*** (0.092)	-0.332*** (0.096)	2.737*** (0.598)	-0.382*** (0.081)
Linear polynomial in latitude and longitude	-0.275*** (0.092)	-0.278** (0.096)	3.440*** (0.556)	-0.373*** (0.077)
Mean dep. var. in interior	4.039 (0.878)	3.299 (0.903)	10.45 (4.828)	1.408 (1.276)
	Auto trade and repair (5)	Communic. (6)	Education (7)	Healthcare (8)
Linear polynomial in distance to border	-0.762** (0.345)	-0.296*** (0.099)	-1.088*** (0.164)	-0.898*** (0.278)
Linear polynomial in latitude and longitude	-0.761** (0.304)	-0.372*** (0.092)	-0.981*** (0.151)	-0.715*** (0.269)
Mean dep. var. in interior	7.979 (3.311)	1.293 (1.280)	4.238 (2.308)	4.677 (2.773)
	Other service (9)	% Primary educ. or less (10)	% Secondary education (11)	% Tertiary education (12)
Linear polynomial in distance to border	-0.218 (0.218)	5.394*** (0.641)	-4.016*** (0.597)	-2.392*** (0.373)
Linear polynomial in latitude and longitude	-0.351** (0.165)	5.351*** (0.527)	-3.700*** (0.487)	-2.663*** (0.406)
Mean dep. var. in interior	4.513 (2.660)	20.70 (4.952)	66.95 (4.801)	8.808 (3.920)
Observations	3572	3572	3572	3572
Clusters	64	64	64	64
Border segments	50	50	50	50
Bandwidth	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Chapter 6

Discussion

In this chapter, the results that are shown in Chapter 5 are being discussed more broadly. The aim is to clarify the economic meaning of the main results of long-run effects of the expulsion of Germans, and to compare how they differ from the results of the previous literature, specifically of Testa (2020). Additional tables are provided here so that results of this thesis and results of Testa (2020), who uses the baseline data in the analysis (i.e. does not account for the potentially confounding factors that this thesis explores) can easily be compared. Also, the meaning of the findings from the sensitivity and robustness checks is discussed here.

The existing literature finds statistically significant negative economic effects of the expulsion of Germans from the Czechoslovak borderlands, but these results can be exaggerated because of some factors that have not been considered by it. Thus, this thesis uses similar methods as the prior literature to identify the long-run economic effects of the Germans' expulsion while also examining the effects of these previously not considered factors – mainly of the military areas, the border zone, and then also the of the 45DL, and pollution in North-West Bohemia.

First, according to Table 6.1, the estimated long-run differences in the economic activity show an adverse effect of the Germans' expulsion on population density, labour force density as well as on the unemployment rate in the borderlands. When the results of this thesis are compared with the results of Testa (2020), we can see that the statistical significance does not change – all three coefficients are statistically significant at 1% level. Nevertheless, the estimated negative effect on the population density is slightly stronger in this thesis, where the military areas and the 10 km wide border zone are excluded

from the regressions. But when only the military areas are excluded, the negative effect becomes weaker compared the one estimated by Testa (2020), which is what we would expect (see Table A.7). Thus, what enlarges the final result is the exclusion of the border zone.

Table 6.1: Long-run Differences in Labour Market at MAL, Comparison, 2011.

	This thesis			Using sample of Testa (2020)		
	In Population density (1)	In Labour force density (2)	Unemployment rate (3)	In Population density (4)	In Labour force density (5)	Unemployment rate (6)
<i>Polynomials in distance to border</i>						
Linear polynomial	-0.335*** (0.105)	-0.343*** (0.108)	2.595*** (0.616)	-0.312*** (0.095)	-0.317*** (0.097)	2.729*** (0.546)
Quadratic polynomial	-0.270** (0.118)	-0.277** (0.119)	1.498** (0.568)	-0.215** (0.104)	-0.219** (0.104)	1.702*** (0.536)
Cubic polynomial	-0.132 (0.124)	-0.148 (0.126)	2.277*** (0.534)	-0.118 (0.116)	-0.131 (0.116)	2.545*** (0.525)
<i>Polynomials in latitude and longitude</i>						
Linear polynomial	-0.248*** (0.093)	-0.251** (0.096)	3.343*** (0.587)	-0.251*** (0.084)	-0.253*** (0.086)	3.623*** (0.520)
Quadratic polynomial	-0.329*** (0.091)	-0.331*** (0.094)	3.303*** (0.629)	-0.332*** (0.080)	-0.330*** (0.083)	3.435*** (0.542)
Cubic polynomial	-0.328*** (0.094)	-0.335*** (0.097)	3.223*** (0.650)	-0.314*** (0.083)	-0.317*** (0.086)	3.294*** (0.553)
Mean dep. var. in interior	4.042 (0.873)	3.302 (0.898)	10.50 (4.815)	4.034 (0.885)	3.294 (0.911)	10.492 (4.809)
Observations	3845	3845	3845	4049	4049	4049
Clusters	70	70	70	71	71	71
Border segments	50	50	50	50	50	50
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011
Exclusion of MA and BZ	Yes	Yes	Yes	No	No	No

Notes: MA denotes military areas and BZ stands for the border zone. All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and polynomials in distance to the MAL interacted with the treatment or polynomials in longitude and latitude. In all regressions Prague and Polish Zaolzie are excluded. In the estimates from this thesis also military areas and border zone are excluded. Observed units are municipalities.

An inspection of the data reveals that the 10 km wide border zone includes some municipalities in the district of Břeclav, Domažlice and Znojmo with very large population densities compared to the rest of the municipalities. So, when these municipalities are then excluded from the analysis it is understandable that the resulting negative effect on the population density is stronger than the one from Testa (2020). This is then also reflected in the estimated effect on the labour force density – the coefficient has a value of -0.343, while in Testa (2020) it is -0.317. Yet the differences are very small as, in this thesis, the estimated long-run effect of the Germans' expulsion on the population density and on the labour force density in the borderlands is only 3.2 pp and 3.6 pp lower, respectively, when compared to Testa (2020).

On the contrary, in this thesis, when the military areas and the border zone are excluded from the analysis, the estimated effect of the Germans' expulsion on the unemployment rate in the borderlands is still positive but slightly weaker compared to the result of Testa (2020), who does leave these areas in the sample. Similar to the case of the two economic factors described in the previous paragraph, when only the military areas are excluded from the analysis, the result is very similar to the one of Testa (2020) (see Table A.7). This suggests that the exclusion of the border zone is again what decreases the magnitude of the effect. Specifically, this is a merit of some municipalities in the Znojmo district (Podhradí nad Dyjí, Zálesí, Lančov, Šafov and others), in which there is a very high unemployment rate. Thus, excluding them from the analysis then pushes the resulting estimate down. But, as in the paragraph above, the final difference between the estimate in this thesis and the one from Testa (2020) is minor – when military areas and the border zone are excluded from the regression the estimated effect of the Germans' expulsion is an increase of 2.6 pp in the unemployment rate when one crosses the MAL, while in Testa (2020) the result is a 2.7 pp increase. Thus, there is only a 0.1 pp difference.

Then, Table 6.2 reveals that in the borderlands there is much lower employment in skill-intensive sectors compared to the interior areas. Specifically, there are drops in the employment in the financial and insurance sector, in communications, healthcare and in education, and also in the auto trade and repair sector. When comparing the results of our analysis with the results of Testa (2020), there are no differences in the statistical significance and the direction as well as the magnitude of the effects is very much the same.

Table 6.2: Long-run Differences in Sectoral Composition at MAL, Comparison, 2011.

	This thesis				Testa (2020)			
	Agricultural sector	Industry	Construction	Transport sector	Agricultural sector	Industry	Construction	Transport sector
	(1)	(2)	(3)	(4)	(13)	(14)	(15)	(16)
Linear polynomial in distance to border	-0.398 (0.457)	-0.453 (0.933)	-0.416 (0.282)	0.065 (0.266)	-0.573 (0.433)	-0.621 (0.860)	-0.447* (0.267)	-0.117 (0.248)
Linear polynomial in latitude and longitude	-0.721 (0.467)	-1.015 (0.800)	-0.043 (0.225)	0.146 (0.280)	-0.752 (0.457)	-1.132 (0.717)	-0.136 (0.223)	-0.034 (0.244)
Mean dep. var. in interior	7.582 (6.339)	26.66 (7.476)	7.218 (3.183)	5.304 (2.660)	7.653 (6.576)	26.672 (7.489)	7.212 (3.197)	5.309 (2.664)
	Finance and insurance	Hospitality	Auto trade and repair	Public	Finance and insurance	Hospitality	Auto trade and repair	Public
	(5)	(6)	(7)	(8)	(17)	(18)	(19)	(20)
Linear polynomial in distance to border	-0.399*** (0.078)	0.159 (0.259)	-0.866*** (0.304)	-0.056 (0.294)	-0.369*** (0.080)	0.365 (0.280)	-0.864*** (0.282)	0.001 (0.274)
Linear polynomial in latitude and longitude	-0.386*** (0.070)	0.255 (0.225)	-0.858*** (0.267)	-0.320 (0.249)	-0.386*** (0.073)	0.440* (0.234)	-0.933*** (0.252)	-0.310 (0.232)
Mean dep. var. in interior	1.412 (1.277)	2.113 (1.580)	7.971 (3.298)	4.464 (2.367)	1.408 (1.273)	2.119 (1.587)	7.959 (3.311)	4.476 (2.371)
	Communi-cations	Education	Healthcare	Other service	Communi-cations	Education	Healthcare	Other service
	(9)	(10)	(11)	(12)	(21)	(22)	(23)	(24)
Linear polynomial in distance to border	-0.322*** (0.089)	-0.972*** (0.207)	-1.002*** (0.258)	-0.208 (0.195)	-0.297*** (0.089)	-0.864*** (0.176)	-0.993*** (0.230)	-0.219 (0.166)
Linear polynomial in latitude and longitude	-0.3949*** (0.087)	-0.862*** (0.173)	-0.718*** (0.242)	-0.345** (0.158)	-0.381*** (0.082)	-0.791*** (0.145)	-0.780*** (0.233)	-0.411*** (0.134)
Mean dep. var. in interior	1.299 (1.275)	4.219 (2.297)	4.686 (2.801)	4.530 (2.653)	1.294 (1.272)	4.203 (2.297)	4.676 (2.797)	4.515 (2.650)
Observations	3845	3845	3845	3845	4049	4049	4049	4049
Clusters	70	70	70	70	71	71	71	71
Border segments	50	50	50	50	50	50	50	50
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011	2011	2011
Exclusion of MA and BZ	Yes	Yes	Yes	Yes	No	No	No	No

Notes: MA denotes military areas and BZ stands for the border zone. All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions Prague and Polish Zaolzie are excluded. In the estimates from this thesis also military areas and border zone are excluded. Observed units are municipalities.

Some differences from the prior literature occur in the analysis of educational attainment, whose results are shown in Table 6.3. The significance levels remain the same as in the prior literature. Nevertheless, in this thesis, the percentage of people with primary education is estimated to increase by 5.4 pp on average when crossing the MAL, while Testa (2020) arrived at the result of 4.9 pp. Naturally, this difference is then connected to differences in the percentage of people with secondary and/or tertiary education. This is because when there are more people with primary education, there must be less people with secondary or tertiary education (or even both). In Table 6.3, we can see that when using the dataset of this thesis the estimates of the percentage of people with secondary and tertiary education are lower by 0.2 and 0.3 pp, respectively, than the estimates of Testa (2020). Though these differences are very small, and they are again caused only by the exclusion of the 10 km border zone (see Table A.8) and not by the exclusion of the military areas. Specifically, exploration of the data revealed that it is caused by some border zone municipalities in the Jihočeský and Plzeňský region, which have very low percentage of people with primary education and a high percentage of people with the higher education (most of them are in the Prachatice and Domažlice districts).

Table 6.3: Long-run Differences in Educational Attainment at MAL, Comparison, 2011.

	This thesis		
	% Primary education or less	% Secondary education	% Tertiary education
	(1)	(2)	(3)
Linear polynomial in distance to border	5.384*** (0.612)	-4.145*** (0.537)	-2.187*** (0.404)
Linear polynomial in latitude and longitude	5.269*** (0.508)	-3.822** (0.456)	-2.410*** (0.420)
Mean dep. var. in interior	20.72 (4.935)	66.96 (4.795)	8.741 (3.907)
Observations	3845	3845	3845
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Exclusion of MA and BZ	Yes	Yes	Yes
	Testa (2020)		
	% Primary education or less	% Secondary education	% Tertiary education
	(4)	(5)	(6)
Linear polynomial in distance to border	4.883*** (0.634)	-3.923*** (0.516)	-1.936*** (0.391)
Linear polynomial in latitude and longitude	4.965*** (0.543)	-3.759*** (0.453)	-2.270*** (0.375)
Mean dep. var. in interior	20.767 (4.980)	66.939 (4.827)	8.716 (3.926)
Observations	4049	4049	4049
Clusters	71	71	71
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Exclusion of MA and BZ	No	No	No

Notes: MA denotes military areas and BZ stands for the border zone. All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions Prague and Polish Zaolzie are excluded. In the estimates from this thesis also military areas and border zone are excluded. Observed units are municipalities.

So far, the results of this thesis seem to be in line with the results of Testa (2020), who studied the same economic factors. The very small differences in the magnitude of some estimates are caused by some municipalities in the 10 km wide border zone, which is being excluded in the analysis in this thesis. On the contrary, when only the military areas are excluded from the analysis, the results are basically identical to those of the prior literature. Nevertheless, also pollution was suspected of having some effect on the long-run development in the borderlands. In Table 6.4, which summarizes corresponding results from Chapter 5, it can be seen that when areas in the North-West Bohemia, which experienced high levels of air pollution during the communist regime, are excluded from the regressions, then the results basically remain consistent with the original results (when only the military areas and border zone are excluded). There are only very small differences in the magnitudes of the estimates, though after excluding the formerly polluted North-West Bohemian municipalities the estimates move in a direction which takes them closer to the results of Testa (2020).

Table 6.4: Long-run Differences at MAL, No Pollution, Comparison, 2011.

	Without polluted areas				With polluted areas			
	In Pop. density (1)	In Labour force den. (2)	Unempl. rate (3)	Fin. and insurance (4)	In Pop. density (13)	In Labour force den. (14)	Unempl. rate (15)	Fin. and insurance (16)
Linear polynomial in distance to border	-0.325*** (0.092)	-0.332*** (0.096)	2.737*** (0.598)	-0.382*** (0.081)	-0.335*** (0.105)	-0.343*** (0.108)	2.595*** (0.616)	-0.399*** (0.078)
Linear polynomial in latitude and longitude	-0.275*** (0.092)	-0.278** (0.096)	3.440*** (0.556)	-0.373*** (0.077)	-0.248*** (0.093)	-0.251** (0.096)	3.343*** (0.587)	-0.386*** (0.070)
Mean dep. var. in interior	4.039 (0.878)	3.299 (0.903)	10.45 (4.828)	1.408 (1.276)	4.042 (0.873)	3.302 (0.898)	10.50 (4.815)	1.412 (1.277)
	Auto trade and repair	Communic.	Education	Healthcare	Auto trade and repair	Communic.	Education	Healthcare
	(5)	(6)	(7)	(8)	(17)	(18)	(19)	(20)
Linear polynomial in distance to border	-0.762** (0.345)	-0.296*** (0.099)	-1.088*** (0.164)	-0.898*** (0.278)	-0.866*** (0.304)	-0.322*** (0.089)	-0.972*** (0.207)	-1.002*** (0.258)
Linear polynomial in latitude and longitude	-0.761** (0.304)	-0.372*** (0.092)	-0.981*** (0.151)	-0.715*** (0.269)	-0.858*** (0.267)	-0.3949*** (0.087)	-0.862*** (0.173)	-0.718*** (0.242)
Mean dep. var. in interior	7.979 (3.311)	1.293 (1.280)	4.238 (2.308)	4.677 (2.773)	7.971 (3.298)	1.299 (1.275)	4.219 (2.297)	4.686 (2.801)
	Other service	% Primary educ. or less	% Secondary education	% Tertiary education	Other service	% Primary educ. or less	% Secondary education	% Tertiary education
	(9)	(10)	(11)	(12)	(21)	(22)	(23)	(24)
Linear polynomial in distance to border	-0.218 (0.218)	5.394*** (0.641)	-4.016*** (0.597)	-2.392*** (0.373)	-0.208 (0.195)	5.384*** (0.612)	-4.145*** (0.537)	-2.187*** (0.404)
Linear polynomial in latitude and longitude	-0.351** (0.165)	5.351*** (0.527)	-3.700*** (0.487)	-2.663*** (0.406)	-0.345** (0.158)	5.269*** (0.508)	-3.822** (0.456)	-2.410*** (0.420)
Mean dep. var. in interior	4.513 (2.660)	20.70 (4.952)	66.95 (4.801)	8.808 (3.920)	4.530 (2.653)	20.72 (4.935)	66.96 (4.795)	8.741 (3.907)
Observations	3572	3572	3572	3572	3845	3845	3845	3845
Clusters	64	64	64	64	70	70	70	70
Border segments	50	50	50	50	50	50	50	50
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011	2011	2011

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, * and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, border zone, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Then, during the first few months after the WWII there was the period of Wild expulsion mainly in the area that was liberated by the Soviet army. This event could have also had some influence on the long-run development of the borderland municipalities in the Soviet-liberation zone. So, what Table 5.4¹ tells us is that the situation on the labour market in the borderlands does not seem to be influenced by the period of Wild expulsion. That is because the estimates are practically the same when one crosses from the Soviet zone to the US zone in the borderlands as well as in the interior. An impact of the Wild expulsion could be confirmed only if there was a significant effect of crossing the 45DL in the borderlands but no effect in the interior since the expulsion was happening in the borderlands, where majority of the German population lived. Then, Table 5.5 also does not provide any evidence of the Wild expulsion having an effect on the long-run development of the sectoral composition in the borderlands. And finally, Table 5.6 shows estimated differences in the educational attainment as one crosses the 45DL in the borderlands and in the interior. This table revealed some interesting information about the percentage of people with secondary education – in the borderlands the percentage increases by 3.3 pp when one crosses the 45DL from the Soviet zone to the US zone. But the same difference is also estimated in the interior municipalities, thus this difference cannot be attributed to the Wild expulsion. It would be the case only if there was a statistically significant difference in the borderlands and no difference in the interior.

So, even the analysis of differences between municipalities surrounding the 45DL did not provide any evidence of the Wild expulsion of Germans in the Soviet zone having any impact on the long-term development of the economy in the borderlands. Even though no effect of the Wild expulsion was identified, there seems to be some variation in the educational attainment around the 45DL. This may be worth exploring in some further research that would focus on the 45DL. Nevertheless, coming back to our topic, this thesis shows that none of the three factors (i.e. military areas, pollution, 45DL) seems to have a confounding effect on the long-run development of the borderlands, and thus they should not be biasing the analysis of the effects of the Germans' expulsion.

¹To easily return to this part of the discussion you can use links in the notes under Table 5.4, Table 5.5, and Table 5.6.

Chapter 7

Conclusion

The main purpose of this thesis was to study the long-run economic effects of the expulsion of Germans from Czechoslovakia after the end of the World War II, and specifically to test the sensitivity of the prior literature's results. In this historical event, the citizenship and property rights of almost all Germans living in the Czech lands were taken from them, and they were permanently expelled. Since until the Munich agreement the Czech and German inhabitants lived in the Sudetenland next to each other for several centuries with a lot of interactions between both nationalities (including marriages), it was an ideal event of forced migration for analysing its impacts on the original country.

The thesis, thus, focused on factors that had high potential to be confounding to the prior literature's results, and that were not addressed by the literature before. Namely these factors were the presence of military areas and the border zone in the data, heavy air pollution in the North-West Bohemia during the communist regime, and impact of Wild expulsion in the Soviet liberation zone right after the end of the World War II. The impact of the military areas, the border zone, and the North-West Bohemian areas was examined by studying how important economic indicators change when one crosses the border of the Sudetenland using the same dataset from municipalities along the Munich Agreement Line as in the prior literature but excluding the corresponding municipalities from the analysis. On the contrary, the possible confounding effects of the Wild expulsion were studied on a smaller set of municipalities along the 1945 US-Soviet demarcation line, and then comparing the results of crossing the demarcation line in the borderlands with those from the interior.

Using the discontinuity in the ethnic composition at the border of the Sudetenland (the Czech borderlands) and employing the regression discontinuity

design, the results showed very minor to no changes when the military areas and the border zone were excluded from the dataset (compared to the results of the prior literature). The same was true even when the North-West Bohemian municipalities that had high levels of air pollution in the past were dropped from the regressions. Other robustness checks also did not reveal anything that would indicate a bias in the results. Then, the analysis of areas around the 1945 demarcation line did not provide evidence for the effect of the Wild expulsion on the economic development of the borderlands. Thus, even though the results of the prior literature were challenged by the described factors, the significant adverse effects of the Germans' expulsion on the borderland economy still persisted.

Therefore, the findings of this thesis are in line with the findings of the existing literature and provide information about the consequences that forced migration can have on the expellees' origin economy. In our case of the Czech borderlands, there is a lasting impact on the development of the borderland economy, which also results in regional inequality within the Czech Republic. Specifically, the identified consequences are higher unemployment, lower population density and employment in high-skilled sectors such as finance or healthcare as well as a lower share of population with secondary and tertiary education in the borderlands, when compared to the Czech interior, that lasts until today. For these reasons, politicians should be very cautious when considering stripping off a citizenship from any minority group, and they should take the described consequences of what a forced expulsion can do to the places that are left behind into account.

Nevertheless, there are some questions that remain to be answered in this topic of the Germans' expulsion from Czechoslovakia. First, the channels of the effects were not explored here. Though the prior literature studied the channels, it may yield different results when the factors addressed in this thesis are controlled for. Second, further research may also explore the role of communism in the development of the borderlands as one of the channels, possibly also through the Council for Mutual Economic Assistance, which was a power tool for the central control of the economy of the socialist states by the Soviet Union. Third, an analysis of separate parts of the borderlands might reveal some interesting information, but the results would be only local with no possibility of generalization. Fourth, this thesis used data on the municipal-level as they were the most granular data available to us. But further research that would collect and use household-level data for the analysis would be able to

study a much narrower area around the Munich Agreement Line. This could reveal more details about the resettlement of the borderlands such as the reasons why people chose to resettle there. Fifth, there remains a concern about comparability of the pre-expulsion and post-expulsion data since in 1930 census people could choose whether they identify as Germans or not, while in 1945 they did not have this choice, which is not controlled for in this thesis. So, this might also be done in some way in some further research.

Finally, it was also addressed by the prior literature that studying a topic of forced migration is often accompanied by concerns about external validity, i.e. about whether the findings can be generalized. That is because most expulsions happen during wars, which always have specific settings. Therefore, further research is needed to explore forced migrations caused by various factors, not just the political ones, and then to compare the results.

References

- Alix-Garcia, J. and D. Saah (2010). “The Effect of Refugee Inflows on Host Communities: Evidence from Tanzania”. In: *The World Bank Economic Review* 24.1, pp. 148–170. URL: <https://www.jstor.org/stable/40647424?seq=1>.
- “Amtliches Deutsches Ortsbuch für das Protektorat Böhmen und Mähren” (1940). In: *Prague: Reichsprotektor in Böhmen und Mähren*.
- Arburg, A. von and T. Staněk (2011). *Vysídlení Němců a proměny českého pohraničí 1945-1951*. Ve Středoklukách: Zdeněk Susa. ISBN: 978-80-86057-67-5.
- Bauer, T. K., S. Braun, and M. Kvasnicka (2013). “The economic integration of forced migrants: Evidence for postwar Germany”. In: *The Economic Journal* 123.571, pp. 998–1024. URL: <https://www.jstor.org/stable/42919252?seq=1>.
- Becker, S. O., K. Boeckh, et al. (2014). “The Empire Is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy”. In: *The Economic Journal* 126.590, pp. 40–74. URL: <https://onlinelibrary.wiley.com/doi/10.1111/eoj.12220>.
- Becker, S. O., I. Grosfeld, et al. (2020). “Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers”. In: *American Economic Review* 110.5, pp. 1430–63. URL: <https://www.aeaweb.org/articles?id=10.1257/aer.20181518>.
- Brandes, D. (2002). *Cesta k vyhnání 1938-1945: plány a rozhodnutí o "transferu" Němců z Československa a z Polska*. Praha: Prostor. ISBN: 80-7260-070-2.
- Braun, S., A. Kramer, et al. (2021). “Local labor markets and the persistence of population shocks: evidence from West Germany, 1939–1970”. In: *Journal of Economic Geography* 21.2, pp. 231–260. URL: <https://academic.oup.com/joeg/article-abstract/21/2/231/5902009?redirectedFrom=fulltext>.

- Braun, S. and T. O. Mahmoud (2011). “The employment effects of immigration: Evidence from the mass arrival of German expellees in post-war Germany”. In: *Kiel Working Paper* 1725. URL: <https://www.econstor.eu/bitstream/10419/49384/1/66629626X.pdf>.
- Calderón, V. and A. M. Ibáñez (2005). “Labor Market Effects of Migration-Related Supply Shocks: Evidence from Internally Displaced Populations in Colombia”. In: *Documentos CEDE, Universidad de los Andes - CEDE* 005851. URL: <https://ideas.repec.org/p/col/000089/005851.html>.
- Daněk, P. (1995). “Towards cultural regionalization of the Czech lands: Sudeten half a century after the transfer”. In: *Scripta Fac. Sci. Nat. Univ. Masaryk. Brun., Geography* 25, pp. 41–60. URL: https://www.sci.muni.cz/geobib/scripta/1995/Scripta_1995_25_Danek.pdf.
- Dejmek, J. et al. (1999). “Fetal Growth and Maternal Exposure to Particulate Matter during Pregnancy”. In: *Environmental Health Perspectives* 107.6, pp. 475–480. URL: https://www.researchgate.net/publication/12960291_Fetal_Growth_and_Maternal_Exposure_to_Part particulate_Matter_during_Pregnancy.
- Dell, M. (2010). “The persistent effects of Peru’s mining mita”. In: *Econometrica* 78.6, pp. 1863–1903. URL: https://scholar.harvard.edu/files/dell/files/ecta8121_0.pdf.
- Eder, CH. and M. Halla (2016). “The Long-lasting Shadow of the Allied Occupation of Austria on its Spatial Equilibrium”. In: *IZA Discussion Paper* 10095. URL: <https://ftp.iza.org/dp10095.pdf>.
- Etnologický ústav AV, Akademie věd ČR (n.d.). *Vojenské prostory a hraniční pásma*. URL: <http://prerusenapamet.eu.cas.cz/vojna.php>. (accessed: 10.3.2022).
- Gelman, A. and G. Imbens (2018). “Why high-order polynomials should not be used in regression discontinuity designs”. In: *Journal of Business and Economic Statistics* 37, pp. 447–56. URL: http://www.stat.columbia.edu/~gelman/research/published/2018_gelman_jbes.pdf.
- Gerlach, D. W. (2017). *The Economy of Ethnic Cleansing: The Transformation of the German-Czech Borderlands after World War II*. Cambridge: Cambridge University Press, 2017. ISBN: 978-1-107-19619-3.
- Glassheim, E. (2005). “Etnické čistky, komunismus a devastace životního prostředí: Vytváření nové identity severočeského pohraničí (1945–1989)”. In: *Soudobé dějiny, Komplex odsunu: vysídlení Němců a české pohraničí po roce 1945*

- 7.3-4, pp. 432–464. URL: <http://www.usd.cas.cz/casopis/soudobe-dejiny-3-4-2005/>.
- Grossmann, J., Š. Jurajda, and F. Roesel (2021). “Forced Migration, Staying Minorities, and New Societies: Evidence from Post-War Czechoslovakia”. In: *CESifo Working Papers* 8950. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3784126.
- Guzi, M., P. Huber, and M. Mikula (2019). “Old sins cast long shadows: The Long-term impact of the resettlement of the Sudetenland on residential migration”. In: *MUNI ECON Working Paper* 2019-07. URL: http://repec.econ.muni.cz/mub/wpaper/wp/econ/WP_MUNI_ECON_2019-07.pdf.
- Hornung, E. (2014). “Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia”. In: *American Economic Review* 104.1, pp. 84–122. URL: <https://www.jstor.org/stable/42920689?seq=1>.
- Jelínek, J. (2011). “Analýza topografických map 1:50000 z období Protektorátu Čechy a Morava”. In:
- King, J. (2005). *Budweisers into Czechs and Germans : a local history of Bohemian politics 1848-1948*. Princeton: Princeton University Press. ISBN: 0-691-12234-2.
- Kotěšovec, F. et al. (2000). “Daily Mortality and Air Pollution in Northern Bohemia: Different Effects for Men and Women”. In: *Central European Journal of Public Health* 8.2, pp. 120–127. URL: https://www.researchgate.net/publication/12461714_Daily_mortality_and_air_pollution_in_Northern_Bohemia_Different_effects_for_men_and_women.
- Kubisa, V. (2010). *Neznámý svět vojenských újezdů*. URL: <https://www.mvcr.cz/clanek/neznamy-svet-vojenskych-ujezdu-974223.aspx>. (accessed: 10.3.2022).
- Kučera, J. (1992). *Odsun nebo vyhnání?: sudetští Němci v Československu v letech 1945-1946*. Praha: H & H. ISBN: 80-85467-32-1.
- Kural, V. and Z. Radvanovský (2002). *"Sudety" pod hákovým křížem*. Ústí nad Labem: Albis international. ISBN: 80-86067-66-1.
- Lee, D. S. and T. Lemieux (2010). “Regression discontinuity designs in economics”. In: *Journal of Economic Literature* 48, pp. 281–355. URL: <https://www.princeton.edu/~davidlee/wp/RDDEconomics.pdf>.
- Mikula, Š. and M. Pytliková (2021). “Air Pollution and Migration: Exploiting a Natural Experiment from the Czech Republic”. In: *IZA Discussion Paper* 14863. URL: <https://docs.iza.org/dp14863.pdf>.

- Moulton, B. R. (1990). “An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units”. In: *The Review of Economics and Statistics* 72.2, pp. 334–338. URL: <https://www.jstor.org/stable/2109724?seq=1>.
- NUV, Národní ústav pro vzdělávání (2010). *Územně správní členění České Republiky*. URL: <http://archiv-nuv.npi.cz/t/informace-o-uzemne-spravnim-cleneni-ceske-republiky.html>. (accessed: 2.4.2022).
- Ochsner, Ch. (2017). “Dismantled once, diverged forever? A quasi-natural experiment of Red Army misdeeds in post-WWII Europe”. In: *ifo Working Paper* 240. URL: https://ideas.repec.org/p/ces/ifowps/_240.html.
- Olivová, V. (2008). *Dějiny nové doby: 1848-2008*. Velké Bílovice: TeMi CZ. ISBN: 978-80-87156-16-2.
- Pecka, J. (1995). *Na demarkační čáře: americká armáda v Čechách v roce 1945*. Praha: Ústav pro soudobé dějiny Akademie věd České republiky. ISBN: 80-85270-47-1.
- Schumann, A. (2014). “Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II”. In: *American Economic Journal: Applied Economics* 6.3, pp. 189–205. URL: <https://www.aeaweb.org/articles?id=10.1257/app.6.3.189>.
- Slapnicka, H. (2000). “Oberösterreichische Demarkationslinien und die Fortsetzung nach Norden”. In: *Oberösterreichische Heimatblätter* 54, pp. 73–76. URL: <https://www.oogeschichte.at/datenbanken/literaturdatenbank-historische-bibliografie>.
- Spurný, M. (2013). “Reliability and the Border. Czech Borderlands Discourse, 1945–49”. In: *Österreichische Zeitschrift für Politikwissenschaft* 42.1, pp. 83–94. URL: <https://webapp.uibk.ac.at/ojs/index.php/OEZP/article/view/1446>.
- Staněk, T. (1991). *Odsun Němců z Československa 1945-1947*. Praha: Academia. ISBN: 80-200-0328-2.
- (1992). *Předpoklady, průběh a důsledky vysídlení Němců z Československa (1918-1948)*. Ostrava: Amosium servis. ISBN: 80-85498-04-9.
- “Sudetendeutsches Ortsnamenverzeichnis” (1987). In: *Sudetendeutsches Ortsnamenverzeichnis: Amtliches Gemeinde- und Ortsnamenverzeichnis der nach dem Münchener Abkommen vom 29. 9. 1938 (Grenzfestlegung vom 20. 11. 1938) zum Deutschen Reich gekommenen sudetendeutschen Gebiete*. Munich: Sudetendeutsches Archiv.

- Suppan, A. (2006). “Austrians, Czechs, and Sudeten Germans as a Community of Conflict in the Twentieth Century”. In: *Center for Austrian Studies - Working Paper* 06.1. URL: <https://conservancy.umn.edu/bitstream/handle/11299/90505/wp061.pdf?sequence=1&isAllowed=y>.
- Tampke, J. (2003). *Czech-German Relations and the Politics of Central Europe: from Bohemia to the EU*. New York: Palgrave Macmillan. ISBN: 0-333-73449-1.
- Testa, P. A. (2020). “The Economic Legacy of Expulsion: Lessons from Postwar Czechoslovakia”. In: URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3251541.
- (2021). *Research: Publications*. URL: <https://www.patrickatesta.com/research>. (accessed: 28.10.2021).
- Tobin, J. (1974). “Notes on the Economic Theory of Expulsion and Expropriation”. In: *Journal of Development Economics* 1, pp. 7–18. URL: <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/520>.
- Topinka, J. (2005). “Zapomenutý kraj: České pohraničí 1948–1960 a takzvaná akce dosídlení”. In: *Soudobé dějiny, Komplex odsunu: vysídlení Němců a české pohraničí po roce 1945* 7.3-4, pp. 534–585. URL: <http://www.usd.cas.cz/casopis/soudobe-dejiny-3-4-2005/>.
- Wyrwich, M. (2020). “Migration restrictions and long-term regional development: evidence from large-scale expulsions of Germans after World War II”. In: *Journal of Economic Geography* 20, pp. 481–507. URL: <https://ideas.repec.org/p/jrp/jrpwpr/2018-002.html>.
- Zahra, T. (2008). *Kidnapped Souls: National Indifference and the Battle for Children in the Bohemian Lands, 1900-1948*. Ithaca: Cornell University Press. ISBN: 978-0-8014-4628-3.
- Zayas, A. M. de (1994). *A Terrible Revenge: The Ethnic Cleansing of the East European Germans, 1944-50*. New York: Saint Martin’s Press. ISBN: 0-312-12159-8.

Appendix A

Appendix

Table A.1: Summary Statistics, 2011 (full sample).

	Observations	Mean	Std. Dev.	Min.	Max.
ln Population density	6206	3.987	0.946	-3.965	7.894
ln Labour force density	6206	3.251	0.967	-5.064	7.214
Unemployment	6206	11.511	5.686	0	52.222
% Agricultural sector	6206	7.111	6.195	0	100
% Industry	6206	25.769	8.119	0	58.974
% Construction	6206	7.135	3.148	0	30.769
% Transport sector	6206	5.301	2.75	0	27.273
% Finance and insurance	6206	1.381	1.262	0	10
% Hospitality	6206	2.435	2.341	0	41.213
% Auto trade and repair	6206	7.861	3.322	0	29.167
% Public	6206	4.54	2.664	0	64.029
% Communications	6206	1.33	1.399	0	14.085
% Education	6206	4.012	2.199	0	22.222
% Healthcare	6206	4.525	2.864	0	54.412
% Other service	6206	4.599	2.806	0	33.333
% Primary education or less	6206	21.797	5.698	0	68.908
% Secondary education	6206	65.919	5.317	25.21	86.111
% Tertiary education	6206	8.364	4.214	0	33.741
Elevation (m)	6206	410.505	144.345	121.833	1144.601
Ruggedness (°)	6206	6.422	3.001	1.053	20.725
Precipitation (mm)	6206	53.047	6.98	40.494	100.068
Temperature (°C)	6206	7.581	0.82	3.262	9.534
Rivers/km ² (km)	6206	1.183	0.52	0	5.1

Notes: In all values Prague and Polish Zaolzie are omitted because they are excluded in all regression of this thesis. Observed units are municipalities.

Table A.2: Summary Statistics, 1930 (full sample).

	Observations	Mean	Std. Dev.	Min.	Max.
% German	325	35.666	41.132	0.024	98.845
Literacy rate	325	98.516	0.709	94.94	99.65
ln Population density	325	4.724	0.623	3.336	9.002
ln Labour force density	325	3.951	0.698	2.413	8.476
Labour force participation	325	46.468	5.318	33.339	61.908
Unemployment	146	13.125	10.186	1.417	58.796
Income per capita (100 Kčs)	145	9.469	4.105	3.669	30.775
% Taxpayers	147	5.698	1.786	2.33	11.65
Convictions per capita	318	7.274	2.057	2.484	16.318
% Roma	325	0.002	0.012	0	0.129
% Jewish	325	0.145	0.316	0	2.825
Major roads/km ² (km)	378	0.216	0.069	0	0.501
Railways/km ² (km)	378	0.094	0.064	0	0.319
% Agricultural sector	325	28.077	13.991	0.769	60.529
% Secondary sector	325	40.602	13.672	16.967	76.309
% Industry	325	33.557	14.183	10.753	72.145
% Mining and other extraction	325	3.59	5.199	0.26	36.218
% Metallurgy and metalwork	325	4.411	3.499	1.429	24.611
% Machinery and auto	325	2.283	2.21	0.311	16.322
% Glasswork	325	1.133	3.687	0	33.643
% Textiles	325	7.13	10.822	0.029	54.741
% Other industry	325	15.01	6.149	6.809	62.857
% Construction	325	7.045	2.364	2.806	17.536
% Transport sector	325	3.473	2.012	1.131	13.615
% Finance and insurance	325	0.401	0.279	0	3.084
% Trade	325	5.597	2.301	2.512	19.469
% Business sector	325	5.999	2.497	2.592	20.841
% Other service	325	6.694	3.779	3.071	29.368

Notes: In all values Prague and Polish Zaolzie are omitted because they are excluded in all regression of this thesis. Observed units are judicial districts, only for income per capita, % taxpayers, and unemployment it is political districts.

Description of the 1945 US-Soviet demarcation line in the Czech lands from Pecka (1995); translation to English by Grossmann, Jurajda, and Roesel (2021):

"The demarcation line was created in May 1945 (see map on the page 60) and it was approximately crossing along the railroad Honí Dvořiště, Velešín, České Budějovice; it overlapped with the main road between Kosov and Kamenný Újezd and headed West towards Vltava valey, Kremž, Brloh and Nová Ves. Further, it followed the road to Netolice, Vitějovice, Strunkovice nad Blatnicí, Bavorov, Vodňany, and Radčice. Passing the quota 466 directed to Chvaletice, Křtětice, Božovice, Ražice, Heřmaň, around Putim, on the left flank of Otava around Písek to Oldřichov. Chlaponice, Mladotice, Nová hospoda and then along the road Písek-Plzeň to Sedlec, Blatná, Lnáře, Kasejovice, towards Životice, Nepomuk, Spálené Poříčí, Nezvěstice, Štáhlavy, and Nord-West via villages Raková, Rokycany, Borek, Svojkovice, Volduchy, Březina, Bezděkov, Stupno, Všenice, Strápole, Kříše. Then turned around Plzeň to Chrást, Třemošná, Horní Bříza, Kaznějov, Nečtiny following the road to Karlovy Vary through villages Třebouň, Toužim, Útvina, Krásné Údolí, East of the city of Teplá along the railroad Bečov nad Teplou-Krásný Jez, following the ridges of Slavkovský forest to Jalový Dvůr near Loket, Vřesová, Jindřichovice, Kraslice, Stříbrná, Bublava and through German teritory to Plavno-Saská Kamenice up to Labe."

Table A.3: Long-run Differences in Labour Market at MAL, 12 km Border Zone, 2011.

	ln Population density	ln Labour force density	Unemployment rate
	(1)	(2)	(3)
Linear polynomial in distance to border	-0.322*** (0.106)	-0.331*** (0.108)	2.587*** (0.615)
Linear polynomial in latitude and longitude	-0.233** (.093)	-0.237** (.095)	3.346*** (.597)
Mean dep. var. in interior	4.042 (0.874)	3.303 (0.899)	10.49 (4.809)
Observations	3806	3806	3806
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Military areas and border zone	Excluded	Excluded	Excluded

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table A.4: Long-run Differences in Labour Market at MAL, 15 km Border Zone, 2011.

	ln Population density	ln Labour force density	Unemployment rate
	(1)	(2)	(3)
Linear polynomial in distance to border	-0.328*** (0.111)	-0.341*** (0.113)	2.675*** (0.652)
Linear polynomial in latitude and longitude	-0.236** (.096)	-0.244** (.098)	3.435*** (.631)
Mean dep. var. in interior	4.057 (0.875)	3.317 (0.899)	10.42 (4.763)
Observations	3688	3688	3688
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Military areas and border zone	Excluded	Excluded	Excluded

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table A.5: Comparison of non-military and military areas, 1930.

	Non-military areas		Military areas		Border zone	
	Borderlands (1)	Interior (2)	Borderlands (3)	Interior (4)	Borderlands (5)	Interior (6)
% German	78.694 (21.172)	1.584 (3.819)	95.753 (3.193)	1.901 (2.848)	90.933 (7.610)	-
Literacy rate	98.862 (0.585)	98.370 (0.684)	98.769 (1.229)	98.640 (0.244)	97.724 (1.341)	-
ln Population density	4.873 (0.650)	4.776 (0.750)	4.113 (0.505)	4.557 (0.313)	4.065 (0.353)	-
Labour force particip.	49.175 (5.310)	45.457 (4.687)	44.370 (4.825)	44.166 (3.471)	41.677 (2.610)	-
Unemployment rate	21.560 (9.705)	9.961 (10.491)	18.878 (9.184)	7.828 (5.635)	13.655 (7.453)	-
Income per capita	11.473 (4.398)	9.368 (4.880)	8.707 (2.812)	9.818 (3.069)	9.139 (3.688)	-
% Taxpayers	6.471 (1.783)	5.588 (2.009)	5.611 (1.596)	6.291 (1.786)	5.578 (1.714)	-
Convicts per capita	7.863 (2.075)	7.262 (2.098)	8.280 (2.777)	7.448 (1.336)	7.639 (1.360)	-
% Roma	0.000 (0.001)	0.001 (0.006)	0.013 (0.037)	0.000 (0.000)	0.000 (0.000)	-
% Jewish	0.148 (0.249)	0.190 (0.489)	0.146 (0.222)	0.034 (0.024)	0.103 (0.134)	-
km Roads per km ²	0.220 (0.062)	0.230 (0.062)	0.198 (0.065)	0.214 (0.067)	0.178 (0.033)	-
km Railways per km ²	0.112 (0.065)	0.099 (0.057)	0.069 (0.058)	0.073 (0.048)	0.061 (0.045)	-
Max. no. of observations	58	85	8	8	12	0
Bandwidth	25-50 km	25-50 km	25-50 km	25-50 km	25-50 km	25-50 km
Year	1930	1930	1930	1930	1930	1930

Notes: Standard deviations are in parenthesis. In all calculations Prague and Polish Zaolzie are excluded. Observed units are districts.

Table A.6: Comparison of non-military and military areas, 2011.

	Non-military areas		Military areas		Border zone	
	Borderlands (1)	Interior (2)	Borderlands (3)	Interior (4)	Borderlands (5)	Interior (6)
In Population density	3.914 (1.021)	4.042 (0.874)	1.604 (1.237)	3.457 (2.682)	3.243 (1.059)	3.600 (0.789)
In Labour force density	3.179 (1.032)	3.302 (0.899)	0.920 (1.167)	2.687 (2.798)	2.525 (1.064)	2.843 (0.822)
Unemployment rate	14.528 (6.277)	10.500 (4.815)	19.662 (7.056)	6.522 (3.927)	17.091 (7.420)	10.904 (4.166)
Agricultural sector	5.654 (4.894)	7.582 (6.339)	13.870 (6.504)	17.088 (29.784)	10.160 (6.152)	10.373 (6.865)
Industry	25.853 (8.500)	26.657 (7.477)	18.548 (12.483)	28.960 (12.006)	21.011 (9.545)	27.182 (7.121)
Construction	6.772 (3.101)	7.218 (3.183)	3.807 (2.558)	4.793 (2.782)	6.538 (3.450)	7.379 (4.056)
Transport sector	5.070 (2.600)	5.304 (2.660)	3.624 (1.261)	4.519 (2.685)	4.307 (2.412)	5.825 (2.910)
Finance and insurance	1.031 (0.914)	1.412 (1.277)	0.304 (0.367)	0.786 (0.715)	0.892 (0.941)	1.258 (1.010)
Hospitality	2.657 (3.387)	2.113 (1.580)	5.487 (6.841)	2.373 (1.732)	5.341 (5.732)	2.430 (2.034)
Auto trade and repair	6.782 (2.682)	7.971 (3.298)	3.517 (2.083)	7.049 (3.508)	6.393 (3.141)	7.302 (4.088)
Public	4.560 (3.631)	4.464 (2.367)	8.226 (5.504)	5.570 (3.303)	4.431 (2.489)	5.042 (2.358)
Communications	0.934 (0.914)	1.299 (1.275)	0.303 (0.435)	1.444 (0.912)	0.654 (0.837)	0.862 (1.076)
Education	3.802 (2.100)	4.219 (2.297)	1.752 (1.395)	3.395 (1.814)	3.191 (2.251)	3.316 (2.252)
Healthcare	4.141 (2.808)	4.686 (2.801)	2.360 (1.402)	4.155 (2.530)	2.699 (1.883)	4.058 (2.551)
Other service	4.174 (2.445)	4.530 (2.653)	5.020 (2.536)	3.587 (1.805)	3.363 (2.491)	3.666 (2.434)
% Primary education or less	25.049 (5.968)	20.719 (4.935)	29.606 (10.086)	18.928 (6.905)	26.737 (6.322)	24.557 (6.115)
% Secondary education	63.541 (5.100)	66.956 (4.795)	56.956 (8.601)	67.881 (3.080)	61.548 (5.929)	65.504 (6.904)
% Tertiary education	6.670 (3.609)	8.741 (3.907)	7.351 (7.284)	9.508 (8.767)	6.283 (3.641)	6.721 (2.993)
No. of observations	949	2,896	6	10	148	41
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011

Notes: Standard deviations are in parenthesis. In all calculations Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table A.7: Long-run Differences in Labour Market at MAL, no military areas, 2011.

	In Population density	In Labour force density	Unemployment rate
	(1)	(2)	(3)
<i>Polynomials in distance to border</i>			
Linear polynomial	-0.310*** (0.094)	-0.315*** (0.096)	2.704*** (0.542)
Quadratic polynomial	-0.214** (0.104)	-0.218** (0.103)	1.706** (0.536)
Cubic polynomial	-0.121 (0.115)	-0.134 (0.116)	2.566*** (0.523)
<i>Polynomials in latitude and longitude</i>			
Linear polynomial	-0.246*** (0.083)	-0.248** (0.086)	3.617*** (0.519)
Quadratic polynomial	-0.322*** (0.081)	-0.321*** (0.083)	3.419*** (0.539)
Cubic polynomial	-0.305*** (0.084)	-0.308*** (0.086)	3.274*** (0.553)
Mean dep. var. in interior	4.036 (0.874)	3.296 (0.899)	10.51 (4.806)
Observations	4033	4033	4033
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Military areas	Excluded	Excluded	Excluded
Border zone	Included	Included	Included

Notes: All reported estimates are coefficients of the variable 'In borderlands'. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and polynomials in distance to the MAL interacted with the treatment or polynomials in longitude and latitude. In all regressions military areas, Prague and Polish Zaolzie are excluded. Observed units are municipalities.

Table A.8: Long-run Differences in Educational Attainment at MAL, no military areas, 2011.

	% Primary education or less	% Secondary education	% Tertiary education
	(1)	(2)	(3)
Linear polynomial in distance to border	4.852*** (0.632)	-3.904*** (0.515)	-1.914*** (0.388)
Linear polynomial in latitude and longitude	4.964*** (0.544)	-3.756** (0.452)	-2.271*** (0.374)
Mean dep. var. in interior	20.77 (4.973)	66.94 (4.832)	8.713 (3.902)
Observations	4033	4033	4033
Clusters	70	70	70
Border segments	50	50	50
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011
Military areas	Excluded	Excluded	Excluded
Border zone	Included	Included	Included

Notes: All reported estimates are coefficients of the variable ‘In borderlands’. Robust standard errors are in parenthesis and clustered at the district level, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions include border segment and district fixed effects, controls for geographic factors, and linear polynomial in distance to the MAL interacted with the treatment or linear polynomial in longitude and latitude. In all regressions military areas, Prague and Polish Zaolzie are excluded. Observed units are municipalities.