Cite as:
Guzi M, Huber P, Mikula S, 2021, The long-term impact of the resettlement of the Sudetenland on residential migration. Journal of Urban Economics, 126, 1–17
https://www.sciencedirect.com/science/article/abs/pii/S009411902100067X

The Long-term impact of the resettlement of the Sudetenland on residential migration*

Martin Guzi, Peter Huber, and Štěpán Mikula†

March 29, 2021

We analyze the long-term impact of the resettlement of the Sudetenland after World War II on residential migration. This event involved expulsion of ethnic Germans and an almost complete depopulation of an area of a country and its rapid resettlement by 2 million Czech inhabitants. Results based on a regression discontinuity design show a highly persistent higher population churn and thus a lower attachment of residents to their region in resettled areas. Descriptive evidence also indicates that resettled settlements still have fewer local club memberships, less frequently organize local social events and had lower turnout in municipal elections until the 1990s. This thus suggests persistently lower levels of local social capital. This finding is consistent with recent theoretical models that suggest a highly persistent impact of the destruction of local social capital on residential migration.

Keywords: Migration, Social Capital, Sudetenland
JEL Codes: N44, Z10, R23, J15
Declaration of interest: none.

*Financial support by the Czech Science Foundation (grant No. 18-16111S) is gratefully acknowledged. We thank Alžběta Hanáčková, Barbora Kuklová and Marek Breza for excellent research assistance and Martin Halla, Benjamin Elsner, Tomáš Dvořák and Peter Egger, as well the participants of the 2017 research seminars at the University of Economics in Prague and the University of Stavanger, 2017, the 10th Geoffrey J.D: Hewings Regional Economics Workshop in Vienna, the 2018 Slovak Economic Association Meeting, the 2018 Scottish Economic Society Annual Conference, the 2019 ISWG workshop in Vienna, the 2019 GFR Winterseminar in Matrei and the 2019 Annual Meeting of the Austrian Economic Association, the 33rd Annual Conference of the European Society for Population Economics in Bath, the 2019 European Economic Association Annual Congress in Manchester, and the 2019 European Association of Labour Economists Conference in Uppsala for helpful comments. Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and Innovations Infrastructures.

†Martin Guzi: Department of Public Economics, Masaryk University. Address: Lipová 41a, Brno, Czech Republic. Email: martin.guzi@econ.muni.cz, Peter Huber: Austrian Institute for Economic Research. Address: Arsenal Objekt 20, Wien, Austria. Email: peter.huber@wifo.ac.at, Štěpán Mikula: Corresponding author. Department of Economics, Masaryk University. Address: Lipová 41a, Brno, Czech Republic. Phone: +420 724 512 946. Email: stepan.mikula@econ.muni.cz.

1 Introduction

A growing body of economic and political science research documents that the disruption of social structures caused by forced mass emigration has long-lasting and sizeable effects on the political attitudes (Acemoglu et al. 2011; Grosfeld et al. 2013), institutional, economic and educational outcomes (Acemoglu et al. 2011; Akbulut-Yuksel and Yuksel 2015; Pascali 2016; Bharadwaj et al. 2008; Testa 2020) and scientific achievements (Waldinger 2010, 2011) of the affected regions as well as on the forced migrants and their descendants (Becker et al. 2020). This literature (see: Becker and Ferrara 2019, for a survey) suggests that these impacts are substantially more long-lasting and sizeable than those found in the related literature on the destruction of physical capital (Brakman et al. 2004; Waldinger 2016). The long-run demographic impact of mass emigration has so far been much less analyzed. Among the exceptions Schumann (2014) focuses on population growth in a receiving region to provide evidence of a highly persistent demographic impact of the mass immigration to West Germany after World War II. The current paper adds to this literature by using the post-WWII mass expulsion of ethnic Germans from the area of today's Czech Republic (the so-called Sudetenland) to study the long-term impact of the subsequent resettlement on migration to and from the sending resettled regions.

This focus may be interesting because migration is generally seen as a behavioural measure of the attractiveness of a region as a place of residence and the attachment of its population to the region (see e.g.: Greenwood et al. 1991). It may, also be interesting because the expulsion and subsequent resettlement we study resulted in an almost complete destruction of the social structures and networks in the affected region. Recent theoretical contributions to the social capital literature (see e.g., David et al. 2010; Bräuninger and Tolciu 2011) suggest that such large shocks may have highly persistent effects on the attachment of residents to the region. In these models there are typically two stable long-run equilibria: One with low and one with high local social capital. Furthermore, since individuals derive utility from income and social capital they may forego financially profitable mobility. As a consequence, the high local social capital equilibrium is associated with low mobility, and the low local social capital equilibrium with high mobility. A sufficiently large shock to local social capital can move regions from one of these equilibria to the other, thereby causing a persistent change in migration rates from and to a region that goes hand in hand with persistent shift in local social capital levels.

The empirical part of this literature confirms this strong impact of current and previous social contacts for mobility and suggests that in particular local contacts are an impediment to mobility, while contacts to other regions enhance it. For example Kan (2007) and Belot and Ermisch (2009) show that people with contacts in their region of residence are

less inclined to move elsewhere, while Büchel et al. (2020) show that mobile individuals prefer to live in places with more nearby contacts. This literature, however, considers the contemporaneous correlation between social capital and migration (i.e. the short-run). Yet recent evidence by Costa et al. (2018) also indicates a persistent long-run impact impact of social contacts on individual's location decisions, as civil war veterans serving in the same military unit tend to live close to each other in the long-run. We therefore add to this literature by analyzing the long-run impact of the resettlement of the Sudetenland, which caused an almost complete destruction of the social structures and networks, on mobility and presenting additional descriptive evidence on the long-run development of local social capital levels in the resettled municipalities.

The resettlement of the Sudetenland is particularly well suited to identify such long-run impacts due to the size of the population exchange, the extraordinary speed with which ethnic Germans had to leave and the rapid resettlement that occurred in the follow-up. In the municipalities we study at least 90% of the population were ethnic Germans, who were subjected to expulsion and by and large completely replaced by Czech speaking settlers within five years. With the exodus of the German speakers also all local social networks and contacts of the resident population were destroyed and the newly arriving population had to re-establish all social contacts anew. This provides for large scale and credibly exogenous shock to local social capital in the affected region. In addition the fleeing population had to leave behind all of their belongings as well as all physical capital, and was rapidly replaced by settlers, who mostly came from other parts of the Czech Republic and took possession of the belongings and physical capital left behind. This limits the role of some potential explanations such as physical capital destruction or cultural or institutional differences, which have been emphasized as alternative sources for long-term differences in regional development in previous research (Grosfeld et al. 2013; Becker et al. 2015; Dell 2010; Alesina and Giuliano 2015).

Finally, this paper also contributes to the still rather scant literature on the resettlement of the Sudetenland (see e.g., Daněk 1995; Testa 2020). Earlier work on this resettlement (Daněk 1995) used descriptive evidence on the district level to show that districts of the Czech Republic that were at least partially located in the Sudetenland, still have a younger, less well educated, more secularized and ethnically more diverse population than other districts of today's Czech Republic. More recently, Testa (2020), in the paper most closely related to this one, uses municipality level data and a spatial regression discontinuity design based on the borders drawn by the Munich Agreement, to show that former German municipalities still had a lower population density, higher rates of unemployment, less skill-intensive industries, and lower levels of education in 2011. He argues that a lack of agglomeration economies and the erosion of property rights caused by resettlement are

potential mechanisms that may have caused these differences. We augment these findings by a detailed analysis of the impact of resettlement on migration and by proposing an additional causal mechanism that may have contributed to these very long-run impacts.

We combine a unique municipality-level administrative data set that includes all permanent residence changes in the years 1971 to 2015 with pre-WWII municipality-level data on the ethnic composition of municipalities in 1930. This allows us to identify the causal effects of resettlement on residential migration by comparing the most strongly affected municipalities with more than 90% ethnic Germans in 1930 (which we refer to as *resettled municipalities*) to municipalities with less than 10% ethnic German residents in 1930 (referred to as *not resettled municipalities*). We use a regression discontinuity (RD) design identification strategy where the inference is based on a precise definition of the border between ethnic German- and Czech-dominated areas before WWII. The results indicate that resettlement led to a long-lasting increase in residential migration to and from the resettled municipalities that survived such important institutional changes as the transition from a planned to a market economy, the dissolution of Czechoslovakia and accession to the European Union, and the many economic changes that occurred in the Czech Republic in that period. Even at the end of our observation period, in 2015 (i.e., 70 years after the resettlement), emigration and immigration rates in resettled municipalities were still substantially (by around 20%) higher than among not resettled municipalities. In addition, at the beginning of the period studied the effects of resettlement on emigration dominate over those on immigration, such that net emigration from resettled municipalities initially increased as well. This effect, however, levels off to zero after the mid-1980s, and is also less robust than the effect on gross-emigration and immigration rates. This suggests that—consistent with theoretical predictions—resettlement mainly resulted in a very long-run reduction in the attachment of the population to the affected regions.

To provide additional descriptive evidence on the potential causal mechanisms for this long-run impact we analyze auxiliary data sets. We explore two mechanisms. The first one assumes that the original settlers moving into the Sudetenland belonged to the more mobile groups of the population at the time of resettlement and may have transferred their values related to mobility to their children, who are the migrants we study at the end of our observation period. The second is based on the assumption that resettlement led to a persistent large scale reduction in local social capital of the resettled territory, and thus moved these regions from the high social capital, low mobility to the low local social capital, high mobility equilibrium suggested by the theoretical literature.

These analyses show that the population in the resettled and not resettled municipalities share similar mobility related values and also do not differ in terms of their behaviour with respect to proxy variables for social capital that is not clearly bound to the locality (such

their inclination to give donations to national organisations, as well as their willingness to attend social events and to participate in voluntary work potentially outside their region of residence). We, however, also find evidence that resettled municipalities had lower turnout rates in municipal elections up to the 1990s and have fewer local activities and clubs as well as lower membership levels in local clubs to this day. This suggests that local social capital is less well developed in the resettled regions. While therefore we cannot fully preclude some role for the parental transmission mechanism in explaining the long run impact of resettlement on migration decisions, we, therefore, argue that—consistent with many of the descriptions of these municipalities by historians (Glassheim 2006; Vaněk 1996; Čapka et al. 2005; Spurný 2011; Matějka 2008) and with theoretical models of the impact of local social capital on migration decisions—persistently lower local social capital in resettled municipalities is likely to have at least contributed to the persistently higher residential in- and out-migration in these municipalities.

The remainder of the paper is structured as follows. Section two provides the historical background of the resettlement. Section three presents the data. Section four introduces the identification strategy used. Section five describes the results. Section six discusses the mechanism driving the results, and section seven concludes.

2 Historical background

2.1 Ethnic Germans in the Czech Republic

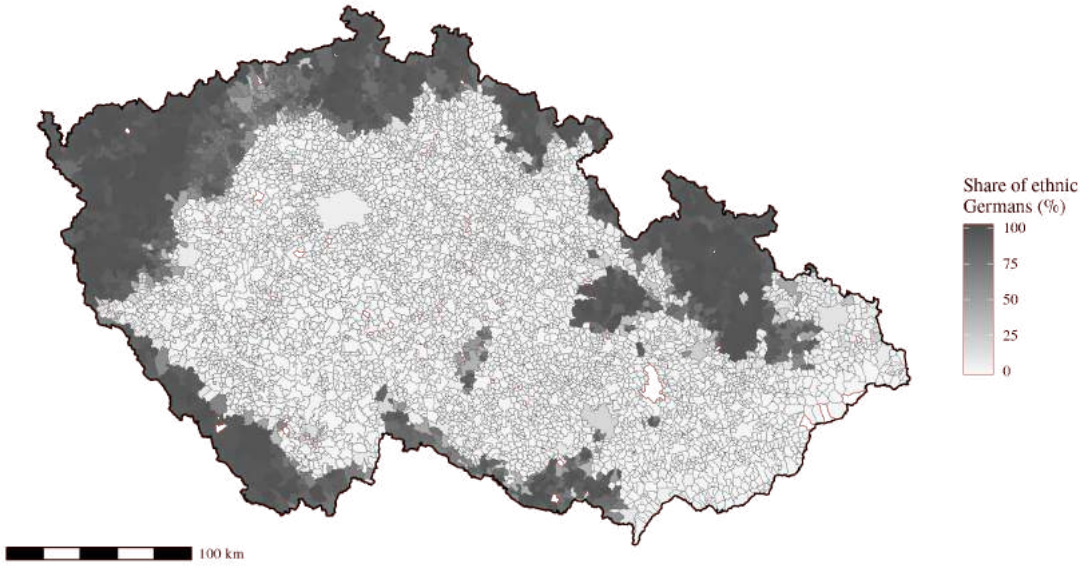
Germans had been settling in the area of today's Czech Republic since the 13th century and, except for a short period from 1938 to 1945 under German occupation, their territory of settlement was always administered by the same state as the rest of today's Czech Republic. Compared with the Jewish population studied in the literature, the ethnic Germans in the Czech territory never belonged to a discriminated group (e.g., Alexander 2008; Meixner 1988). During the Austro-Hungarian Empire, of which today's Czech Republic was a part from the 16th century to 1918, they tended to be more privileged relative to the Czech population partly because they spoke the official language of the empire.

The German speaking population was, however, clearly segregated from the Czech population. According to the last Czechoslovak census prior to World War II the German speaking population mainly resided in a well defined territory in the north, west, and south of today's Czech Republic, the so-called Sudetenland (see Figure 1),[1] and the number municipalities where both Germans and Czechs resided in equal numbers was rather low

1. The word Sudetenland is used in common language to denote the territories settled by the ethnic Germans before World War II.

(see Figure 2). The differences in economic and population structure between the two territories were small according to the results of this census. In particular, while the region where German speakers dominated were less agricultural and more industrial (see Tables A.1 and A.2 in the Appendix) the share of migrants living in the two regions, which is the only proxy for migration provided in the 1930 census, were rather similar in the Czech and German speaking parts of today's Czech Republic (see Figure 3).[2]

Figure 1: Share of ethnic Germans in municipalities in 1930

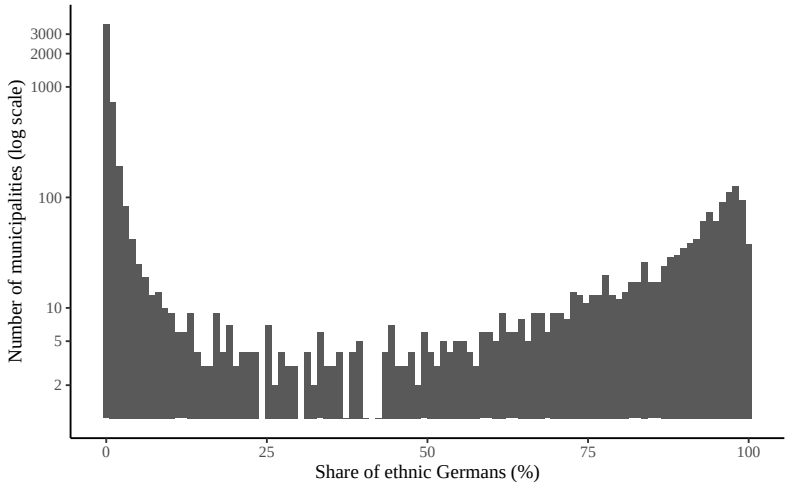


Source: CZSO, own calculations.
Note: The ethnic German population is defined according to the primarily spoken language. The 1930 municipality-level data on ethnic Germans is harmonized with 6,168 municipalities defined in the 2011 census by using matching rules provided by the CZSO.

Historic accounts also suggest that the relationships between the two population groups were relatively unproblematic for most of the time. Ethnic tensions between Czechs and Germans arose only in the second half of the 19th century and continued after the break-up of the Austro-Hungarian Empire, when ethnic Germans comprised 29.5% of the population, according to the 1930 population census (see CZSO 2014). Historical records document a number of complaints from German representatives during the interwar period inter alia about limited access to employment in the state bureaucracy, the closing of German language schools, and the asymmetric impact of land reforms in the 1920s. Yet, according to many accounts (e.g., Glassheim 2000) minority policy in Czechoslovakia was one of the most liberal in Central and Eastern Europe at the time. Ethnic tensions severely intensified

2. Data on employment structure and share of migrants are not available at municipal level in 1930 census. Therefore, we use data aggregated at the level of 330 judicial districts. For the definition of 1930 judicial districts see Figure A.1 in the Appendix.

Figure 2: Distribution of municipalities by the share of ethnic Germans in 1930



Source: CZSO, own calculations.
Note: The ethnic German population is defined according to the primarily spoken language. The 1930 municipality-level data on ethnic Germans is harmonized with 6,168 municipalities defined in the 2011 census by using matching rules provided by the CZSO.

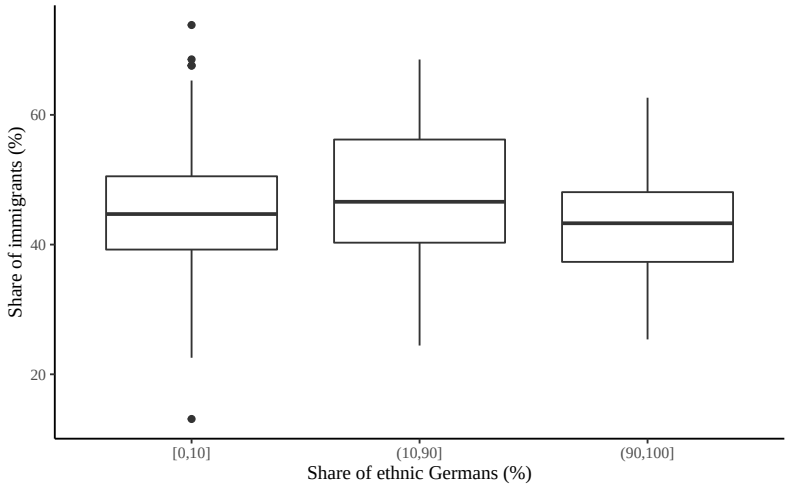
only with the economic crisis in 1933 and the increasing popularity of German nationalist political parties. Under the Munich Agreement in 1938 the Sudetenland was annexed by the German Reich and remained under German rule until the end of WWII.

2.2 Expulsion of the Germans and resettlement

In the aftermath of WWII, the ethnic Germans were held responsible for the Nazi atrocities and considered traitors. This perspective ultimately led to their expulsion from the country, which started with the end of WWII in May 1945 and proceeded in two phases. The initial phase, referred to as *"wild expulsion",* was poorly organized and controlled within a vague legal framework. Up to 800,000 Germans left the country in this phase until the autumn of 1945 (Wiedemann 2016). The second more organized phase continued from January to October 1946 and followed the agreements of the Potsdam Conference. The mass deportation during the second phase reduced the prewar population of ethnic Germans of 3 million to 200 to 300 thousand (Gerlach 2017). The share of ethnic Germans in the Czech population decreased from 29.5% (based on the 1930 census) to 1.8%[3] based on the first postwar census in 1950 (CZSO 2014).

3. Data on ethnicity from the 1930 and 1950 censuses are not fully comparable due to methodological changes. Nevertheless, estimates of historians are very similar to the 1950 census data. According to Staněk (1991) there were 216,545 inhabitants of non-Slavic origin in 1947. The remaining Germans were also subjected to an internal relocation policy (Dvořák 2013).

Figure 3: Share of immigrants from other regions in total population in judicial districts in 1930



Source: CZSO, Urbánní a regionální laboratoř (URRlab) UK (2015), own calculations.
Note: The ethnic German population is defined according to the primarily spoken language. For definition of 1930 judicial districts see 330 see Figure A.1 in the Appendix.

The process of resettlement occurred in parallel with the expulsion and was also rather rapid. Initially, people were encouraged through newspapers and radio broadcasting to seize German properties and were supported by Czech soldiers, militias, and security forces (Glassheim 2000). During the *wild expulsion,* (i.e. within the first year of expulsion) between 500,000 and 900,000 new settlers arrived in the Sudetenland.[4] In addition Wiedemann (2016) states that the most massive inflows of settlers occurred until 1947 (i.e. one year after the beginning the second phase of deportation and two years after the end of WWII) and that resettlement was largely over by the end of 1952, with only modest inflows continuing on until the end of the 1950s.[5] Similarly, Gerlach (2010) finds that almost 2 million new ethnic Czech settlers had arrived to Sudetenland by May 1947, while earlier estimates by Radvanovský (2001) suggest an influx of 1.5 million over the first 2 years of resettlement. In sum thus resettlement was rather rapid and was clearly already completed in 1971 when our period of analysis starts.

4. Radvanovský (2001) estimates that 514,515 settlers arrived in Sudetenland by September 16, 1945. Wiedemann (2016) states that their numbers reached 696,554 by mid-October and 862,706 by the end of 1945.

5. This extraordinary speed of resettlement is also documented in a number of anecdotes according to which new settlers ended up cohabiting with those to be expelled when deportation trains were delayed (e.g. Wiedemann (2016), p. 106).

2.3 The process of property acquisition by new settlers

For the Czech citizens resettlement to the Sudetenland offered a unique opportunity to improve their economic and social status by acquiring a house and small piece of land (the official limit was 0.13 square kilometers)[6] and thus also seizing the expellees' other property,[7] as well as to obtain a better job, or to become a national administrator[8] of seized properties (Wiedemann 2016). The key institutions for assigning property to the new settlers were about 100 local resettlement committees (*"Osidlovací komise"*). They were responsible for collecting applications and redistributing land and property. They were also responsible for appointing national administrator for the vacated enterprises. Wiedemann (2016) reports that, given the "spontaneous" nature of the wild expulsion the work of the resettlement committees was rather complicated, inconsistent and also intransparent. Committees had only limited knowledge of the situation in general and occasionally allocated properties that were already taken and also caused many other irregularities. Settlers had to pay for the acquired property. Prices were low, and amounted to one to three yearly rents. Ten percent of the total price was due at the time of property acquisition. The remainder was payable in the following 15 years. Settlers could pay in cash or kind and a substantial part of the liabilities was never paid (Wiedemann 2016).

The process of property redistribution with respect to enterprises and arable land was, however, also rather inconsequential, as the communist party decided to collectivize the land and the industrial property throughout the Czech Republic after the 1948 coup d'état. This process started in 1949 and was by and large completed in the mid-1950s (i.e. shortly after resettlement had been completed). After this the Czech Republic remained to be a country with one of the lowest shares of private property of land and firms throughout the communist era, even among COMECON countries, and private ownership of "means of production" was virtually non-existent until the early 1990s. The decisions of the resettlement committees were therefore relevant only for home-ownership, as this was the only part of the property of the expelled ethnic Germans that remained in private hands after the collectivization.

6. Čapka et al. (2005) document that 91% of the settlers in the villages around Mikulov previously owned land with acreage less than 0.03 square kilometers, and the typical acreage redistributed to settlers around Mikulov was 0.05 to 0.08 square kilometers.

7. The expellees could keep 30 kg and later 50 kg of their belongings (excluding valuables) during the more organized phase of the expulsion (Gerlach 2017).

8. This was a trustee who could manage an enterprise (or large farm) on behalf of the state with the prospect of becoming the owner (Gerlach 2017).

2.4 Characteristics of settlers

The weak regulation (and often chaotic nature) of the resettlement makes it difficult to map the socioeconomic characteristics of settlers. Wiedemann (2016), Čapka et al. (2005) and Školl (1983) show, however, that settlers were young, often married couples, landless persons, small farmers, second-born children (with low prospects for family inheritance), or individuals who had worked in the civil services or non-agricultural sector. It is also highly likely that the settlers belonged to the more mobile groups in population.

The resettlement policy aimed to attract people primarily from areas that were geographically close and climatically similar in order to increase the chances of settlers establishing economically and socially functional communities (Wiedemann 2016). Historical research documents that settlers almost exclusively moved from other regions of today's Czech Republic[9] and over rather short distances. For example, Školl (1983) documents that only 16% of the settlers in Břeclav district moved less than 10 km, and 39% moved more than 100 km (Figure A.2 and Table A.3 in the Appendix). Settlers likely moved in smaller groups, however. Detailed statistics on the original municipalities of settlers in the Břeclav district reveal that 12% of settlers were from the same miunicipality of origin and 29% of settlers originated from the four most important municipalities (Table A.4 in the Appendix). In addition, Čapka et al. (2005) describe Sobotín in North Moravia as a municipality with a large group of settlers from the same original municipality, but even in this example, the settlers from the same original municipality comprise around 15% of the 1950 population.

These major migratory movements resulted in two types of municipalities. *Not resettled municipalities* were inhabited by ethnic Czechs before WWII and were therefore not subject to expulsion and resettlement. The *resettled municipalities* were inhabited by ethnic Germans before WWII and thus lost most of their original population together with their human and social capital. The resettlement process brought new inhabitants to the emptied municipalities, who were allowed to seize the property left behind by the expelled but had to re-establish all local social contacts anew. In consequence resettlement also implied a massive destruction of local social capital in the resettled regions.

9. Školl (1983) notes that in the Břeclav district, located at the border with Slovakia and Austria, 90.3% of settlers were Czech, 2.1% were Slovak, and 7.6% were of other nationalities (Table A.5 in the Appendix).

3 Data

3.1 Migration and population data

We investigate how the resettlement of Sudetenland affected the subsequent migration behavior in the resettled municipalities. The immigration and emigration rates as the key dependent variables of this analysis are taken from administrative records of permanent residence changes provided by the Czech Statistical Office (CZSO) for the period 1971 to 2015.[10] These rates provide the number of movers from and to municipalities as a percentage of the population on January 1 and apply to a period when according to historical sources resettlement had clearly been completed, with the German-speaking population having been gone from the municipality and the country for 20 years or more.[11] This data is highly reliable because residents of former Czechoslovakia (and of the current Czech Republic) are legally obliged to register changes of their permanent address. This registration also defines the constituency in municipal elections and is associated with preferential access to local public services such as healthcare, elementary schools, kindergartens, and subsidized accommodations for university students. We obtain the pre-WWII share of ethnic Germans in each municipality from the 1930 population census (the last census before WWII), in which ethnicity is defined according to the language primarily spoken in the household.[12] We use this data to identify resettled and not resettled municipalities and to construct the ethnic border.

Table 1 presents descriptive statistics of the pooled sample. We report averages and standard errors of migration variables, population, and the share of ethnic Germans in 1930 separately for the resettled and not resettled municipalities. The emigration rates were on average higher in the 1970s and 1980s compared with later periods, and the immigration rates are more stable over time. Both emigration and immigration rates are consistently higher in the resettled municipalities relative to not resettled ones throughout the observation period, but this difference narrows in later periods. Figure 4 adds to this information by presenting the size distribution of resettled and not resettled municipalities for individual census years. It shows that although resettled municipalities were in general larger prior to resettlement, their size reduced substantially after resettlement and they

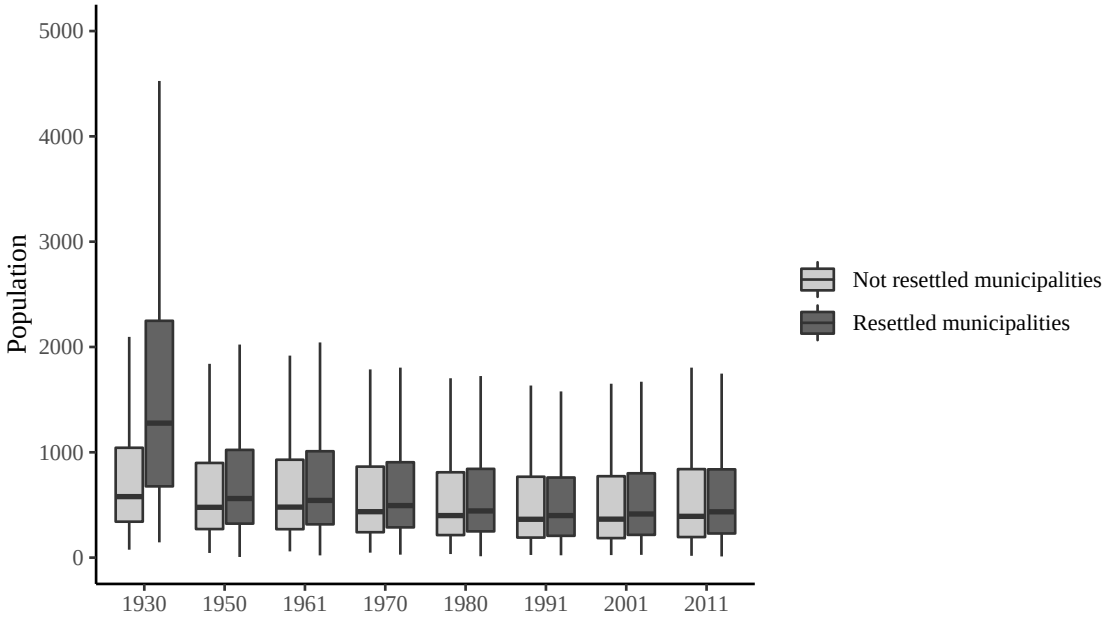
10. Data is available at www.czso.cz/csu/czso/databaze-demografickych-udaju-za-obce-cr. We have data for 6,168 municipalities due to some records being lost prior to digitization and due to definition changes; this is an unbalanced panel. The baseline model is estimated with pooled data on all municipalities, and we address concerns that relate to missing observations in the robustness analysis in the Appendix to the paper.

11. Annual population data are obtained from the CZSO. We remove obvious outliers (i.e., the top 1% of emigration and immigration rates) from the data. In the robustness analysis we show that the inclusion of outliers slightly increases the estimated effects.

12. Data from the 1930 census was digitized by the authors (see Appendix B). Matching rules provided by the CZSO were applied to harmonize municipalities such as in the 2011 census.

also continued to loose population until the 1980s. Afterwards the size distribution of both types of municipalities remained rather stable and the population of both types of municipalities was also rather similar.

Figure 4: The evolution of the size distribution of resettled and not resettled municipalities (1930 to 2011)



Source: CZSO, own calculations.

3.2 Local social capital and values data

In addition, to provide evidence on potential causal mechanisms, we collected data on values and different forms of social capital at a municipality level. This includes administrative data on voter turnout (i.e., the percentage of eligible voters who cast a ballot) in all free and voluntary municipal elections held in the Czech Republic provided by the CZSO. This, following the social capital literature (Knack 1992; Hotchkiss and Rupasingha 2018), is used as a proxy for local social capital. Further this includes survey data on values, civic participation, and charitable activities collected in 2003 and 2004, respectively.[13] These survey data are, according to our review of the sources, the only publicly available data in the Czech Republic that allow us to examine shared norms and values as well as social capital at the municipality level. The 2003 survey (Majerová et al. 2003) was conducted in two waves and provides information on both local social capital as well as forms of social capital not

13. This data is described in more detail in the Appendix B of this paper.

Table 1: Descriptive statistics

		Municipalities			Difference
		All	Resettled	Not resettled	(2)−(3)
	Period	(1)	(2)	(3)	(4)
Municipalities (n)		6,168	751	4,808	
Share of ethnic Germans (%)	1930	18.43	95.92	0.48	95.43***
		(0.45)	(0.10)	(0.02)	
Population (n)	1930	1,725.54	1,997.30	1,319.73	677.57**
		(172.51)	(117.27)	(202.09)	
Emigration rate (%)	1971–1979	3.50	5.05	3.19	1.85***
		(0.008)	(0.03)	(0.008)	
	1980–1989	3.11	4.34	2.88	1.46***
		(0.008)	(0.03)	(0.008)	
	1990–1999	2.44	3.12	2.29	0.83***
		(0.006)	(0.02)	(0.007)	
	2000–2009	2.37	3.09	2.21	0.88***
		(0.006)	(0.02)	(0.007)	
	2010–2015	2.51	3.26	2.34	0.92***
		(0.008)	(0.03)	(0.008)	
Immigration rate (%)	1971–1979	2.38	3.38	2.14	1.24***
		(0.009)	(0.03)	(0.009)	
	1980–1989	2.35	3.14	2.19	0.95***
		(0.008)	(0.03)	(0.009)	
	1990–1999	2.41	3.21	2.24	0.97***
		(0.008)	(0.03)	(0.009)	
	2000–2009	3.03	3.64	2.90	0.74***
		(0.009)	(0.03)	(0.01)	
	2010–2015	3.05	3.53	2.94	0.59***
		(0.01)	(0.03)	(0.01)	
Net immigration rate (%)	1971–1979	−1.12	−1.67	−1.05	−0.62***
		(0.01)	(0.04)	(0.01)	
	1980–1989	−0.76	−1.20	−0.69	−0.51***
		(0.01)	(0.03)	(0.01)	
	1990–1999	−0.02	0.09	−0.05	0.14***
		(0.009)	(0.03)	(0.01)	
	2000–2009	0.67	0.55	0.70	−0.14***
		(0.009)	(0.03)	(0.01)	
	2010–2015	0.54	0.28	0.61	−0.33***
		(0.01)	(0.04)	(0.01)	

Note: Columns (1) to (3) contain means and standard errors (in parentheses). Column (4) contains differences in means between resettled and not resettled municipalities and their statistical significance with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

bound to the region of the resettled and not resettled municipalities and their inhabitants. In the first wave, mayors of 1,324 municipalities were asked about the abundance of certain forms of local social capital in the municipality (i.e the frequency of local events organized by local clubs and local green activities). In the second wave, 1,287 residents in 223 municipalities were asked about memberships in local clubs, participation at social event events and donations, and volunteer work in potentially national organisations. Among these variables the first is a measure of local social capital, while the others pertain to aspects of social capital that are not necessarily bound to the municipality of residence. The 2004 survey (Majerová et al. 2004) asked 1,518 respondents in 220 municipalities about the importance of values in their life (i.e., nature and environment, a job, relationships, faith and spiritual values, hobbies, housing, friendship, family life and children, and material conditions).

3.3 Geographical data

Since our identification strategies require geographic information, we geocode all municipality-level data using the reference points defined by the CZSO[14] and obtain altitude and terrain roughness data for each municipality using remotely sensed elevation data from Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global.[15] Altitude is measured as the elevation at the reference point. Terrain roughness is calculated as the average of terrain roughness index around the reference point[16] such that higher values indicate rougher terrain.

4 Identification strategy

Formerly German-dominated municipalities were clearly segregated from the Czech-dominated municipalities and mainly located in the borderlands of today's Czech Republic. This segregation as well as the specific location of formerly German-dominated municipalities makes direct comparison of resettled and not resettled municipalities problematic as they could differ in unobserved characteristics that are correlated with their respective status (such as local labor market conditions, market access, etc.). This may lead to OLS estimation leading to biased estimates. To address this concern we identify causal effects by

14. These are placed in the social center of municipalities (i.e., typically in front of the town hall or church)

15. The SRTM data for the Czech Republic are transformed into approximately 30×40 meter tiles, which is the maximum homogeneous resolution available for the SRTM data of the Czech Republic. The terrain roughness index is calculated as the mean of the absolute differences between the altitude of a tile and the altitudes of its eight surrounding tiles (Wilson et al. 2007).

16. We use a 1.6 km radius throughout because this is the median radius of municipalities in the Czech Republic.

using a RD-design (Dell 2010; Becker et al. 2015; Egger and Lassmann 2015; Oto-Peralías and Romero-Ávila 2017). We estimate the following specification:

$$y_{it} = \gamma \mathrm{RM}_i + \beta \mathbf{Z}_{it} + f(d_i) + \phi_r + \phi_t + \phi_s + \xi_{it} \tag{1}$$

where y_{it} is one of the three outcome variables (emigration, immigration, and net immigration rates) defined for municipality i in year t, RM_i is an indicator variable for a resettled municipality, and $\mathbf{Z}_i$ includes geographical variables (log of altitude, the terrain roughness, the log of shortest distance to the Czech border and second order polynomials of the degree of latitude and longitude of the municipality). The year fixed effects ϕ_t are included to control for changes in institutions and economic developments affecting all municipalities alike, ϕ_r are region fixed effects that control for any region-specific features of the outcome variable (or equivalently for the border segment).[17] As the migration rates tend to be slightly higher in larger municipalities (see Figure A.4 in the Appendix) we also include population fixed effects ϕ_s (decile of the 1930 population). These account for potential differences in migratory moves between municipalities of different sizes.[18] Variable η_{it} is the error term.

To define the ethnic border, we treat all municipalities with a minority (less than 50%) of ethnic Germans in 1930 as Czech municipalities and all municipalities with a majority of ethnic Germans in 1930 as German municipalities.[19] The ethnic border is then located at the border of German municipalities contingent to a Czech municipalities. This delineation yields a compact Czech territory in the center of today's Czech Republic surrounded by what once were German municipalities in the south, west, and north of the country and a few former German enclaves surrounded by Czech municipalities and Czech enclaves surrounded by former German municipalities. From these, we removed two enclaves (one resettled and one not resettled municipality, respectively) that were smaller than 10 square kilometers to ensure that we could define a border region sufficiently wide for an RD-regression. The distance of municipalities to the ethnic border is defined by the distance of their reference points to the ethnic border.

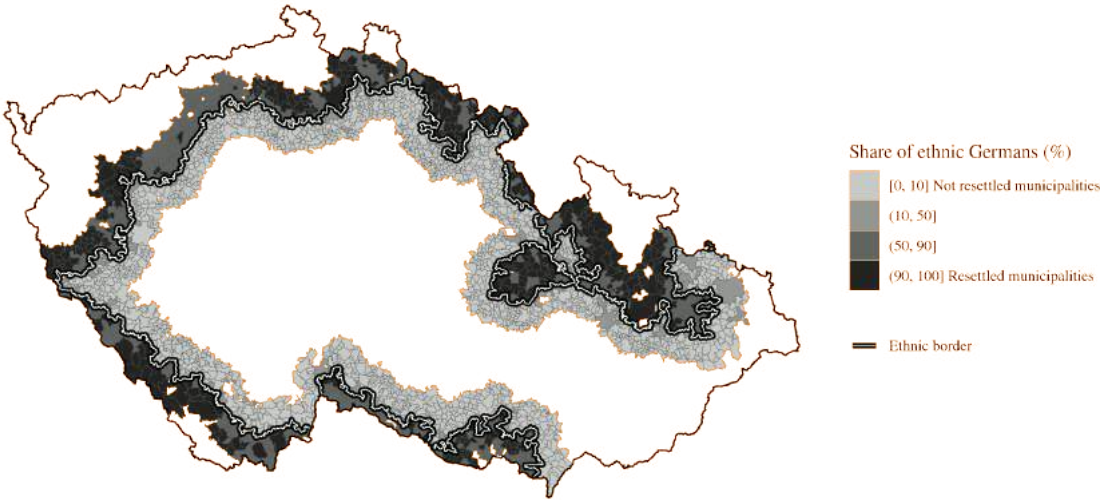
One issue with respect to the RD-design based identification strategy is that we measure distance to the border by defining municipality centers that are rarely located at the defined border. This implies that we have only few municipalities in the vicinity of the border (see

17. The Czech Republic is administratively divided into 14 regions *("kraj").* We checked that the borders of regions do not coincide with the ethnic border (Figure A.3 in the Appendix).

18. Population in 1930 is used as a control to avoid including endogenous controls and to account for the fact that the once larger municipalities in the Sudetenland may provide more infrastructure and cultural amenities than municipalities outside the Sudetenland.

19. We use the administrative areas of municipalities from a series of maps (ArcČR 500 Version 3.3) developed by ARCDATA PRAHA to construct the ethnic border. These are also used for all geospatial visualization (https://www.arcdata.cz/produkty/geograficka-data/arccr-500).

Figure 5: The ethnic border between Czech- and German-dominated municipalities and municipalities within a 15 km band around the border

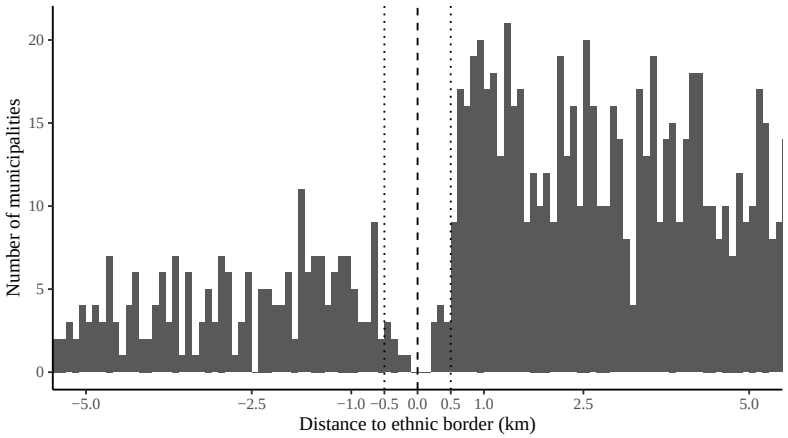


Source: CZSO, own calculations.

Figure 6). As pointed out by e.g. Cunningham (2021) this can make RD-estimates based on non-parametric optimal bandwidth estimators unreliable and inrobust. In our baseline estimates we therefore rely on a RD-approach that includes all resettled and not resettled municipalities located within 15 kilometers of the ethnic border to ensure symmetry in the number of observations at certain distances on both sides of the ethnic border (Figure 5). In our robustness section we, however, also provide estimates based on local linear regressions as well as on a different identification strategy based on spatial matching, to assess the impact of the border definition and the chosen RD-estimator on results.

The central identification assumptions of the regression discontinuity design are that (1) all unobserved confounding variables affecting migration decisions vary smoothly at the ethnic border but that (2) the share of ethnic Germans has a distinct jump at that border. Figure 7 therefore shows discontinuity plots for the geographic variables and the share of the German-speaking population in 1930 living on each side of the ethnic border for the estimation sample as the only municipality level pre-treatment variables for which we have data. There are no visible discontinuities in altitude, terrain roughness, distance to the Czech border, and population in 1930, but there is a sizeable discontinuity in the share of Germans at the cut-off.

Figure 6: Distance to ethnic border



Source: CZSO, own calculations.
Note: Resettled municipalities have negative distances to the ethnic border, and not resettled municipalities have a positive one. Municipalities are aggregated to bins of 0.1 km.

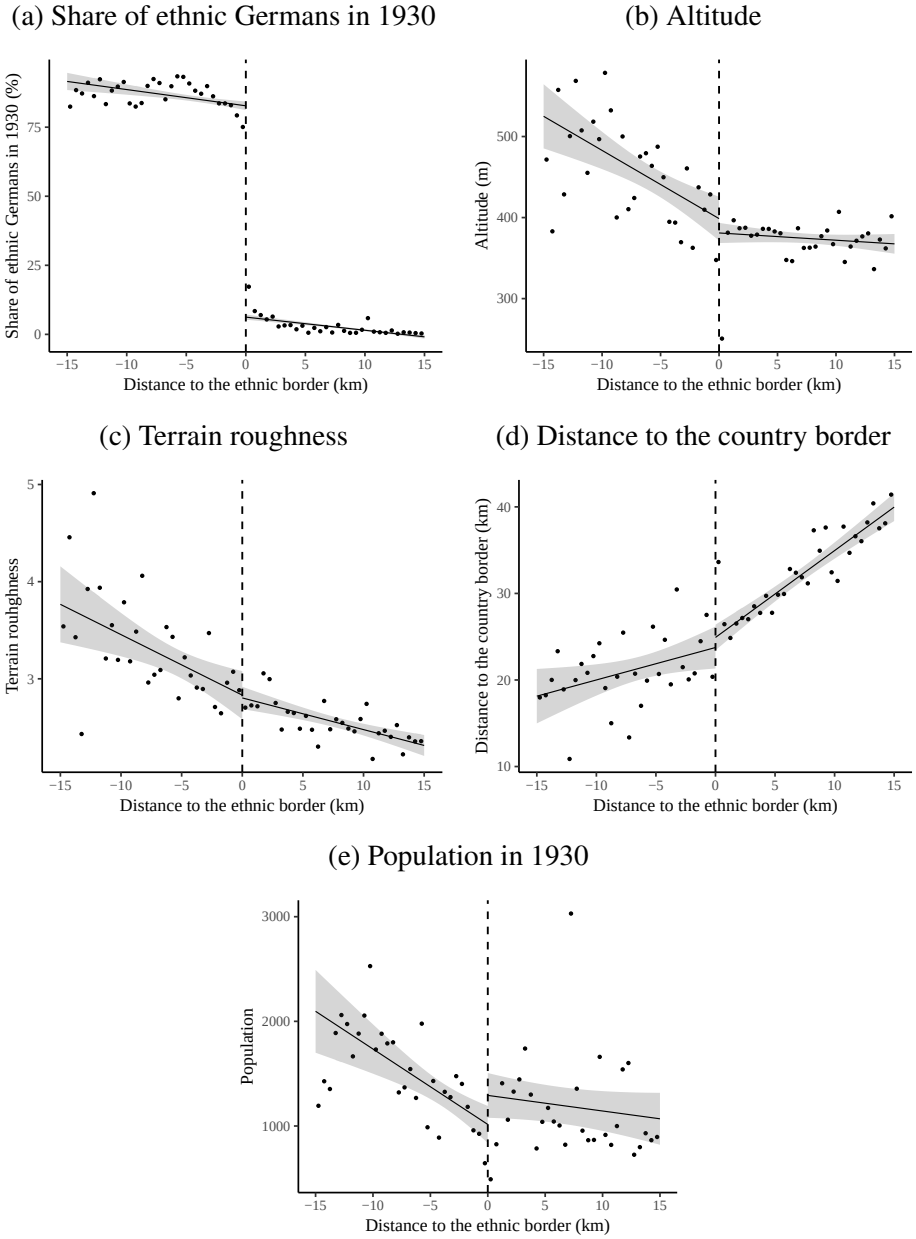
5 Results

5.1 Baseline results

Figure 8 shows similar regression discontinuity plots for the three dependent variables for the pooled data over all time periods in the analysis. This figure suggests a distinct jump in the emigration an immigration rates as well as the net immigration rate at the ethnic border. According to this figure municipalities on the former German speaking side of the border were characterized by in average higher emigration and immigration rates as well as lower net immigration rates over the years 1971 to 2015.

While these graphs are indicative of a potential causal effect of the resettlement on mobility, the estimates they present fail to control for other confounding factors that may have contributed to different migration levels between the two regions and also pool across the entire time period from 1971 to 2015. In particular the later fact may be problematic on account of the data in the early observation period also including the subsequent mobility of the settlers, originally moving into the region. As these settlers almost by definition belonged to the more mobile groups of the Czech population at the time of resettlement this may bias the causal estimates upward. Table 2 therefore presents several baseline estimates of the parameter of interest (γ) of Equation (1). The first panel (of two columns) presents estimates for the average mobility rates over the entire observation period using the complete complete panel information, while the second panel, to reduce potential issues with serial correlation of error terms, focuses on the average mobility rate for the entire

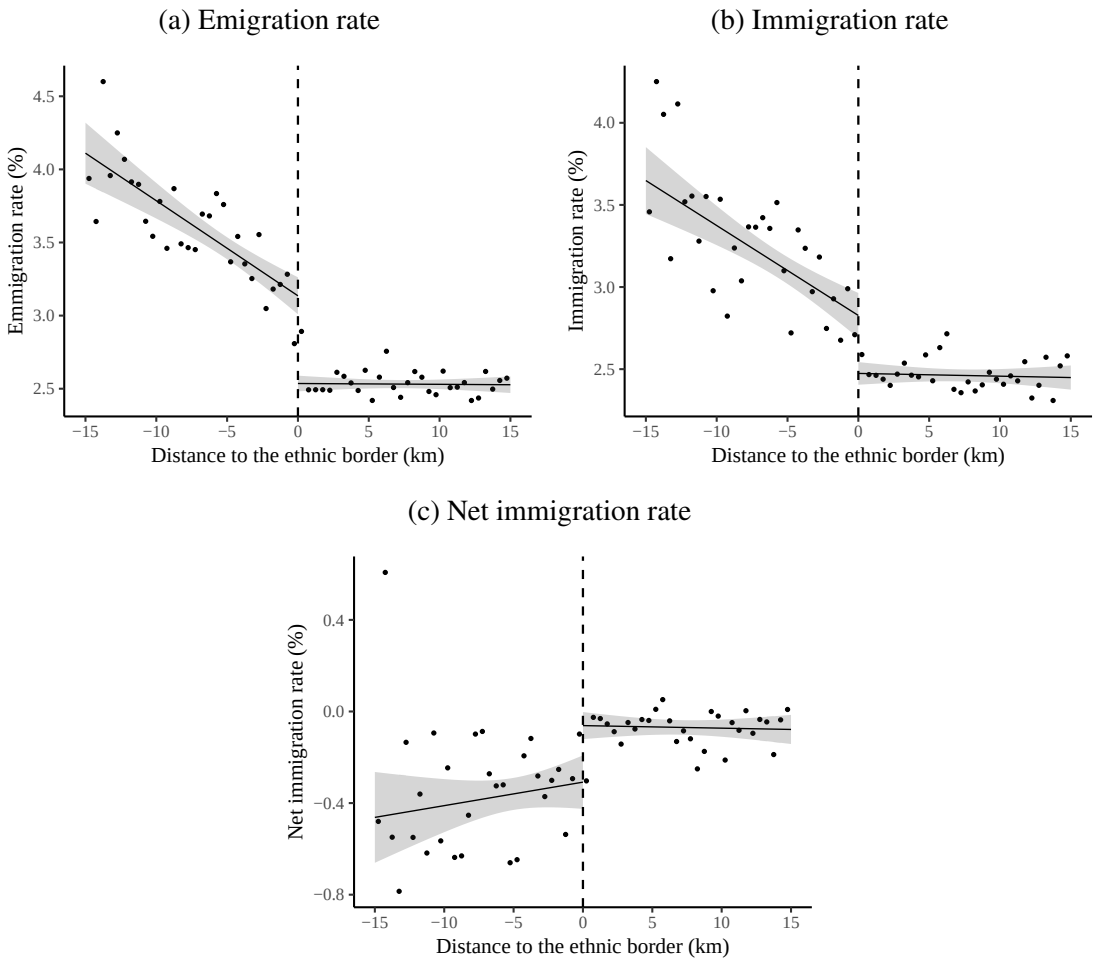
Figure 7: Discontinuity plots of geographic variables, population, and share of ethnic Germans



Source: CZSO, SRTM, own calculations.
Note: Resettled municipalities have negative distances to the ethnic border, and not resettled municipalities have a positive one. Municipalities are aggregated to bins of 0.5 km represented by points. The smoothing lines are parametric linear predictions and shaded areas represent 95% confidence intervals.

period. Panels three and four by contrast present estimates for two different sub-periods (1971 to 1989 and 1990 to 2015). Further each of these panels presents estimates for two different RD-polynomials (in their two columns) using standard errors clustered by

Figure 8: Discontinuity plots of migration rates at the ethnic border



Source: CZSO, own calculations.
Note: Resettled municipalities have negative distances to the ethnic border, and not resettled municipalities have a positive one. Municipalities are aggregated to bins of 0.5 km represented by points. The smoothing lines are parametric linear predictions and shaded areas represent 95% confidence intervals.

municipality for panel estimates and robust standard errors for estimates using aggregated data.

The estimates focusing on the entire period, irrespective of whether they apply to the full panel or averages over the period, suggest that resettlement increased emigration rates from the resettled municipalities by 0.7 percentage points and immigration rates by approximately 0.5 to 0.6 percentage points per year on average. Both these effects are statistically significant at the 0.1% level in all estimations. In addition, according to the results net immigration decreased by 0.1 to 0.2 percentage points per year on average, with these effects being statistically significant only when considering panel estimates.

Results, however, differ for the sub-periods from 1971 to 1989 and 1990 to 2015. In the first period the effects on emigration rates are substantially and effects on immigration rates slightly higher than for the overall period and also effects on net immigration are negative and statistically significant at the 1% significance level. In this earlier period resettlement led to a substantial increase of both emigration (of 1.2 to 1.3 percentage points) and immigration (0.6 to 0.7 percentage points) and also caused net immigration rates to these regions to fall (or equivalently net emigration from these regions to rise). This may, however, be due to the original settlers belonging to the more mobile groups of their population and some of them potentially also having been disappointed with their new live in the resettled municipalities.

Table 2: Impact of the resettlement on migration rates

	Panel estimates		Cross-sectional estimates using data aggregated by period					
			1971–2015		1971–1989		1990–2015	
	RD-polynomial		RD-polynomial		RD-polynomial		RD-polynomial	
	1st order	2nd order	1st order	2nd order	1st order	2nd order	1st order	2nd order
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable: Emigration rate (‰)							
Resettled municipality (=1)	0.735***	0.690***	0.685***	0.664***	1.303***	1.196***	0.474***	0.400***
	(0.063)	(0.092)	(0.066)	(0.096)	(0.105)	(0.156)	(0.062)	(0.090)
Adjusted R^2	0.169	0.169	0.444	0.444	0.453	0.454	0.359	0.359
Observations	82423	82423	2104	2104	1882	1882	2104	2104
	Dependent variable: Immigration rate (‰)							
Resettled municipality (=1)	0.504***	0.465***	0.573***	0.506***	0.589***	0.675**	0.512***	0.417**
	(0.070)	(0.102)	(0.084)	(0.122)	(0.129)	(0.207)	(0.090)	(0.129)
Adjusted R^2	0.091	0.091	0.320	0.319	0.268	0.269	0.314	0.314
Observations	82423	82423	2104	2104	1882	1882	2104	2104
	Dependent variable: Net immigration rate (‰)							
Resettled municipality (=1)	−0.231***	−0.225*	−0.111	−0.159	−0.714***	−0.521**	0.038	0.017
	(0.062)	(0.092)	(0.083)	(0.121)	(0.133)	(0.198)	(0.085)	(0.126)
Adjusted R^2	0.096	0.096	0.102	0.102	0.234	0.235	0.124	0.124
Observations	82423	82423	2104	2104	1882	1882	2104	2104

Note: Estimates from Equation (1). Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Somewhat more surprisingly the increased mobility also applies to the 1990 to 2015 period, where the original adult settlers were over 60 years old and the majority of migration moves was accounted for by the children of the original settlers. In particular while the effects of resettlement on net migration are statistically insignificant and also very close to zero for this period, emigration and immigration rates in the resettled municipalities still exceeded those in not resettled municipalities by 0.4 to 0.5 percentage points. This suggests that in this very long-run, resettlement mainly resulted in increased population churning and thus a reduction in the attachment of the population to the affected regions.

5.2 Effects over time

Figure 9 extends these findings by plotting estimates of year-treatment interaction terms, based on a linear RD-specification.[20] It, therefore, provides a more complete picture of the dynamics of the effects of resettlement on migration. The effects once more indicate a particularly pronounced increase in emigration and immigration rates in the treated municipalities in the 1970s and 1980s. In the 1970s, emigration rates were as much as 2 percentage points and immigration rates as much as 1.5 percentage points higher in the resettled municipalities (relative to a baseline emigration rate of 3.2% and an immigration rate of 2.1% in the not resettled municipalities). The overall effect on net immigration in this period was negative, that is, resettled municipalities' population decreased because of emigration.

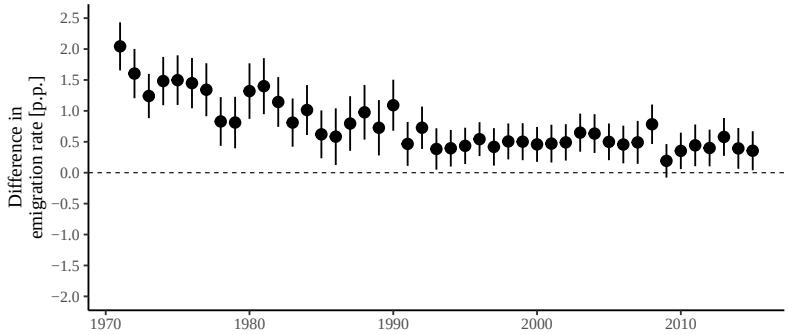
By the 1990s the gap in net immigration rates, however, narrowed and stabilized at a level comparable to those of the not resettled municipalities for the remainder of the observation period. By contrast, gross emigration from the resettled municipalities remained statistically significantly higher than in not resettled municipalities in all years except for 2013. Furthermore, while these differences declined during the 1970s and 1980s, they remained at approximately 0.4 percentage points higher in the resettled than in not resettled municipalities thereafter. Similar observations apply to immigration rates, although these differences are not statistically significant in all years.

Thus while impacts on net migration rates levelled off during the 1980s, the impact on population churning persisted over a period over which the original settlers in the resettled areas went through almost their entire life cycle, as an individual born at the beginning of resettlement in 1945 would have been 70 years old in 2015. This stability of the estimates for the effects on immigration and emigration from the 1990s on, implies that the effect was affected neither by the many political and institutional changes experienced by the Czech Republic in that time period nor by the rapidly changing macroeconomic environment.

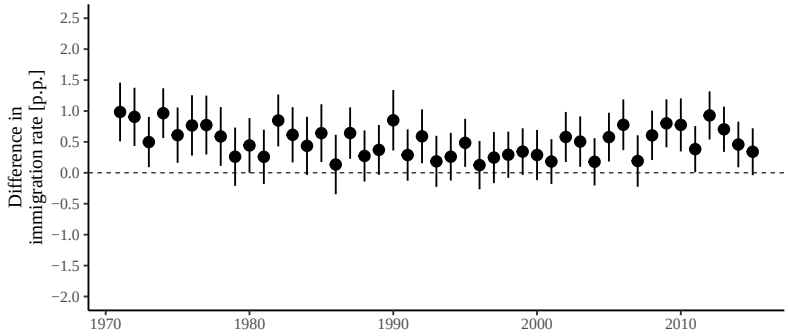
20. This is chosen because it minimizes the Bayesian Information Criterion (BIC).

Figure 9: Impact of the resettlement on migration rates by year

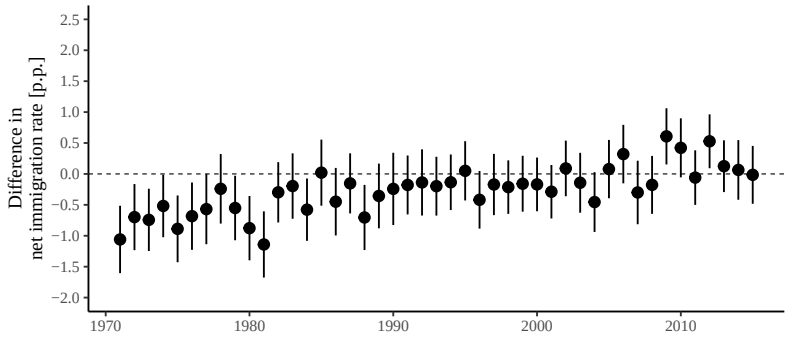
(a) Emigration rate



(b) Immigration rate



(c) Net immigration rate



Note: Figure reports estimates of the impact of the resettlement on the migration rates from Equation (1). Only municipalities within a 15 km band around the ethnic border are considered. Results control for the interaction of the year fixed effects with the dummy variable for resettled municipality, interaction of year fixed effect with a first-order polynomial forcing function, and region fixed effects, year fixed effects, population fixed effect, log of altitude, terrain roughness, and log of the distance to the country border. Points represent the point estimates and the associated bars the 95% confidence intervals of the estimates. Standard errors are clustered by municipality.

There are no visible changes in the causal impact at the time of fall of the Iron Curtain and the start of economic transition (1989), dissolution of Czechoslovakia (1993), and the accession to the European Union (2004). Similarly, the estimates are very stable throughout the deep (so-called transition) recession at the beginning of the 1990s, the Czech currency crisis in 1997 and the following recession, the economic crisis in the late 2000s, and the great depression of 2009.

5.3 Effects by education groups

Table 3 further extends these finding, by using data on changes in permanent residence by education groups available from CZSO for the years 1993 to 2004, to estimate Equation (1) separately for the highly and less highly educated.[21] This data pertains to a shorter time period than that used in baseline estimates.[22] Nonetheless it indicates that the impact of resettlement on population churning in the later period of analysis was more pronounced among the more highly educated, who according to most results in the social capital literature (see Putnam 1995; Helliwell and Putnam 1999; Glaeser et al. 2002) both contribute and depend less on social capital.

Thus, resettlement led to a sizeable and statistically highly significant increase in emigration rates (of around 2.1 percentage points) of the highly educated. Similarly their immigration rates increased 2.8 to 3.8 percentage points. Among the less educated effects were more modest, as their emigration rate was by 0.4 to 0.5 percentage points higher in the resettled than in not resettled municipalities and their immigration rates were by 0.4 to 0.6 percentage points higher. Results for net immigration rates, by contrast, are once more much less robust for both the highly and the less educated and statistically significant only for the highly educated, when using a first-order RD-polynomial.

5.4 Falsification tests and robustness

One concern regarding these findings is that—as aforementioned—the resettled municipalities are located closer to the German border. This could threaten the causal interpretation of our results if municipalities closer to the border of the country are in general less attractive for living. This may be of particular relevance in the context of communist Czechoslovakia before 1989 because in those times, due to heavy policing of the border, borders to "Western countries" were particularly unattractive places to live. We therefore conduct two falsification tests to address this concern. In these we counter-factually shift

21. Less highly educated are persons with compulsory education and highly educated with more than compulsory education.

22. There are no annual municipality-level data on population by education. Therefore, in this analysis we calculate yearly migration rates using 1991 census data on population in each education group.

Table 3: Impact of the resettlement on migration rates by education

	Sample			
	Primary education		Secondary and tertiary education	
	RD-polynomial		RD-polynomial	
	1st order	2nd order	1st order	2nd order
	(1)	(2)	(3)	(4)
	Dependent variable: Emigration rate (%)			
Resettled municipality (=1)	0.533***	0.368**	2.139***	2.058***
	(0.079)	(0.114)	(0.341)	(0.472)
Adjusted R^2	0.077	0.078	0.044	0.044
Observations	25188	25188	25164	25164
	Dependent variable: Immigration rate (%)			
Resettled municipality (=1)	0.605***	0.372†	3.772***	2.841**
	(0.173)	(0.208)	(0.733)	(0.993)
Adjusted R^2	0.042	0.042	0.062	0.062
Observations	25188	25188	25164	25164
	Dependent variable: Net Immigration rate (%)			
Resettled municipality (=1)	0.072	0.004	1.632**	0.783
	(0.147)	(0.177)	(0.575)	(0.824)
Adjusted R^2	0.008	0.008	0.016	0.016
Observations	25188	25188	25164	25164

Note: Panel estimates for Equation (1) for period 1993–2004. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, and region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomials of longitude and latitude. Symbols represent statistical significance of differences in means between resettled an not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

the ethnic border of the Sudetenland 5 kilometers away from, respectively, 5 kilometers toward the border of the Czech Republic in sub-samples of resettled and not resettled municipalities. These placebo treatments (Table 4) indicate no statistically significant differences between municipalities at the counterfactual ethnic border for emigration and net immigration rates irrespective of where the border is shifted. The point estimates level off to zero for the counter-factual shift of the ethnic border away from the border of the Czech Republic and are only slightly larger in case of counterfactual shift toward the country border. However, none of these estimates is statistically significant.

To address other concerns about the baseline estimates we conduct a series of further robustness checks based on panel data estimates. These address issues related to the potential impact of spatial auto-correlation of error terms on the results (see Kelly 2019; 2020, for a discussion), the choice of the estimation method for the regression RD-specification and other potential methodological choices made in the estimation of the baseline specification The main results of these robustness tests are presented in Table 5 and more detailed descriptions are available in the Appendix C, which also reports results for more settings and estimates with higher order RD-polynomials.[23]

To assess the potential impact of spatial auto-correlation on results we adopted the correction suggested by Kelly (2020) and additionally implemented a permutation test where the empirical distribution of treatment effects under the null hypothesis is obtained from 1000 placebo treatments resulting from counter factually shifting the ethnic border between 20 and 35 km away from the border of the Czech Republic.[24] The results (reported in the first two rows of Table 5) suggest only minor increases in standard errors for immigration and net immigration rate, which do not affect statistical significance, for estimates based on the spatial auto-correlation corrected standard errors and also no change for the inference based on permutation tests.

Further to assess the impact of the RD-estimation method on results we re-estimated the regression in Equation (1) using more flexible specifications allowing for different RD-polynomials on both sides of the cut-off (by allowing for all combinations of first to third-order polynomials on both sides of the border) and also considered results from estimating local linear regressions based on several optimal bandwidth selection procedures.[25] Results for estimates based on different functional forms of the RD-Polynomials at the two sides of the border (presented in the Appendix in Table C.3) do not differ from

23. Equivalent baseline results estimated with 3rd order RD-polynomials are presented in Table A.6 in the Appendix.

24. Artificial shifts in the other direction cannot be considered, because the ethnic border lies too close to the border of the Czech Republic. Consequently this test rests on the assumption that the auto-correlation structure at the ethnic border is identical to that the counter factually constructed border.

25. These included using the Mean Square Error (MSE) and CER-optimal bandwidth selector to define optimal bandwidth.

Table 4: Placebo tests

	Placebo test			
	Placebo test on the sample of resettled municipalities (shift of the ethnic border by 5 km towards the country border)		Placebo test on the sample of not resettled municipalities (shift of the ethnic border by 5 km outwards the country border)	
	RD-polynomial		RD-polynomial	
	1st order	2nd order	1st order	2nd order
	(1)	(2)	(3)	(4)
	Dependent variable: Emigration rate (‰)			
Resettled municipality (=1)	0.100	0.054	0.010	0.002
	(0.121)	(0.142)	(0.042)	(0.053)
Adjusted R^2	0.227	0.227	0.112	0.112
Observations	15070	15070	67353	67353
	Dependent variable: Immigration rate (‰)			
Resettled municipality (=1)	0.163	0.179	−0.006	−0.024
	(0.126)	(0.148)	(0.063)	(0.086)
Adjusted R^2	0.070	0.070	0.081	0.081
Observations	15070	15070	67353	67353
	Dependent variable: Net immigration rate (‰)			
Resettled municipality (=1)	0.063	0.125	−0.016	−0.026
	(0.119)	(0.128)	(0.059)	(0.082)
Adjusted R^2	0.126	0.126	0.091	0.091
Observations	15070	15070	67353	67353

Note: Panel estimates from specifications equivalent to Equation (1) with counter-factually shifted ethnic border. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Table 5: Overview of robustness tests

		Dependent variable		
		Emigration rate (%)	Immigration rate (%)	Net Immigration rate (%)
		(1)	(2)	(3)
(1)	Standard errors adjusted for spatial autocorrelation	0.735^{***}	0.504^{***}	-0.231^{***}
		(0.062)	(0.076)	(0.068)
(2)	Statistical significance determined in a permutation test	0.735^{***}	0.504^{***}	-0.231^{***}
		(0.021)	(0.069)	(0.068)
(3)	Local linear regression	$0.209^{\dagger}$	0.398^{***}	$-0.156^{\dagger}$
		(0.107)	(0.093)	(0.091)
(4)	Local linear regression: Exclusion of 0.5 km band	0.778^{***}	0.519^{***}	-0.285^{***}
		(0.060)	(0.064)	(0.068)
(5)	Local linear regression: Optimum bandwidth doubled	0.568^{***}	0.464^{***}	-0.182^{**}
		(0.061)	(0.054)	(0.056)
(6)	(Spatial) matching estimates	0.782^{***}	0.607^{***}	-0.175^{**}
		(0.053)	(0.061)	(0.053)
(7)	Defining resettled and not resettled municipalities using a 50% cutoff	0.647^{***}	0.414^{***}	-0.233^{***}
		(0.043)	(0.048)	(0.044)
(8)	Exclusion of a part of ethnic border close to Bavaria (West Germany)	0.659^{***}	0.456^{***}	-0.203^{**}
		(0.064)	(0.072)	(0.065)
(9)	Inclusion of linear time trends for each region and period (before/after 1990)	0.738^{***}	0.505^{***}	-0.233^{***}
		(0.063)	(0.070)	(0.063)

Note: Table contain estimates of impact of resettlement on residential migration. Rows (1), (2), (7)–(9) contain estimates from (extended) Equation (1). Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Rows (3)–(5) contain results obtained from local linear regression using MSE optimum bandwidth selector. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, first-order RD-polynomial, region (defined by nearest foreign country) fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Row (6) contains estimates based on spatial matching strategy. Results control for year fixed effects, matched-pair fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Values in brackets are standard errors adjusted for spatial auto-correlation in row (1); standard deviations of the empirical distributions obtained in permutation tests in row (2); robust standard errors in rows (3)–(5); standard errors clustered by municipality in rows (6)–(9). Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

the baseline. Estimates obtained from local linear regressions show a substantial decline in estimated effects and are highly sensitive to the choice of specification (see row 3 in Table 5 and Table C.4 in the Appendix). This is due to the low number of observation located directly at the border, which makes local linear estimates based on optimal bandwidth selection procedures unstable and sensitive to outliers. Thus when we exclude the 0.5 km band along the ethnic border, which contains 10 not resettled and 7 resettled municipalities, and adjust distances accordingly the local linear estimates in row (4) converge to parametric ones. Similarly doubling the bandwidth, that reduces sensitivity to sparse observations in the proximity of ethnic border, yields point estimates reported in row (5), that are close to baseline results.

Further, to assess the impact of additional methodological choices we made in implementing the baseline RD-analysis, we conducted a number of additional robustness tests. These included implementing an alternative identification strategy based on spatial matching (Becker et al. 2015)[26] to test whether our results are corroborated by a method which does not require a definition of the ethnic-border (see row 6 in Table 5 for results); defining resettled municipalities as municipalities with a share of ethnic Germans of 50% or more and all others as being not resettled, to assess the impact of focusing only on municipalities with more than 90% respectively 10% ethnic Germans on results (see row 7 for results) and excluding the municipalities located close to Bavaria (West Germany) to avoid potential asymmetric impact from the opening of the border in 1990 on the attractiveness of the Czech-German border region as a place of residence (see row 8 for results); controlling for different trends in regional development by extending (1) with linear time trends for each region and period prior and after 1990 (see row 9 for results).[27]

These additional estimates are once more highly robust and comparable to our baseline results. Throughout the estimated coefficients for the emigration rate range between 0.7 and 0.9 percentage points, and those for the immigration rate between 0.5 and 0.7 percentage points. The changes in results are thus mostly smaller than 0.1 percentage points relative to our baseline estimates. The estimates for the net immigration rate are less robust.

In summary, this evidence is consistent with a strong medium-term causal negative impact of resettlement on net immigration on new municipalities, which could be associated with return or onward migration of the original settlers. More important, this evidence

26. Row (6) in Table 5 shows estimation results on within-pair comparison of resettled and not resettled municipalities that are located within 10 km from each other and did not differ in population by more than 250 in 1930.

27. We also conducted robustness tests to check how our results are affected by the ethnic border irregularities in the northeast of the Czech Republic, the use of a balanced panel data, the exclusion of outliers and the use of municipality size classes based on the 1950 population as a control affect results. Moreover, we estimate the baseline specification separately for northern, western, and southern part of the country. These additional checks provide no further insights over the previous analysis. They are, therefore, reported in the Appendix to the paper only.

suggests a long-lasting effect of resettlement on immigration and emigration rates and thus the population churn. This implies a persistently lower attachment of people in new municipalities that continued to persist even 70 years after the beginning of resettlement.

6 Mechanisms

The fact that the fleeing population from Sudetenland left behind all of their belongings as well as all physical capital, and was replaced by settlers, who mostly came from other regions of the Czech Republic and took possession of the belongings and physical capital left behind suggests that cultural differences, differences in administrative rules and effects from the destruction of physical capital, which have been emphasized as alternative sources for long-term differences in regional development in previous research (Grosfeld et al. 2013; Becker et al. 2015; Dell 2010; Alesina and Giuliano 2015), are unlikely to explain the long-run impacts found.

As a consequence we consider the plausibility of two alternative potential causal mechanism for the long-term reduction in the attachment of the population to their region of residence caused by resettlement. The first, is based on the self-selection of settlers to resettled municipalities: Since the Sudetenland was in all likelihood resettled by the more mobile groups among the Czech population at the time, it could be that migration-related values were transferred across generations among the population of the resettled municipalities such that these are still populated by more mobile population groups to this day. The second, accords with many of the descriptions of the resettled municipalities by historians (Vaněk 1996; Glassheim 2006; Matějka 2008; Čapka et al. 2005; Spurný 2011) and is derived directly from the predictions theoretical models of the impact of social capital on migration decisions (e.g., David et al. 2010; Bräuninger and Tolciu 2011). It assumes that the destruction of local social capital in resettled municipalities could have moved these municipalities to a low social capital and high mobility equilibrium.[28]

Although, due to data constraints, we cannot provide a causal inference with respect to these mechanisms, we examine a number of further data sets (described in Section 3), to assess their plausibility. These provide information on the differences in values of the population of resettled and not resettled municipalities and their activities related to local social capital (the participation in local elections, the number of clubs and local green activities in a municipality and membership in local clubs) as well as to forms of social

28. The empirical findings by Charnysh (2019) also supports the plausibility of this causal mechanism. She studies the impacts of the settlement process that followed post-WWII westward shift of Poland and finds that districts settled by migrants heterogeneous in place of origin are lower in prevalence of volunteer fire brigades, characteristic linked to local social capital level, as of 1989.

capital that are not bound to a specific locality (such as attending social events in general and volunteering to work with or to give donations to potentially national aid organisations).

6.1 Values

Regarding social values, we examine data from a survey on the values of the old (aged 60–74 years) and young (aged 18–29 years) generation in the Czech Republic (*Český venkov 2004: život mladých a starých lidí"*). We run a linear regression on the responses to the value-related questions in this survey on a dummy variable that takes the value of one if the respondent resided in a resettled municipality and zero if the respondent resided in an not resettled municipality. In these regressions we control for gender, age, previous migration, education, labor market status, marital status, household size and municipality population, and county fixed effects.[29] The results (Table 6) suggest no statistically significant differences in the values between the resettled and not resettled municipalities among the old or young residents. This finding applies to all values except the importance attributed to hobbies by old inhabitants; however, this could just as well be caused by a type II error.

29. Due to the smaller number of municipalities covered in this survey the regression discontinuity design approach is feasible – see Figure B.1 in the Appendix.

Table 6: Differences in values between resettled and not resettled municipalities

	Things and values very important in personal life (=1)									
	Nature, environment	Job, occupation	Relationships	Faith, spiritual values	Hobbies	Housing	Friendship	Health	Family life and children	Material conditions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Regression results					
					Panel A: Young cohort (18–29)					
Resettled municipality (=1)	0.049	0.113	−0.032	−0.029	0.076	0.042	0.072	−0.046	−0.062	0.004
	(0.075)	(0.080)	(0.084)	(0.046)	(0.077)	(0.075)	(0.066)	(0.090)	(0.087)	(0.088)
Adjusted R^2	0.039	0.050	0.043	0.039	0.021	0.018	0.037	0.044	0.226	0.016
Observations	680	680	680	680	680	680	680	679	680	679
					Panel B: Old cohort (60–74)					
Resettled municipality (=1)	−0.019	0.030	−0.045	−0.015	0.231***	0.006	0.004	−0.078†	−0.076	0.108
	(0.087)	(0.096)	(0.094)	(0.080)	(0.086)	(0.081)	(0.092)	(0.045)	(0.084)	(0.103)
Adjusted R^2	0.026	−0.011	0.022	0.084	0.006	0.022	0.010	0.043	0.109	0.036
Observations	611	602	609	608	610	611	609	610	610	610
					Descriptive statistics					
					Panel C: Young cohort (18–29)					
Not resettled municipalities	0.579	0.661	0.441	0.073	0.253	0.659	0.511	0.822	0.657	0.394
	(0.020)	(0.019)	(0.020)	(0.011)	(0.018)	(0.019)	(0.020)	(0.016)	(0.019)	(0.020)
Resettled municipalities	0.591	0.750	0.443	0.057	0.250	0.727	0.534	0.747	0.602	0.398
	(0.053)	(0.046)	(0.053)	(0.025)	(0.046)	(0.048)	(0.054)	(0.047)	(0.053)	(0.053)
					Panel D: Old cohort (60–74)					
Not resettled municipalities	0.680	0.304	0.564	0.256	0.218	0.568	0.543	0.930	0.762	0.270
	(0.019)	(0.019)	(0.020)	(0.018)	(0.017)	(0.020)	(0.020)	(0.010)	(0.017)	(0.018)
Resettled municipalities	0.674	0.287	0.506	0.227	0.295	0.607	0.568	0.932	0.670	0.386
	(0.050)	(0.050)	(0.053)	(0.045)	(0.049)	(0.052)	(0.053)	(0.027)	(0.050)	(0.052)

Table reports estimated coefficients on an indicator variable for resettled municipalities after controlling for municipality (log of altitude, terrain roughness, log of distance to the country border, region fixed effect, and population fixed effect) and personal (age group, education, labor market status, marital status, household size, and for being born in the municipality of residence) characteristics. Standard errors clustered by municipality are reported in parentheses with: $^{\dagger} = p < 0.1$, $^{**} = p < 0.05$, $^{***} = p < 0.01$, $^{***} = p < 0.001$. Descriptive statistics are the mean and the standard deviations (in parentheses).

6.2 Social Capital

With respect to social capital we follow the literature (Knack 1992; Hotchkiss and Rupasingha 2018) and use voter turnout in municipal elections at the municipality level as one dependent variable[30] In addition, from the first wave of the survey "*Český venkov 2003 – situace před vstupem do EU.*" we use the responses of mayors to a question on the number of events organized by local clubs and green activities (i.e., collective activities to improve the environment and living conditions in the municipality) in their municipality. These data are explicitly related to a specific municipality and are therefore measures of local social capital. From the second wave of the survey we consider the responses to questions on whether the interviewed individuals are a member of a local club in their municipality and whether they participated in an event, have made a donation, or have committed to conducting voluntary work that are not necessarily located in the respondent's municipality. Of these variables only the first refers to a specific municipality and therefore measures local social capital. For the other three the questions posed in the questionnaire (explicitly or implicitly) also include contacts outside the specific municipality and therefore refer to forms of social capital not directly related to a specific locality.

Among these variables the voter turnout allows for using RD-design as in Section 5, on account of applying to all Czech municipalities. According to the results (Table 7),[31] voter turnout in municipal elections was statistically significantly lower in the resettled than the not resettled municipalities with the point estimates of this impact ranging from 2.5 to 2.7 percentage points on average. In addition, the results of year by year RD-design estimates for voter turnout (Figure 10) suggest that these differences between the resettled and not resettled municipalities have reduced continuously since the times of the fall of the Iron Curtain and attained a similar value in both types of municipalities in the 2010 and 2014 elections. Thus, resettlement also led to a reduced voter turnout in municipal elections, that was, however, not as long-lived as the impact on the population churn.

For the survey based variables, by contrast, an RD-design analysis is not feasible, on account of the surveys being based on only a sample of municipalities (see Figure B.3 in the Appendix), as for the values data, we therefore, only regress the responses at the municipality level to these questions on a dummy variable indicating that a particular municipality was a resettled municipality, and the same municipality-level controls also used in the remainder of this analysis. For the individual-level data, we additionally control for gender, age, home-ownership status, education, marital status, household size, and labor market status.

30. In municipal elections citizens are not allowed to vote outside their constituency. This rules out potential bias caused by, e.g., correspondence voting.

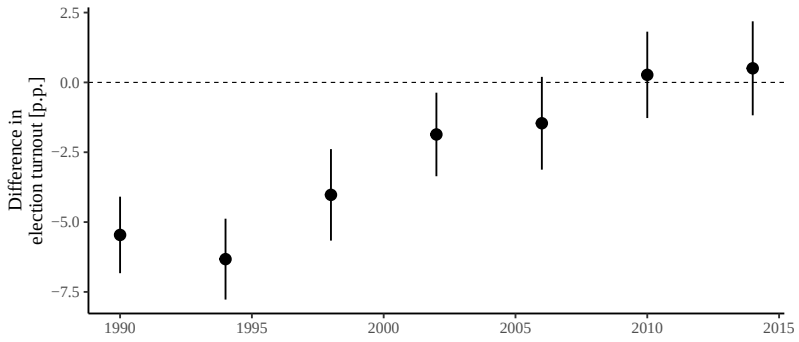
31. For descriptive statistic on voter turnout see Table A.7 in the Appendix.

Table 7: Impact of the resettlement on migration rates on voter turnout in municipal elections

	RD-polynomial	
	1st order	2nd order
	(1)	(2)
Resettled municipality (=1)	−2.538***	−2.718**
	(0.653)	(1.006)
Adjusted R^2	0.573	0.574
Observations	14423	14423

Note: Estimates from the specification equivalent to Equation (1) with voter turnout in municipal elections held between 1990 and 2014 being the dependent variable. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Figure 10: Impact of the resettlement on voter turnout in municipal elections by year



Note: Figure reports estimates of the impact of the resettlement on the voter turnout in municipal elections from the specification equivalent to Equation (1). Municipalities within a 15 km band around the ethnic border are considered. Results control for the interaction of the year fixed effects with the dummy variable for resettled municipality, interaction of year fixed effect with a first-order polynomial forcing function, and region fixed effects, year fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and second order polynomials of longitude and latitude. Points represent the point estimates and the associated bars the 95% confidence intervals of the estimates. Standard errors are clustered by municipality.

The results in Table 8 show statistically significant differences between the resettled and not resettled municipalities in variables associated with local social capital (i.e. the participation in events organized by local clubs, green activities, and club membership in the municipalities). These differences are also quantitatively important because they indicate that there are on average 2.8 fewer events organized by clubs and 0.9 fewer green events organized in the resettled compared with not resettled municipalities in 2003 and that the probability to participate in a local club or civil society organization is 0.1 to 0.2 percentage points lower in the resettled compared with the not resettled municipalities. The same does, however, not apply to indicators associated with non-localized forms of social capital (the frequency with which people attend events, provide donations, and participate in voluntary work). Consistent with our hypothesis this implies lower local social capital in resettled municipalities, and thus suggests that a persistently lower level of local social capital in the resettled municipalities may have contributed to the lower attachment of the population to their region.

Table 8: Differences in events organization, attendance, club membership and charitable activities between resettled and not resettled municipalities

	Panel A: Wave 1		Panel B: Wave 2			
	Events organized by local clubs (n)	Green activities (n)	Club membership (=1)	Events attendance (=1)	Donations (=1)	Voluntary work (=1)
	(1)	(2)	(3)	(4)	(5)	(6)
			Regression results			
Resettled municipality (=1)	−2.909*	−0.882***	−0.150†	0.003	−0.032	−0.030
	(1.297)	(0.310)	(0.082)	(0.062)	(0.083)	(0.066)
Adjusted R^2	0.086	0.032	0.055	0.115	0.094	0.069
Observations	1,181	1,181	1,525	1,525	1,523	1,523
			Descriptive statistics			
Not resettled municipalities	10.598	2.233	0.428	0.596	0.586	0.248
	(0.462)	(0.093)	(0.013)	(0.013)	(0.013)	(0.012)
Resettled municipalities	8.535	1.208	0.314	0.599	0.465	0.256
	(0.802)	(0.177)	(0.035)	(0.037)	(0.038)	(0.033)

Regression results report estimated coefficients on an indicator variable for resettled municipalities. Regressions in panels A and B control for the log of altitude, terrain roughness, log of distance to the country border, population fixed effects, and region fixed effects. Municipality-level regressions in the Panel A also control for age structure (shares of 20—29, 30—39, 40—49, 50—59, and 60—69 in total population) and the shares of the low and high educated in population 15+ in 2001. Individual-level regressions in Panel B control for gender, age group, home-ownership status, education, marital status, household size, and labor market status. Values in parentheses are robust standard errors in Panel A, and standard errors are clustered by municipality in Panel B: $^{\dagger} = p < 0.1$, $^{**} = p < 0.05$, $^{***} = p < 0.01$, $^{***} = p < 0.001$. Descriptive statistics are the mean and the standard deviations (in parentheses) for resettled and not resettled municipalities.

7 Conclusions

In this paper, we exploit the resettlement of the Sudetenland after WWII as a natural experiment to analyze the long-term impact of the destruction of the social structures of municipalities affected by this resettlement on residential migration. This large and unexpected resettlement involved the relocation of 2 million Czech inhabitants to municipalities previously populated by ethnic Germans. We find that resettlement led to a very long-term increase of the population churn from and to the affected municipalities. Even at the end of our observation period (i.e. 70 years after the resettlement) annual immigration and emigration rates were still by 0.4 to 0.5 percentage points higher (relative to a baseline of around 2.5%) in the affected resettled municipalities than in the unaffected not resettled ones without there being any indication of a decrease in these differences. In addition, in the period from 1970 to about 1989 next to a more pronounced increase in emigration and immigration also net immigration rates to the resettled municipalities was lower by 0.5 to 0.7 percentage points. This effects, however, levels off by the 1990s and may due to the migratory behaviour of the original settlers.

The longer-lived impact of resettlement on the population churn implies that even in 2015 the population in resettled municipalities was still less attached to their place of residence. Furthermore, the higher population churn survived important institutional changes and major macroeconomic crises that occurred in the Czech Republic in that time. These include the fall of the communist regime and the Iron Curtain, transition from a planned to a market economy, separation from the Slovak Republic, and accession to the European Union. Our evidence therefore suggests a persistent and long-term causal impact of resettlement on the attachment of the population to the regions affected.

Exploring the plausibility of different potential causal mechanisms for this long-run and stable impact of resettlement on the attachment of residents, we provide descriptive evidence that inhabitants of resettled municipalities share the same values as residents of municipalities that were not resettled and also do not differ in terms of their behaviour with respect to proxy variables for social capital not bound to a specific locality (such as their inclination to give donations to national organisations, as well as their willingness to attend social events and to participate in voluntary work potentially outside their region of residence). We, however, find evidence that resettled municipalities to this day are lower values in measures of local social capital (such as the number of local social activities organized and membership in local clubs). In addition, up until the early 2000s, resettled municipalities had lower voter turnout in municipal elections than municipalities not resettled. This finding leads us to conclude that lower local social capital was a likely contributor to the lower attachment of the population to the resettled municipalities.

These findings thus corroborate previous empirical evidence on the long-term impact that the destruction of the social structures of a region has on its economic, social, and political development. In particular, the finding of higher population churn and lower civic participation in clubs and similar communal activities in the resettled municipalities is consistent with a number of recent theoretical and empirical contributions on the role of local social capital in shaping migration decisions. These show that if local social capital yields utility and is endogenously invested in by the resident population two stable equilibria arise: one with low local social capital and high migration and one with high local social capital and low migration (Bräuninger and Tolciu 2011 or David et al. 2010). Interpreting these results from this perspective, the destruction of local social capital in resettled municipalities may have moved the resettled municipalities to a low local social capital and a high migration equilibrium while not resettled municipalities continued to be in high local social capital and low migration equilibrium. From a wider perspective this evidence points to the highly persistent effects that destruction of the local social structure of a region (as was the case in Sudetenland) can have on the mobility of residents from and to that region in the long-run.

References

Acemoglu, Daron, Tarek A Hassan, and James A Robinson. 2011. "Social structure and development: A legacy of the Holocaust in Russia." *The Quarterly Journal of Economics* 126 (2): 895–946.

Akbulut-Yuksel, Mevlude, and Mutlu Yuksel. 2015. "The long-term direct and external effects of Jewish Expulsions in Nazi Germany." *American Economic Journal: Economic Policy* 7 (3): 58–85.

Alesina, Alberto, and Paola Giuliano. 2015. "Culture and institutions." *Journal of Economic Literature* 53 (4): 898–944.

Alexander, Manfred. 2008. *Kleine Geschichte der böhmischen Länder.* Reclam.

Becker, Sascha O, Katrin Boeckh, Christa Hainz, and Ludger Woessmann. 2015. "The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy." *The Economic Journal* 126 (590): 40–74.

Becker, Sascha O, and Andreas Ferrara. 2019. "Consequences of forced migration: A survey of recent findings." *Labour Economics* 59:1–16.

Becker, Sascha O, Irena Grosfeld, Pauline Grosjean, Nico Voigtlander, and Ekaterina Zhuravskaya. 2020. "Forced migration and human capital: Evidence from post-WWII population transfers." *American Economic Review* 110 (5): 1430–63.

Belot, Michele, and John Ermisch. 2009. "Friendship ties and geographical mobility: evidence from Great Britain." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 172 (2): 427–442.

Bharadwaj, Prashant, Asim Khwaja, and Atif Mian. 2008. "The big march: migratory flows after the partition of India." *Economic and Political Weekly,* 39–49.

Brakman, Steven, Harry Garretsen, and Marc Schramm. 2004. "The strategic bombing of German cities during World War II and its impact on city growth." *Journal of Economic Geography* 4 (2): 201–218.

Bräuninger, Michael, and Andreia Tolciu. 2011. "Should I stay or should I go? Regional mobility and social capital." *Journal of Institutional and Theoretical Economics JITE* 167 (3): 434–444.

Büchel, Konstantin, Maximilian V Ehrlich, Diego Puga, and Elisabet Viladecans-Marsal. 2020. "Calling from the outside: The role of networks in residential mobility." *Journal of urban economics* 119:103277.

Čapka, František, Lubomír Slezák, and Jaroslav Vaculík. 2005. *Nové osídlení pohraničí českých zemí po druhé světové válce.* Akademické nakladatelství CERM.

Charnysh, Volha. 2019. "Diversity, institutions, and economic outcomes: Post-WWII displacement in Poland." *American Political Science Review* 113 (2): 423–441.

Costa, Dora L, Matthew E Kahn, Christopher Roudiez, and Sven Wilson. 2018. "Persistent social networks: Civil war veterans who fought together co-locate in later life." *Regional science and urban economics* 70:289–299.

Cunningham, Scott. 2021. *Causal inference: The mixtape.* Yale University Press.

CZSO. 2014. *Národní struktura obyvatel.* Technical report. https://www.czso.cz/documents/10180/20551765/170223-14.pdf. Prague: Czech Statistical Office.

Daněk, Petr. 1995. "Towards cultural regionalisation of the Czech Lands: Sudeten half a century after the transfer." *Scripta Fac. Brun* 25:41–60.

David, Quentin, Alexandre Janiak, and Etienne Wasmer. 2010. "Local social capital and geographical mobility." *Journal of Urban Economics* 68 (2): 191–204.

Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica* 78 (6): 1863–1903.

Dvořák, Tomáš. 2013. *Vnitřní odsun 1947–1953: Závěrečná fáze očisty pohraničí v politických a společenských souvislostech poválečného Československa.* Brno: Matice Moravská.

Egger, Peter H, and Andrea Lassmann. 2015. "The causal impact of common native language on international trade: Evidence from a spatial regression discontinuity design." *The Economic Journal* 125 (584): 699–745.

Gerlach, David. 2010. "Beyond Expulsion: The Emergence of "Unwanted Elements" in the Postwar Czech Borderlands, 1945–1950." *East European Politics and Societies* 24 (2): 269–293.

———. 2017. *The Economy of Ethnic Cleansing: The Transformation of the German-Czech Borderlands after World War II.* Cambridge University Press.

Glaeser, Edward L, David Laibson, and Bruce Sacerdote. 2002. "An economic approach to social capital." *The economic journal* 112 (483): F437–F458.

Glassheim, Eagle. 2000. "National mythologies and ethnic cleansing: The expulsion of Czechoslovak Germans in 1945." *Central European History* 33 (4): 463–486.

———. 2006. "Ethnic cleansing, communism, and environmental devastation in Czechoslovakia's Borderlands, 1945–1989." *The Journal of Modern History* 78 (1): 65–92.

Greenwood, Michael J., Gary L. Hunt, Dan S. Rickman, and George I. Treyz. 1991. "Migration, Regional Equilibrium, and the Estimation of Compensating Differentials." *The American Economic Review* 81 (5): 1382–1390. ISSN: 00028282. http://www.jstor.org/stable/2006927.

Grosfeld, Irena, Alexander Rodnyansky, and Ekaterina Zhuravskaya. 2013. "Persistent antimarket culture: a legacy of the pale of settlement after the holocaust." *American Economic Journal: Economic Policy* 5 (3): 189–226.

Helliwell, John F, and Robert D Putnam. 1999. *Education and social capital.* Technical report. National Bureau of Economic Research.

Hotchkiss, Julie L, and Anil Rupasingha. 2018. "Individual Social Captial and Migration."

Kan, Kamhon. 2007. "Residential mobility and social capital." *Journal of Urban Economics* 61 (3): 436–457.

Kelly, Morgan. 2019. "The standard errors of persistence."

———. 2020. "Understanding persistence."

Knack, Stephen. 1992. "Civic norms, social sanctions, and voter turnout." *Rationality and Society* 4 (2): 133–156.

Majerová, Věra, et al. 2003. *Český venkov 2003, Situace před vstupem do EU.* Prague: PEF CZU Praha.

———. 2004. *Český venkov 2004, Život mladých a starých lidí.* Prague: PEF CZU Praha.

Matějka, Ondřej. 2008. *Social Capital in Contemporary Sudetenland,* edited by Joseph D Lewandowski and Milan Znoj, 277–293. Newcastle, UK: Cambridge Scholars.

Meixner, Rudolf. 1988. *Geschichte der Sudetendeutschen.* Nürnberg: Helmut Preussler Verlag.

Oto-Peralías, Daniel, and Diego Romero-Ávila. 2017. "Historical frontiers and the rise of inequality: the case of the frontier of Granada." *Journal of the European Economic Association* 15 (1): 54–98.

Pascali, Luigi. 2016. "Banks and development: Jewish communities in the Italian Renaissance and current economic performance." *Review of Economics and Statistics* 98 (1): 140–158.

Putnam, Robert D. 1995. "Tuning in, tuning out: The strange disappearance of social capital in America." *PS: Political science & politics* 28 (4): 664–684.

Radvanovský, Zdeněk. 2001. "Integrationsprobleme bei der Wiederbesiedlung der deutschen Siedlungsgebiete in den böhmischen Ländern nach 1945." In *Vorträge der Tagungen des Collegium Carolinum in Bad Wiessee vom 20. bis 22. November 1992 und 19. bis 21. November 1993,* 143–161. München, November.

Schumann, Abel. 2014. "Persistence of population shocks: Evidence from the occupation of West Germany after World War II." *American Economic Journal: Applied Economics* 6 (3): 189–205.

Školl, Jaroslav. 1983. *Nové osídlení okresu Břeclav po roce 1945.* Mikulov: Okresní archív Břeclav.

Spurný, Matěj. 2011. *Nejsou jako my: Česká společnost a menšiny v pohraničí 1945–1960.* Praha: Antikomplex.

Staněk, Tomáš. 1991. *Odsun Němců z Československa, 1945–1947.* Academia.

Testa, Patrick A. 2020. "The Economic Legacy of Expulsion: Lessons from Postwar Czechoslovakia." Ueaa132, *The Economic Journal* (December). ISSN: 0013-0133. https://doi.org/10.1093/ej/ueaa132. eprint: https://academic.oup.com/ej/advance-article-pdf/doi/10.1093/ej/ueaa132/34702029/ueaa132.pdf. https://doi.org/10.1093/ej/ueaa132.

Urbánní a regionální laboratoř (URRlab) UK. 2015. *Historická data v GIS.* http://www.historickygis.cz/.

Vaněk, Miroslav. 1996. *Nedalo se tady dýchat: Ekologie v českých zemích v letech 1968 až 1989.* Praha: Ústav pro soudobé dějiny AV ČR.

Waldinger, Fabian. 2010. "Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany." *Journal of Political Economy* 118 (4): 787–831.

Waldinger, Fabian. 2011. “Peer effects in science: Evidence from the dismissal of scientists in Nazi Germany.” *The Review of Economic Studies* 79 (2): 838–861.

———. 2016. “Bombs, brains, and science: The role of human and physical capital for the creation of scientific knowledge.” *Review of Economics and Statistics* 98 (5): 811–831.

Wiedemann, Andreas. 2016. *"Pojď s námi budovat pohraničí!": osídlování a proměna obyvatelstva bývalých Sudet 1945–1952.* Prague: Prostor.

Wilson, Margaret FJ, Brian O’Connell, Colin Brown, Janine C Guinan, and Anthony J Grehan. 2007. “Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope.” *Marine Geodesy* 30 (1–2): 3–35.

The Long-term impact of the resettlement of the Sudetenland on residential migration

Appendix

A Appendix: Additional Figures and Tables

Table A.1: Share of employees on working age population (%) by economic sector in 1930

	Judicial districts by share of ethnic Germans (%)			F-test	Difference
	(90, 100]	(10, 90]	[0, 10]	(F-statistic)	(3)−(1)
	(1)	(2)	(3)	(4)	(5)
Auxiliary retail services	0.837	0.560	0.482	1.908	−0.355†
	(0.167)	(0.124)	(0.099)		
Construction	8.311	10.033	9.612	1.306	1.301
	(0.501)	(0.786)	(0.557)		
Crop and animal production	19.335	29.878	40.839	8.494***	21.504***
	(4.636)	(4.791)	(2.758)		
Education	1.452	2.082	1.577	13.654***	0.126
	(0.076)	(0.137)	(0.049)		
Financial and insurance activities	0.639	0.532	0.732	0.275	0.093
	(0.084)	(0.088)	(0.185)		
Forestry, hunting and fishing	1.444	1.608	1.284	0.212	−0.160
	(0.367)	(0.393)	(0.290)		
Gastronomy	2.348	1.591	1.485	8.366***	−0.863***
	(0.152)	(0.126)	(0.128)		
Health services	0.719	0.672	0.809	0.410	0.090
	(0.042)	(0.084)	(0.106)		
Manufacture of clothes	3.531	3.708	4.240	1.227	0.709
	(0.287)	(0.331)	(0.308)		
Manufacture of food products	4.388	4.614	6.076	2.166	1.688
	(0.456)	(0.623)	(0.591)		
Manufacture of footwear	1.819	1.591	2.153	0.496	0.334
	(0.179)	(0.283)	(0.392)		
Manufacture of glass	3.931	2.916	0.427	2.352	−3.504†
	(2.915)	(1.462)	(0.163)		
Manufacture of chemicals	0.571	0.688	0.411	0.246	−0.160
	(0.294)	(0.512)	(0.182)		
Manufacture of leather	0.564	0.622	0.672	0.097	0.108
	(0.090)	(0.169)	(0.157)		
Manufacture of machinery	3.080	2.343	3.355	0.569	0.275
	(0.699)	(0.374)	(0.585)		
Manufacture of metals	0.046	0.128	0.394	0.361	0.348
	(0.036)	(0.111)	(0.314)		
Manufacture of other products	0.233	0.166	0.070	9.355***	−0.163***
	(0.055)	(0.039)	(0.011)		
Number of districts	71	85	174		

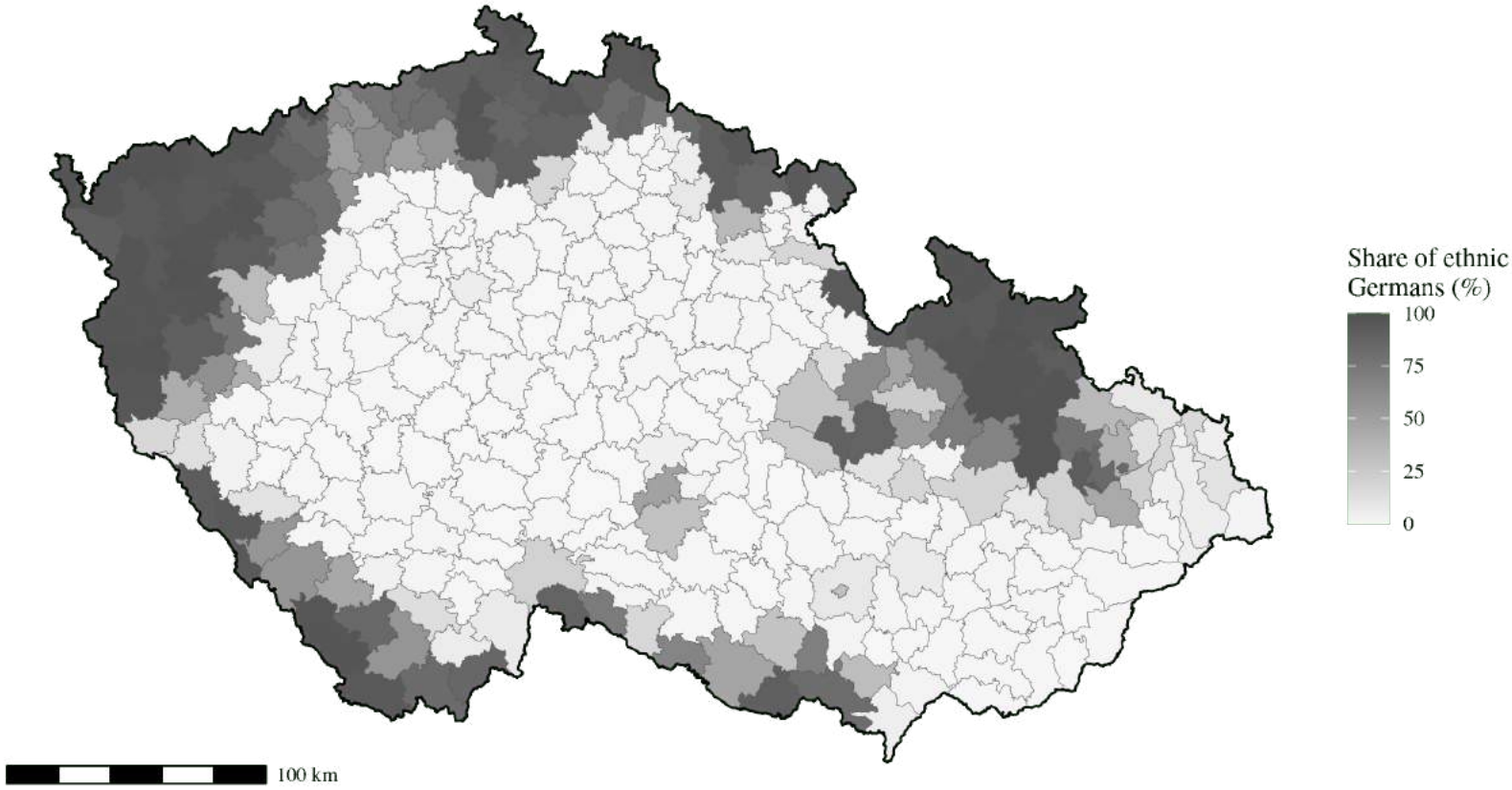
Note: Columns (1) to (3) contain means and standard errors (in parentheses). Column (4) contains F-test for systematic differences across judicial districts groups and Column (5) contains differences in means between districts with less than 10% and more than 90% of ethnic Germans: $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

Table A.2: Share of employees on working age population (%) by economic sector in 1930

	Judicial districts by share of ethnic Germans (%)			F-test	Difference
	(90, 100]	(10, 90]	[0, 10]	(F-statistic)	(3)−(1)
	(1)	(2)	(3)	(4)	(5)
Manufacture of paper	1.076	0.321	0.287	4.262*	−0.789*
	(0.453)	(0.130)	(0.078)		
Manufacture of textiles	35.176	12.910	5.037	10.811***	−30.139***
	(8.931)	(5.523)	(2.193)		
Metals processing	6.720	4.794	5.687	0.607	−1.033
	(1.227)	(0.721)	(0.856)		
Military	0.620	4.313	4.332	1.048	3.712
	(0.303)	(2.067)	(1.670)		
Mining	2.186	5.147	0.297	2.799†	−1.889†
	(1.741)	(3.397)	(0.128)		
Other occupations	0.853	0.681	0.618	0.403	−0.235
	(0.120)	(0.137)	(0.168)		
Other public services	0.430	0.480	0.408	0.312	−0.022
	(0.097)	(0.082)	(0.042)		
Other transport services	2.303	1.396	1.391	4.580*	−0.913*
	(0.454)	(0.147)	(0.124)		
Postal services	0.771	0.741	0.821	0.190	0.050
	(0.092)	(0.099)	(0.082)		
Printing and artistic creation	0.852	0.432	0.458	1.923	−0.394†
	(0.169)	(0.141)	(0.123)		
Production of electricity, gas, water, etc.	0.454	0.248	0.292	1.268	−0.162
	(0.092)	(0.060)	(0.070)		
Public administration	1.372	1.574	1.549	0.170	0.177
	(0.127)	(0.175)	(0.215)		
Quarrying	2.628	3.069	3.226	0.215	0.597
	(0.545)	(0.674)	(0.542)		
Railway transport	2.897	3.191	2.804	0.119	−0.093
	(0.759)	(0.746)	(0.379)		
Timber processing	4.900	4.140	5.018	0.224	0.118
	(0.419)	(0.396)	(0.891)		
Wholesale and retail trade	5.724	5.418	5.462	0.066	−0.261
	(0.463)	(0.615)	(0.465)		
Number of districts	71	85	174		

Note: Columns (1) to (3) contain means and standard errors (in parentheses). Column (4) contains F-test for systematic differences across judicial districts groups and Column (5) contains differences in means between districts with less than 10% and more than 90% of ethnic Germans: $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

Figure A.1: Share of ethnic Germans in judicial districts in 1930



Source: CZSO, Urbánní a regionální laboratoř (URRlab) UK (2015).
Note: The ethnic German population is defined according to the primarily spoken language.

Table A.3: Resettlement process in Břeclav area: Number and share of settlers by the distance of origin

Municipality	Settlers from Czech Republic by the distance of origin				Total	Total
	≤ 5km	≤ 10km	≤ 30km	≤ 100km	Settlers	Population
	(1)	(2)	(3)	(4)	(5)	(6)
Břeclav	195 (4.4%)	875 (19.5%)	1,719 (38.4%)	2,930 (65.5%)	4,476	11,010
Valtice	112 (5.8%)	290 (15.1%)	712 (37.2%)	956 (49.9%)	1,915	3,000
Hustopeče	70 (5.0%)	349 (25.0%)	756 (54.2%)	922 (66.1%)	1,395	2,652
Dolní Dunajovice	0 (0.0%)	0 (0.0%)	302 (25.4%)	654 (55.0%)	1,190	1,689
Lednice	126 (10.7%)	299 (25.4%)	584 (49.5%)	792 (67.2%)	1,179	1,740
Pohořelice	98 (10.0%)	326 (33.4%)	728 (74.6%)	924 (94.7%)	976	2,961
Drnholec	0 (0.0%)	0 (0.0%)	27 (2.9%)	528 (56.4%)	937	1,484
Mikulov	0 (0.0%)	0 (0.0%)	8 (0.9%)	444 (47.4%)	937	5,337
Novosedly	0 (0.0%)	0 (0.0%)	25 (2.7%)	387 (41.3%)	936	1,062
Zaječí	213 (23.3%)	389 (42.6%)	660 (72.3%)	724 (79.3%)	913	1,214
Popice	35 (3.9%)	146 (16.3%)	640 (71.3%)	674 (75.1%)	897	979
Vlasatice	13 (1.6%)	54 (6.5%)	107 (12.8%)	457 (54.6%)	837	988
Pouzdřany	138 (18.7%)	177 (24.0%)	276 (37.4%)	505 (68.4%)	738	1,067
Sedlec	0 (0.0%)	34 (5.2%)	164 (25.3%)	405 (62.5%)	648	803
Strachotín	65 (10.0%)	65 (10.0%)	402 (62.1%)	451 (69.7%)	647	789
Jevišovka	0 (0.0%)	0 (0.0%)	23 (3.7%)	72 (11.5%)	627	808
Perná	0 (0.0%)	0 (0.0%)	34 (5.6%)	111 (18.3%)	608	776
Starovice	0 (0.0%)	327 (55.0%)	453 (76.1%)	468 (78.7%)	595	742
Přítluky	329 (56.9%)	419 (72.5%)	477 (82.5%)	502 (86.9%)	578	745
Bulhary	6 (1.1%)	63 (11.2%)	204 (36.1%)	437 (77.3%)	565	708
Horní Věstonice	0 (0.0%)	0 (0.0%)	161 (31.1%)	348 (67.3%)	517	549
Mušov	27 (5.4%)	33 (6.5%)	129 (25.6%)	356 (70.6%)	504	559
Březí	0 (0.0%)	0 (0.0%)	54 (11.4%)	272 (57.4%)	474	1,208
Pavlov	0 (0.0%)	0 (0.0%)	69 (15.2%)	245 (53.8%)	455	658
Brod nad Dyjí	0 (0.0%)	0 (0.0%)	103 (23.4%)	275 (62.5%)	440	587
Poštorná	84 (19.2%)	122 (27.9%)	159 (36.3%)	243 (55.5%)	438	3,073
Nový Přerov	5 (1.2%)	5 (1.2%)	18 (4.3%)	23 (5.4%)	423	510
Cvrčovice	42 (11.1%)	42 (11.1%)	97 (25.7%)	138 (36.5%)	378	591
Úvaly	0 (0.0%)	0 (0.0%)	89 (25.6%)	207 (59.7%)	347	n.a.
Milovice	16 (4.7%)	21 (6.2%)	114 (33.7%)	237 (70.1%)	338	425
Klentnice	0 (0.0%)	0 (0.0%)	54 (16.4%)	245 (74.5%)	329	373
Bavory	0 (0.0%)	0 (0.0%)	31 (10.7%)	145 (49.8%)	291	342
Dolní Věstonice	0 (0.0%)	0 (0.0%)	99 (40.4%)	149 (60.8%)	245	412
Dobré Pole	0 (0.0%)	0 (0.0%)	36 (15.7%)	73 (31.9%)	229	460
Charvatská Nová Ves	76 (34.2%)	132 (59.5%)	144 (64.9%)	153 (68.9%)	222	1,725
Nová Ves	120 (55.6%)	153 (70.8%)	216 (100.0%)	279 (129.2%)	216	269
Smolín	5 (2.4%)	19 (9.3%)	31 (15.1%)	102 (49.8%)	205	304
Nejdek	8 (4.5%)	37 (20.7%)	121 (67.6%)	136 (76.0%)	179	212
Pasohlávky	0 (0.0%)	0 (0.0%)	31 (21.8%)	54 (38.0%)	142	282
Nové Mlýny	4 (4.3%)	8 (8.6%)	57 (61.3%)	57 (61.3%)	93	143
Hlohovec	3 (5.6%)	19 (35.2%)	21 (38.9%)	25 (46.3%)	54	1,307
Kurdějov	15 (60.0%)	16 (64.0%)	21 (84.0%)	21 (84.0%)	25	524
Total	1,805 (6.4%)	4,420 (15.7%)	10,156 (36.1%)	17,126 (60.9%)	28,138	55,067

Source: Školl (1983), own elaboration.

Table A.4: Resettlement process in Břeclav area: Number and share of settlers from 4 most important municipalities of origin within Czech Republic

Municipality	Settlers from n most important origins				Total	Total
	$n = 1$	$n = 2$	$n = 3$	$n = 4$	Settlers	Population
	(1)	(2)	(3)	(4)	(5)	(6)
Břeclav	416 (9.3%)	677 (15.1%)	935 (20.9%)	1,091 (24.4%)	4,476	11,010
Valtice	125 (6.5%)	224 (11.7%)	301 (15.7%)	367 (19.2%)	1,915	3,000
Hustopeče	85 (6.1%)	159 (11.4%)	229 (16.4%)	294 (21.1%)	1,395	2,652
Dolní Dunajovice	122 (10.3%)	203 (17.0%)	249 (20.9%)	291 (24.4%)	1,190	1,689
Lednice	70 (5.9%)	120 (10.2%)	170 (14.4%)	218 (18.5%)	1,179	1,740
Pohořelice	131 (13.4%)	262 (26.8%)	296 (30.3%)	330 (33.8%)	976	2,961
Drnholec	99 (10.6%)	195 (20.8%)	237 (25.3%)	272 (29.0%)	937	1,484
Mikulov	54 (5.7%)	92 (9.8%)	130 (13.9%)	161 (17.2%)	937	5,337
Novosedly	88 (9.4%)	137 (14.6%)	177 (18.9%)	216 (23.1%)	936	1,062
Zaječí	161 (17.6%)	250 (27.4%)	302 (33.1%)	350 (38.3%)	913	1,214
Popice	123 (13.7%)	184 (20.5%)	239 (26.6%)	293 (32.7%)	897	979
Vlasatice	48 (5.7%)	87 (10.4%)	119 (14.2%)	150 (17.9%)	837	988
Pouzdřany	95 (12.9%)	128 (17.3%)	161 (21.8%)	191 (25.9%)	738	1,067
Sedlec	51 (7.9%)	91 (14.0%)	120 (18.5%)	144 (22.2%)	648	803
Strachotín	100 (15.5%)	177 (27.4%)	242 (37.4%)	277 (42.8%)	647	789
Jevišovka	11 (1.8%)	21 (3.3%)	29 (4.6%)	37 (5.9%)	627	808
Perná	176 (28.9%)	210 (34.6%)	233 (38.4%)	256 (42.2%)	608	776
Starovice	74 (12.4%)	146 (24.5%)	208 (35.0%)	269 (45.2%)	595	742
Přítluky	329 (56.9%)	370 (64.0%)	393 (68.0%)	413 (71.5%)	578	745
Bulhary	41 (7.3%)	82 (14.5%)	117 (20.7%)	152 (26.9%)	565	708
Horní Věstonice	107 (20.7%)	207 (40.0%)	241 (46.6%)	268 (51.8%)	517	549
Mušov	38 (7.5%)	69 (13.7%)	92 (18.3%)	114 (22.6%)	504	559
Březí	34 (7.3%)	69 (14.5%)	99 (21.0%)	130 (27.4%)	474	1,208
Pavlov	57 (12.6%)	99 (21.9%)	119 (26.1%)	134 (29.4%)	455	658
Brod nad Dyjí	46 (10.4%)	88 (20.0%)	130 (29.6%)	157 (35.7%)	440	587
Poštorná	51 (11.6%)	84 (19.2%)	109 (24.9%)	128 (29.2%)	438	3,073
Nový Přerov	21 (5.0%)	26 (6.1%)	31 (7.3%)	36 (8.5%)	423	510
Cvrčovice	49 (13.0%)	72 (19.0%)	86 (22.8%)	95 (25.1%)	378	591
Úvaly	29 (8.4%)	54 (15.6%)	78 (22.5%)	98 (28.2%)	347	n.a.
Milovice	33 (9.8%)	55 (16.3%)	75 (22.2%)	93 (27.5%)	338	425
Klentnice	69 (20.9%)	103 (31.4%)	134 (40.7%)	165 (50.0%)	329	373
Bavory	73 (25.0%)	107 (36.8%)	134 (46.0%)	153 (52.6%)	291	342
Dolní Věstonice	42 (17.2%)	77 (31.2%)	92 (37.5%)	103 (42.2%)	245	412
Dobré Pole	40 (17.5%)	65 (28.4%)	77 (33.6%)	88 (38.4%)	229	460
Charvatská Nová Ves	55 (24.8%)	82 (36.9%)	103 (46.4%)	122 (55.0%)	222	1,725
Nová Ves	17 (7.9%)	34 (15.7%)	51 (23.6%)	67 (31.0%)	216	269
Smolín	28 (13.7%)	40 (19.5%)	50 (24.4%)	57 (27.8%)	205	304
Nejdek	33 (18.4%)	53 (29.6%)	68 (38.0%)	82 (45.8%)	179	212
Pasohlávky	19 (13.5%)	34 (24.3%)	46 (32.3%)	54 (37.7%)	142	282
Nové Mlýny	29 (31.2%)	37 (39.8%)	44 (47.3%)	49 (52.7%)	93	143
Hlohovec	10 (18.5%)	14 (25.9%)	17 (31.5%)	20 (37.0%)	54	1,307
Kurdějov	8 (32.0%)	16 (62.6%)	20 (78.6%)	24 (94.6%)	25	524
Total	3,288 (11.7%)	5,300 (18.8%)	6,783 (24.1%)	8,007 (28.5%)	28,138	55,067

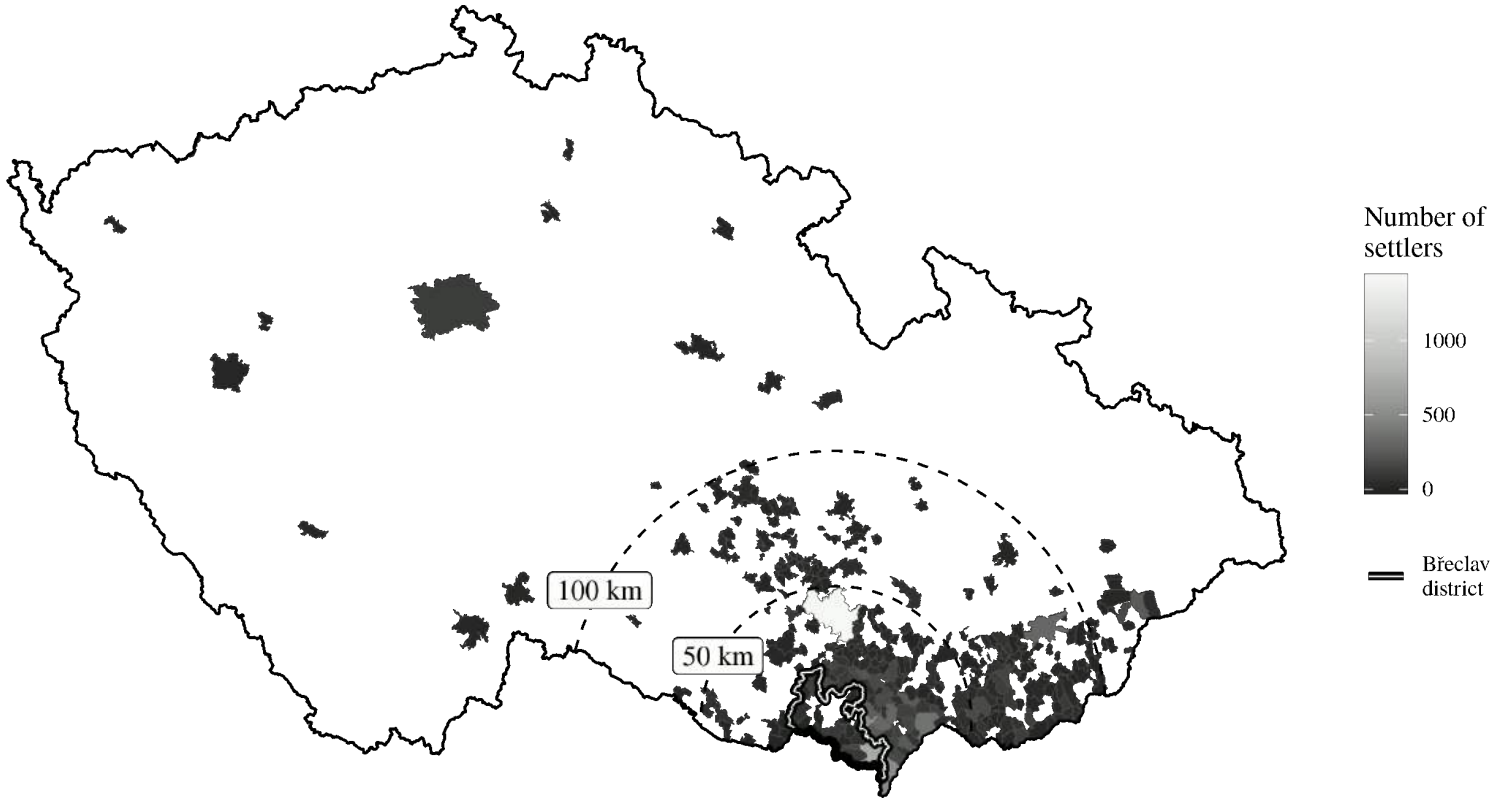
Source: Školl (1983), own elaboration.

Table A.5: Resettlement process in Břeclav area: Number and share of settlers by country of origin

Municipality	Czech Republic	Bulgaria	Slovakia	Soviet Union	Yugoslavia	Austria	Other	Total Settlers	Total Population
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Břeclav	4,181 (93.4%)	3 (0.1%)	0 (0.0%)	17 (0.4%)	7 (0.2%)	62 (1.4%)	27 (0.6%)	4,476	11,010
Valtice	1,486 (77.6%)	7 (0.4%)	88 (4.6%)	3 (0.2%)	17 (0.9%)	55 (2.9%)	39 (2.0%)	1,915	3,000
Hustopeče	1,323 (94.8%)	1 (0.1%)	4 (0.3%)	10 (0.7%)	6 (0.4%)	15 (1.1%)	8 (0.6%)	1,395	2,652
Dolní Dunajovice	1,102 (92.6%)	0 (0.0%)	8 (0.7%)	13 (1.1%)	5 (0.4%)	2 (0.2%)	1 (0.1%)	1,190	1,689
Lednice	992 (84.1%)	7 (0.6%)	27 (2.3%)	0 (0.0%)	50 (4.2%)	25 (2.1%)	9 (0.8%)	1,179	1,740
Pohořelice	809 (82.9%)	4 (0.4%)	48 (4.9%)	61 (6.2%)	3 (0.3%)	12 (1.2%)	11 (1.1%)	976	2,961
Drnholec	815 (87.0%)	11 (1.2%)	77 (8.2%)	0 (0.0%)	0 (0.0%)	31 (3.3%)	4 (0.4%)	937	1,484
Mikulov	796 (85.0%)	2 (0.2%)	19 (2.0%)	2 (0.2%)	20 (2.1%)	1 (0.1%)	7 (0.7%)	937	5,337
Novosedly	785 (83.9%)	81 (8.7%)	18 (1.9%)	0 (0.0%)	33 (3.5%)	3 (0.3%)	4 (0.4%)	936	1,062
Zaječí	860 (94.2%)	0 (0.0%)	15 (1.6%)	4 (0.4%)	0 (0.0%)	9 (1.0%)	14 (1.5%)	913	1,214
Popice	873 (97.3%)	0 (0.0%)	0 (0.0%)	19 (2.1%)	0 (0.0%)	2 (0.2%)	2 (0.2%)	897	979
Vlasatice	676 (80.8%)	22 (2.6%)	59 (7.0%)	32 (3.8%)	8 (1.0%)	2 (0.2%)	3 (0.4%)	837	988
Pouzdřany	707 (95.8%)	0 (0.0%)	4 (0.5%)	11 (1.5%)	0 (0.0%)	0 (0.0%)	3 (0.4%)	738	1,067
Sedlec	564 (87.0%)	0 (0.0%)	17 (2.6%)	0 (0.0%)	46 (7.1%)	2 (0.3%)	6 (0.9%)	648	803
Strachotín	588 (90.9%)	0 (0.0%)	12 (1.9%)	47 (7.3%)	0 (0.0%)	1 (0.2%)	0 (0.0%)	647	789
Jevišovka	158 (25.2%)	445 (71.0%)	5 (0.8%)	0 (0.0%)	13 (2.1%)	0 (0.0%)	2 (0.3%)	627	808
Perná	513 (84.4%)	0 (0.0%)	4 (0.7%)	1 (0.2%)	2 (0.3%)	1 (0.2%)	1 (0.2%)	608	776
Starovice	589 (99.0%)	0 (0.0%)	0 (0.0%)	4 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	595	742
Přítluky	561 (97.1%)	0 (0.0%)	6 (1.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.3%)	578	745
Bulhary	544 (96.3%)	0 (0.0%)	5 (0.9%)	1 (0.2%)	0 (0.0%)	3 (0.5%)	3 (0.5%)	565	708
Horní Věstonice	497 (96.1%)	0 (0.0%)	8 (1.5%)	0 (0.0%)	0 (0.0%)	4 (0.8%)	4 (0.8%)	517	549
Mušov	470 (93.3%)	10 (2.0%)	8 (1.6%)	0 (0.0%)	2 (0.4%)	2 (0.4%)	0 (0.0%)	504	559
Březí	398 (84.0%)	0 (0.0%)	4 (0.8%)	1 (0.2%)	18 (3.8%)	0 (0.0%)	0 (0.0%)	474	1,208
Pavlov	398 (87.5%)	0 (0.0%)	15 (3.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (0.9%)	455	658
Brod nad Dyjí	421 (95.7%)	4 (0.9%)	15 (3.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	440	587
Poštorná	400 (91.3%)	0 (0.0%)	6 (1.4%)	4 (0.9%)	0 (0.0%)	6 (1.4%)	1 (0.2%)	438	3,073
Nový Přerov	91 (21.5%)	327 (77.3%)	1 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.2%)	423	510
Cvrčovice	248 (65.6%)	0 (0.0%)	23 (6.1%)	83 (22.0%)	11 (2.9%)	1 (0.3%)	2 (0.5%)	378	591
Úvaly	333 (96.0%)	0 (0.0%)	14 (4.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	347	n.a.
Milovice	335 (99.1%)	0 (0.0%)	3 (0.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	338	425
Klentnice	321 (97.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (0.6%)	0 (0.0%)	0 (0.0%)	329	373
Bavory	272 (93.5%)	0 (0.0%)	0 (0.0%)	11 (3.8%)	8 (2.7%)	0 (0.0%)	0 (0.0%)	291	342
Dolní Věstonice	272 (111.0%)	0 (0.0%)	4 (1.6%)	0 (0.0%)	4 (1.6%)	0 (0.0%)	4 (1.6%)	245	412
Dobré Pole	176 (76.9%)	0 (0.0%)	1 (0.4%)	0 (0.0%)	20 (8.7%)	14 (6.1%)	0 (0.0%)	229	460
Charvatská Nová Ves	215 (96.8%)	0 (0.0%)	4 (1.8%)	0 (0.0%)	0 (0.0%)	3 (1.4%)	0 (0.0%)	222	1,725
Nová Ves	158 (73.1%)	6 (2.8%)	17 (7.9%)	7 (3.2%)	0 (0.0%)	0 (0.0%)	20 (9.3%)	216	269
Smolín	158 (77.1%)	0 (0.0%)	0 (0.0%)	47 (22.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	205	304
Nejdek	161 (89.9%)	0 (0.0%)	7 (3.9%)	4 (2.2%)	4 (2.2%)	0 (0.0%)	3 (1.7%)	179	212
Pasohlávky	103 (72.5%)	0 (0.0%)	27 (19.0%)	0 (0.0%)	0 (0.0%)	11 (7.7%)	0 (0.0%)	142	282
Nové Mlýny	85 (91.4%)	0 (0.0%)	5 (5.4%)	0 (0.0%)	0 (0.0%)	1 (1.1%)	2 (2.2%)	93	143
Hlohovec	43 (79.6%)	0 (0.0%)	1 (1.9%)	1 (1.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	54	1,307
Kurdějov	117 (468.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	8 (32.0%)	0 (0.0%)	25	524
Total	24,594 (90.3%)	930 (3.4%)	579 (2.1%)	383 (1.4%)	279 (1.0%)	276 (1.0%)	187 (0.7%)	28,138	55,067

Source: Školl (1983), own elaboration. Note: Category *"other"* covers settlers from Poland, Germany, Romania, France, Hungary, Argentina, United Kingdom, Italy, Norway, Canada, Israel, Ecuador and United States.

Figure A.2: Origins of settlers in Břeclav area



Source: Školl (1983), own elaboration.

Figure A.3: The ethnic border and district borders

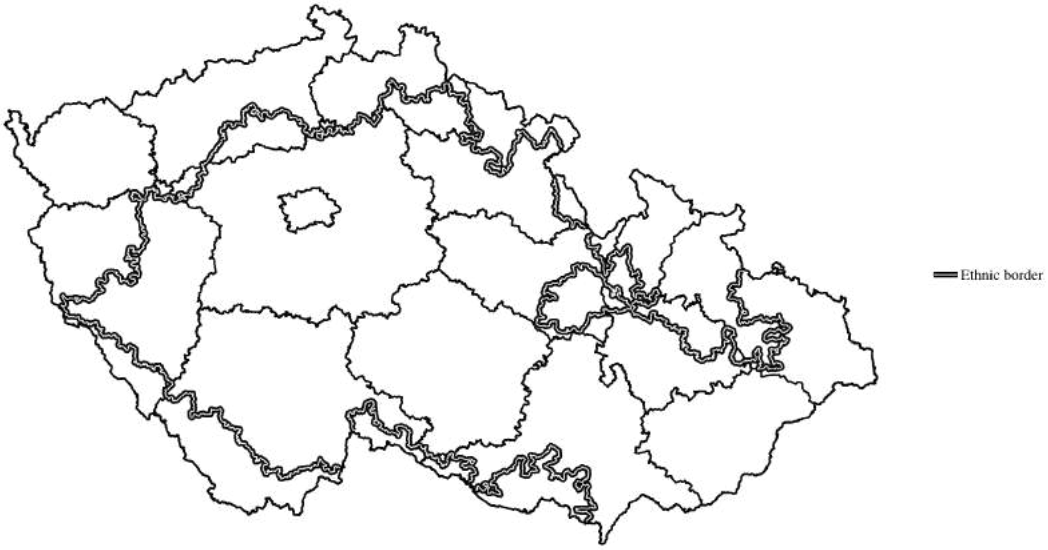


Figure A.4: Migration rates by population

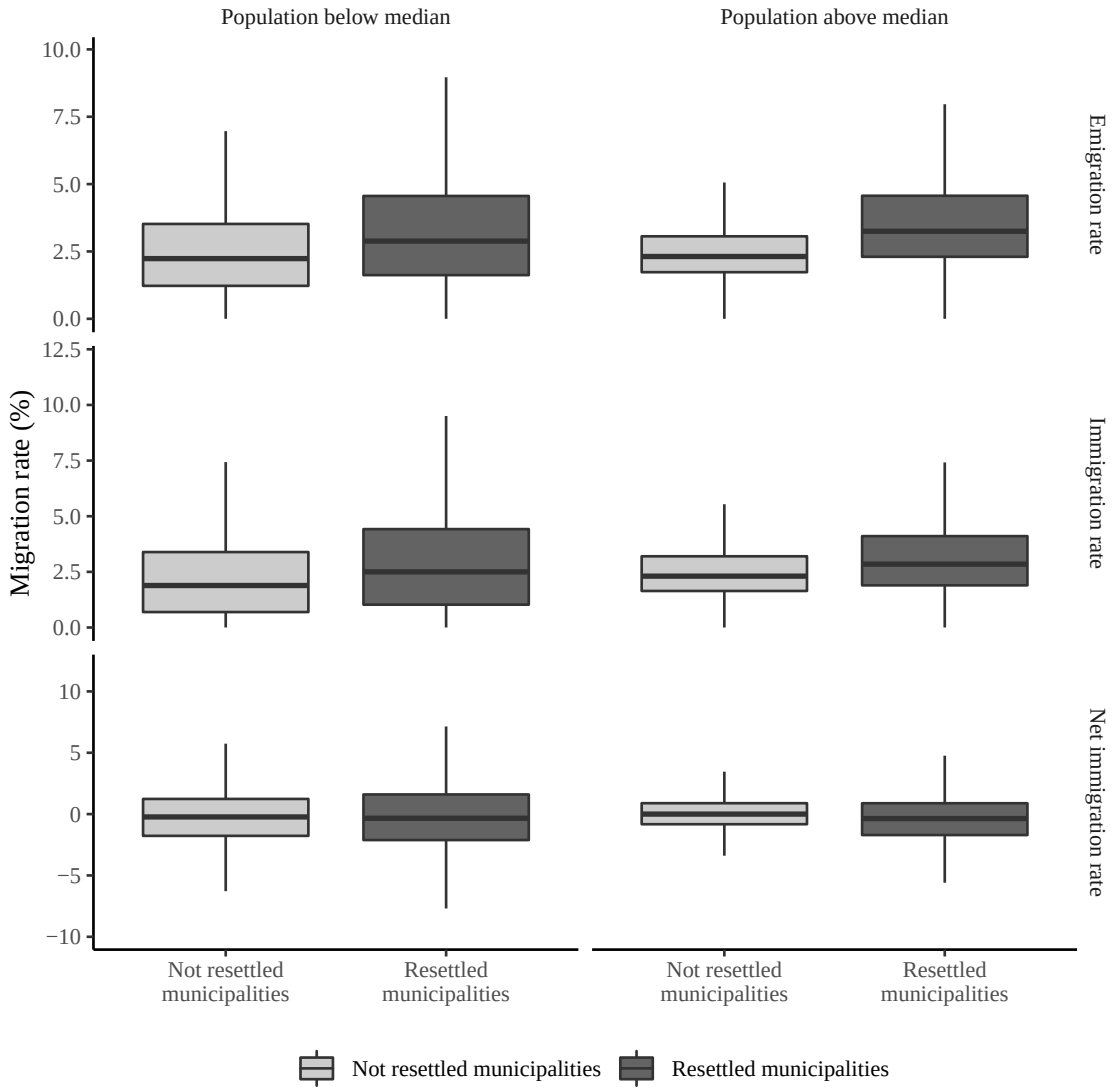


Table A.6: Impact of the resettlement on migration rates (3rd order RD-polynomial)

	Panel estimates	Cross-sectional estimates using data aggregated by period		
		1971–2015	1971–1989	1990–2015
	(1)	(2)	(3)	(4)
	Dependent variable: Emigration rate (%)			
Resettled municipality (=1)	0.570***	0.551***	0.990***	0.290*
	(0.133)	(0.143)	(0.223)	(0.133)
Adjusted R^2	0.169	0.443	0.454	0.359
Observations	82423	2104	1882	2104
	Dependent variable: Immigration rate (%)			
Resettled municipality (=1)	0.405**	0.464**	0.778*	0.271
	(0.148)	(0.167)	(0.311)	(0.176)
Adjusted R^2	0.091	0.319	0.268	0.314
Observations	82423	2104	1882	2104
	Dependent variable: Net immigration rate (%)			
Resettled municipality (=1)	−0.166	−0.087	−0.212	−0.019
	(0.131)	(0.165)	(0.286)	(0.176)
Adjusted R^2	0.096	0.101	0.236	0.124
Observations	82423	2104	1882	2104

Note: Estimates from Equation (1) in the main text with 3rd order RD-polynomial. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, longitude, latitude, and squared longitude and latitude squared. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Table A.7: Descriptive statistics for turnout in municipal elections

Elections	Municipalities			Difference
	All	Resettled	Not resettled	(2)−(3)
	(1)	(2)	(3)	(4)
1990	85.72	78.22	87.59	−9.37***
	(0.14)	(0.36)	(0.15)	
1994	78.18	70.05	80.36	−10.31***
	(0.13)	(0.40)	(0.13)	
1998	67.37	61.13	69.36	−8.23***
	(0.17)	(0.50)	(0.18)	
2002	66.17	62.44	67.59	−5.14***
	(0.16)	(0.46)	(0.17)	
2006	63.27	59.93	64.53	−4.59***
	(0.16)	(0.48)	(0.17)	
2010	63.41	61.34	64.42	−3.08***
	(0.15)	(0.46)	(0.17)	
2014	60.62	58.37	61.65	−3.27***
	(0.16)	(0.47)	(0.17)	

Note: Columns (1) to (3) contain means and standard errors (in parentheses). Columns (4) contains difference in means with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

References

Školl, Jaroslav. 1983. *Nové osídlení okresu Břeclav po roce 1945.* Mikulov: Okresní archív Břeclav.

Urbánní a regionální laboratoř (URRlab) UK. 2015. *Historická data v GIS.* http://www.historickygis.cz/.

B Appendix: Description of data set and merging definitions of communities

B.1 Migration data

Data on permanent residence changes within a year was obtained from the website of the Czech Statistical Office (CZSO, https://www.czso.cz/csu/czso/databaze-demografickych-udaju-za-obce-cr, last accessed April 4th 2019). These data are compiled from administrative historical records on annual residence changes from 1971 to 2015 by the CZSO. Issues related to changing definitions of borders and territories of municipalities are not applicable as CZSO matches them to a consistent data set throughout. The data report the number of people who move in and out of a municipality per year and provide information on the population of a municipality on January 1st of a year. The latter are imputed by CZSO using the last census and annual migration, mortality and natality statistics. The emigration and immigration rates as well as the net immigration rates are used as dependent variables and are defined as the number of residential moves across municipality borders over a year in percent of the population on January 1st of the year.

The data are not a balanced panel because some records were lost prior to their digitization. The number of missing observations is highest in the late 1980s. In this period observations on only 66% of the municipalities are available. Data availability swiftly improves after the fall of the communist regime in 1989. As of the year 1992 96.7% of the data are observed and this ratio has grown ever since. Throughout the main part of the paper we use the maximum number of observations possible and thus focus on unbalanced panel data. We address issues relating to missing observations in additional robustness checks, where we present results based on a balanced panel (see Appendix C).

Furthermore, in processing the data we remove outliers by removing the top 1% of observations with respect to emigration and immigration rates that may be due to large region-specific shocks (such as e.g. the destruction of a number of municipalities due to open pit lignite mining in the Northwest of the Czech Republic – see Vaněk 1996). In our robustness checks in Appendix C we show that this has only minor impacts on the estimated causal effects and if anything increases them. This implies that by omitting outliers, estimates are more conservative.

B.2 Census Data

The key source of historical demographic data of municipalities is the publicly available database *"Historický lexikon obcí"* (available at ast accessed April 4th 2019). The database

is compiled and maintained by the CZSO. It contains municipality level data from general censuses on the total population going back to 1869 using the definition of municipalities in 2011. This data is our source of information on the population and the number of houses in a municipality in 1930.

Unfortunately, data on the ethnic Germans in a municipality, which was contained in the 1930 census and is based on a question on the language spoken in the household was, not digitized. We therefore inputted this data from the printed statistical overviews of the 1930 census published in the 1930s for the 11,797 municipalities reported in the 1930 census. These municipalities do not accord with the definition of the 6,258 municipalities in 2011, that are the observational units for all other data. To derive consistent data we therefore applied the same aggregation rules as used by the CZSO in the construction of time series for the *"Historický lexikon obcí"* and aggregated pre-WWII data to the 2011 municipalities.[32] Using these aggregation rules we succeeded in matching 6168 municipalities (98.6%) with their 1930 counterparts. 4,210 of these municipalities were matched 1 to 1, and 865 municipalities were matched 2 to 1. Only 307 current municipalities consist of more than five 1930 municipalities with the maximum number of municipalities being found in today's Prague, the country capital, which consisted of 60 historical municipalities in 1930.

In our robustness checks we also use data from the decennial population censuses held since 1980. These are the only source on the age and highest completed education of the population of the municipalities. The post-WWII general censuses were held in 1950, 1961, 1970, 1980, 1991, 2001, and 2011. Data on the age and education structure at municipality level are available only for 1980 onward because non-digitized records from the previous censuses were lost to a flood in 2002. We therefore obtained the 2011 census data at the CZSO website (https://www.czso.cz/csu/czso/otevrena_data_pro_vysledky_scitani_lidu_domu_a_bytu_2011_sldb_2011, last accessed April 4th 2019), while data from the 1980, 1991 and 2001 censuses aggregated on the level of 2011 municipalities were provided to us by the CZSO upon request.

B.3 Data on election turnout

Data on election turnout for all municipal elections from 2006 to 2014, was downloaded from the official website for election results run by CZSO (https://www.volby.cz/opendata/opendata.htm), last accessed April 4th 2019). For earlier elections the data were provided by CZSO upon request. These data contain the number of eligible voters as well as the

32. The necessary aggregation rules were kindly provided by the CZSO. They are based on tracking individual neighborhoods. For example, if a municipality was divided and merged with multiple municipalities the aggregation rules associate the historical records of the municipality with a municipality which absorbed the largest neighborhood. By applying these rules the CZSO can construct long-term time-series on population for current municipalities.

number of ballots cast in all the free municipal elections with voluntary voting held in today's Czech Republic since 1990 for each municipality. They can therefore be used to calculate turnout rates. These are defined as the share of votes cast in percent of the eligible voters for each election. This measure is highly reliable, as voting outside the community in which voters are registered as residents is not possible in municipal elections in the Czech Republic.

B.4 Data on the resettlement of Břeclav area

There is no comprehensive data set available on the resettlement process due to its often chaotic nature. To our knowledge the study by Školl (1983)) on the resettlement of the Břeclav area, South Moravia, is the only data source covering a compact region. This study presents data on the resettlement collected from two principal sources: individual 1950 census questionnaires, and records on the land redistribution. Školl's study contains basic information on origins of settlers. For the settlers from today's Czech Republic he lists individual municipalities of the origin while for settlers from Slovakia and foreign countries the study contains only districts or country of origin. The presented data sometimes report only information on the number of families that moved and sometimes only the number of persons, while in some instances (12% of all cases) it contains both families and persons. We use these overlaps to impute the number of persons moving based on the average family size (3.8 individuals) to estimate the number of moving persons for all records where this information is missing in the original data.

B.5 Survey data on social capital

The paper also uses two surveys on the social values and social capital in the Czech Republic. While several surveys in the Czech Republic include questions on local social capital and various values, to the best of our knowledge, only these two provide geolocated respondents at a municipality level and thus allow to determine whether respondents lived in resettled or not resettled municipalities.

B.5.1 Survey *"Český venkov 2004: život mladých a starých lidí"*

Data on social values is taken from the survey *"Český venkov 2004: život mladých a starých lidí"* (Majerová et al. 2004), which was conducted among young (aged 18–29) and old (aged 60–74) respondents in small municipalities (with less than 2000 inhabitants) in 2004 using a standardized questionnaire. Data are available from the Czech Social Science Data Archive (http://archiv.soc.cas.cz/en, last accessed April 4th 2019). They contain observations on 751 young respondents and on 767 old respondents.

We exclude respondents living in municipalities with a population in excess of 2,000 inhabitants according to the 2001 census (12 young respondents and 7 old respondents) and with a self-reported age outside the range declared by the survey definition (4 young respondents older than 29 and 17 old respondents younger than 60) from the sample. This leaves us with 735 young and 738 old respondents. Young respondents are drawn from 275 and old respondents from 278 different municipalities. The geographical distribution of the municipalities and municipality types covered are reported in Figure B.1.

The questions we use asked respondents on their on their values. These were elicited through the battery of questions listed in Table B.1. As respondents evaluated all listed values independently the answers tend to be strongly skewed towards positive evaluations (i.e. *"Very important"* and *"Rather important"*) with *"Very important"* being chosen in 51% cases by young and in 52% by the old group (see Figure B.2). The importance of spiritual values is the only significant exception from the pattern as only 7% of young respondents consider them being *"Very important"* (see Table 6). Therefore, in the main part of the paper these variables are encoded as an indicator variable that takes on a value of 1 if a respondent answered a question with *"Very important"* and zero else. We, however, conducted robustness checks with linear regressions using the full set of values for responses. These robustness checks suggest that our choice of coding has few consequences for the qualitative results (see Appendix C below).

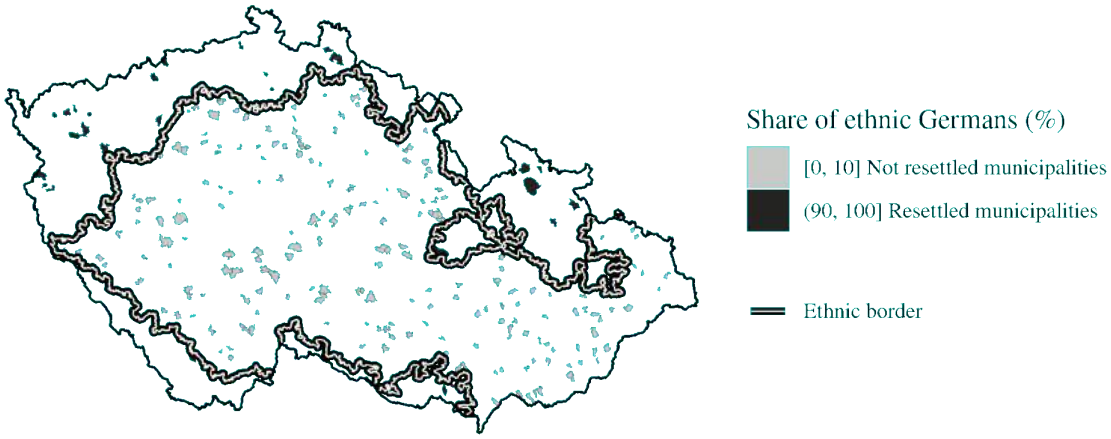
Table B.1: Questions used from survey *"Český venkov 2004: život mladých a starých lidí"*

Wording of the question	
English	Czech (original)
In the following questions, we will focus on your personal values and goals. *People in their lives value some things more and some things less. How important are following things in your life? Please, assign to each of listed values an importance for your life nowadays. Please, try to use the whole scale from 1 to 4 from the CARD while evaluating.*	*V následujících otázkách se zaměříme na Vaše životní hodnoty a cíle.* *Lidé si ve svém životě některých věcí cení více, některých méně. Jak důležité jsou následující věci pro Váš osobní život? Přiřaďte prosím každé z vyjmenovaných hodnot důležitost, kterou jim ve svém současném životě přikládáte. Při hodnocení podle KARTY se, prosím, snažte využívat celou stupnici od jednička do čtyřky.*
CARD:	*KARTA:*
1. *Very important*	1. *Velmi důležité*
2. *Rather important*	2. *Spíše důležité*
3. *Rather unimportant*	3. *Spíše nedůležité*
4. *Not important at all*	4. *Zcela nedůležité*
VALUES:	*HODNOTY:*
1. *Nature, environment*	1. *Příroda, životní prostředí*
2. *Job, occupation*	2. *Práce, zaměstnání*
3. *Relationships*	3. *Mezilidské vztahy*
4. *Faith, spiritual values*	4. *Víra, duchovní hodnoty*
5. *Interests, hobbies*	5. *Zájmy, koníčky*
6. *Housing*	6. *Bydlení*
7. *Friendship*	7. *Přátelství, parta*
8. *Health status*	8. *Zdravotní stav*
9. *Family life, children*	9. *Rodinný život, děti*
10. *Material conditions*	10. *Materiální podmínky*

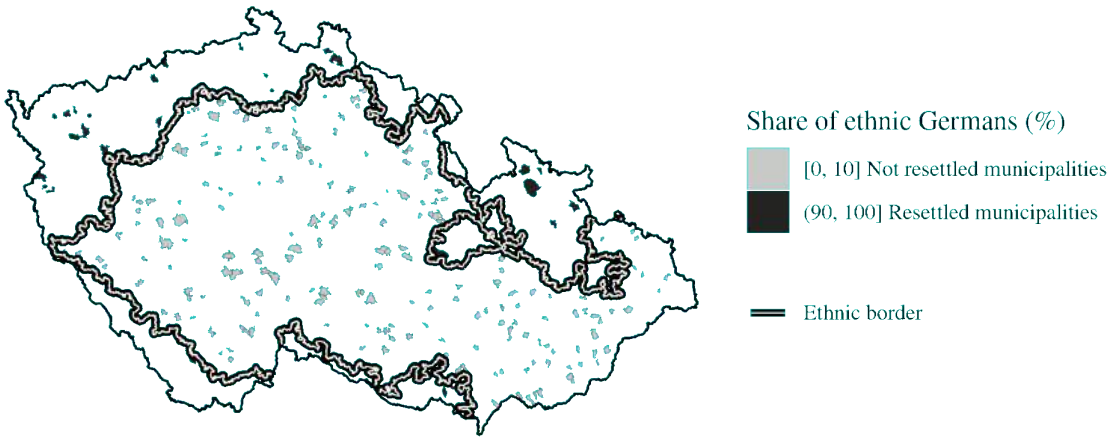
Source: *"Český venkov 2004: život mladých a starých lidí"* (Majerová et al. 2004)

Figure B.1: Location of respondents by municipality type in *"Český venkov 2004: život mladých a starých lidí"*

(a) Young respondents



(b) Old respondents



Source: *"Český venkov 2004: život mladých a starých lidí"* (Majerová et al. 2004), own calculations

B.5.2 Survey *"Český venkov 2003: situace před vstupem do EU"*

The second survey used is the survey *"Český venkov 2003: situace před vstupem do EU"* (Majerová et al. 2003). It too was conducted in municipalities with less than 2,000 inhabitants in 2003. In its first wave this survey asked mayors of municipalities on the number of events organized by local clubs and green activities in their municipality, while in the second wave individuals were asked on whether they are member of a club, participated in an event, have made donation or have committed to conduct voluntary work. We use this survey to evaluate the stock of local social capital in the municipalities of the Czech Republic. Data from both waves of the survey can be downloaded from the Czech Social Science Data Archive (http://archiv.soc.cas.cz/en, last accessed April 4$^{\text{th}}$ 2019).

The first wave was a correspondence expert opinion survey. In this 2,000 mayors of small municipalities (with up to 2000 inhabitants) were asked to assess various characteristics of their municipality. 1,324 (66%) of the mayors participated. From these we exclude 37 municipalities with a population above the declared threshold of 2,000 inhabitants according to the 2001 population census. This leaves us with 1,287 usable observations (see top panel of Figure B.3 for their spatial distribution and municipality types).

We focus on two questions asked in this first wave. The first is the reported number of events organized by local clubs; the second is the number of *"green activities"*. For the number of events, mayors were asked to report the number of events organized by rather broad categories of clubs: Voluntary firefighters, sport clubs, hunters, breeders, Red cross, women union, and others. The survey used the following question:

> *"Označte spolky a zájmová sdružení, které aktivně působí ve Vaší obci a napište počet akcí, které daný spolek pořádal pro veřejnost za poslední rok."*
>
> *"Choose clubs and voluntary organizations, which are active in your village and fill in the number of public events the club organized in the last year."*

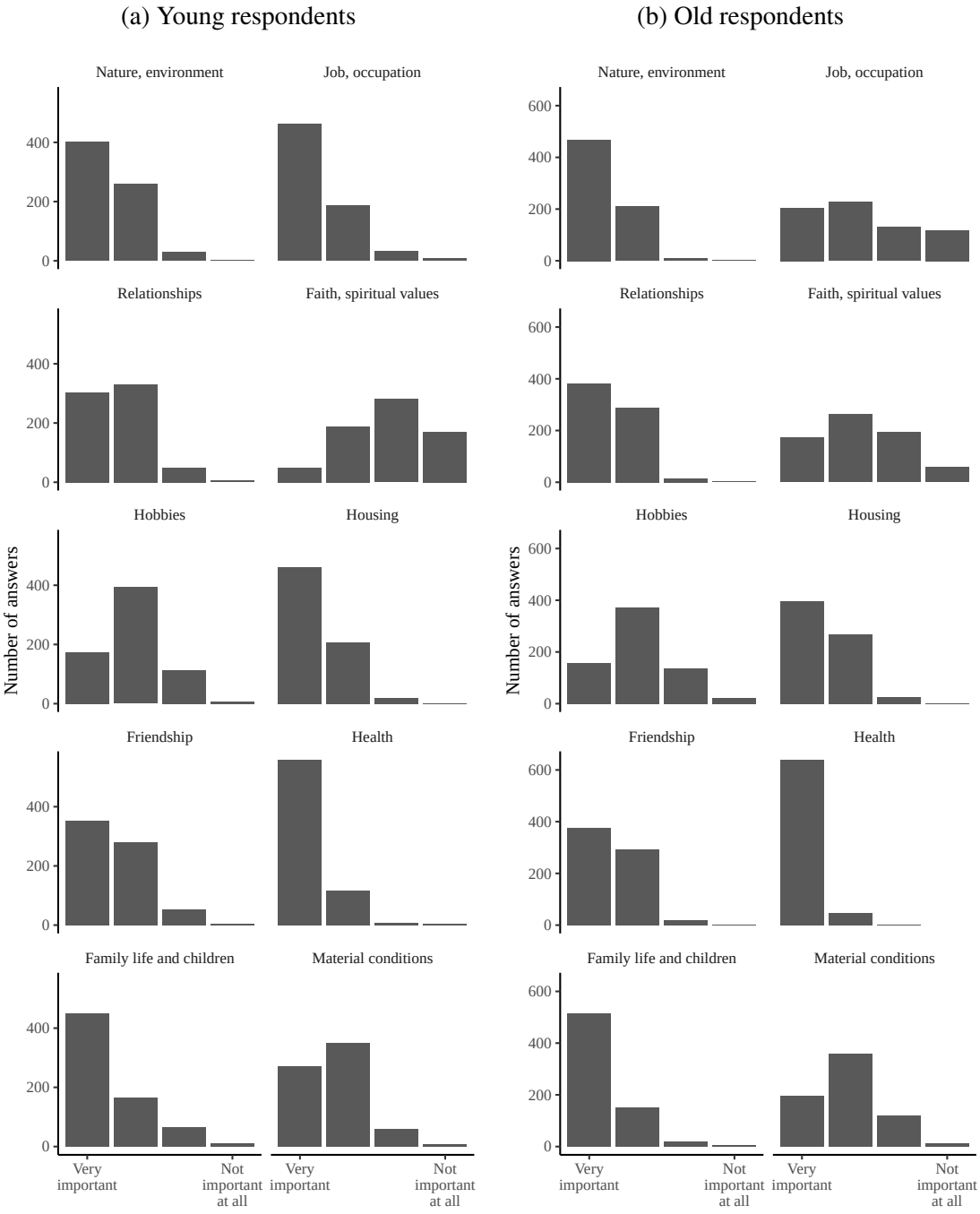
We use the reported number of events as a dependent variable.

For *"green activities"* — i.e. collective activities aimed to improve the environment and living conditions in the municipality (e.g. "spring cleaning", football field reconstruction, etc.) with the participation of locals — mayors were asked:

> *"Pořádají se u Vás akce pro zvelebování obce, kterých se účastní obyvatelé obce?"*
>
> *"Are there green activities in your village, which locals participate in?"*

If respondents answered this question affirmatively they were also required to list the number of events held. We used this number as the dependent variable in the estimates.

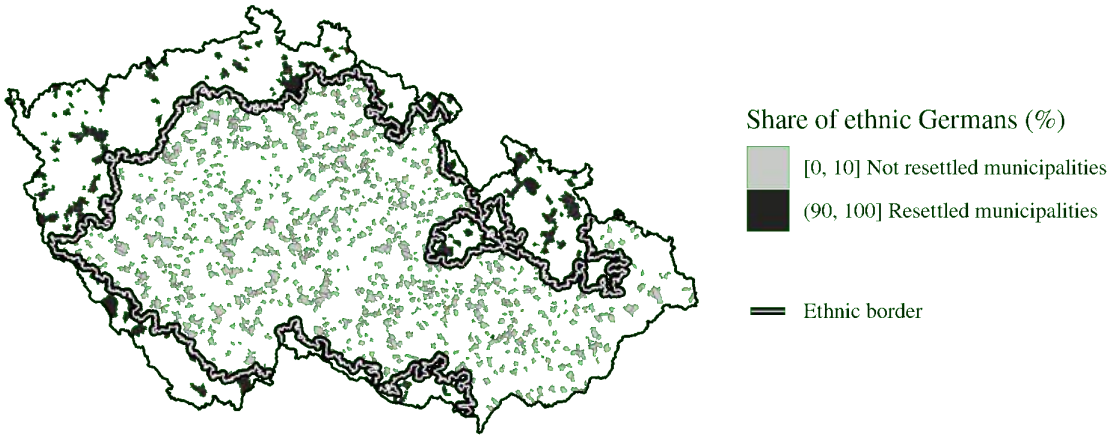
Figure B.2: Distribution of the responses to value questions in *"Český venkov 2004: život mladých a starých lidí"*



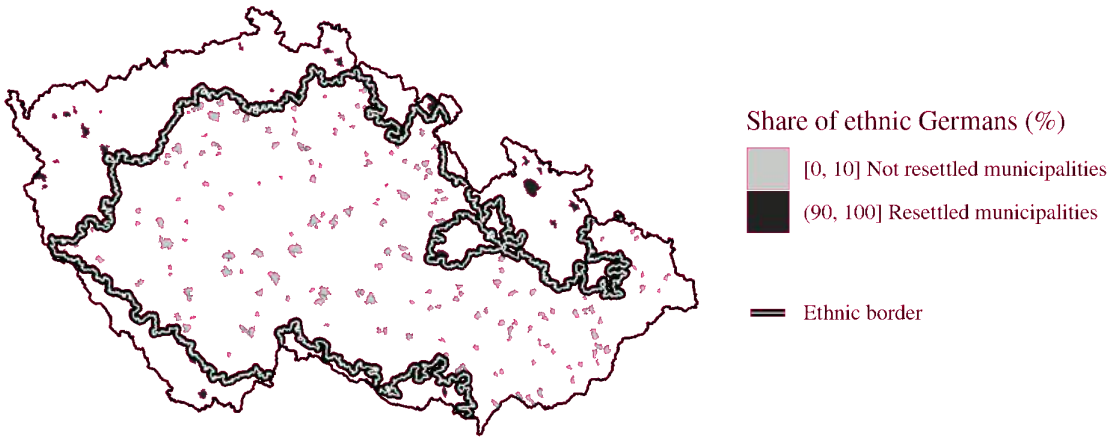
Source: *"Český venkov 2004: život mladých a starých lidí"* (Majerová et al. 2004), own calculations

Figure B.3: Location of respondents by municipality type in *"Český venkov 2003: situace před vstupem do EU"*

(a) First Wave (mayors)



(b) Second Wave (residents)



Source: *"Český venkov 2003: situace před vstupem do EU"* (Majerová et al. 2003), own calculations

The second wave of the survey focused on individuals living in a selection of small municipalities in the Czech Republic. Interviews were conducted among 1,634 respondents. We exclude 27 respondents from municipalities with a population above the declared threshold (of 2,000) in the 2001 census. This leaves us with 1,617 observations living in 223 different municipalities (see bottom panel of Figure B.3 for their spatial distribution and municipalities).

The first question we use, addresses the *"use"* of voluntary organizations and explicitly focuses on the municipality of residence by the respondent and members of his/her household:

> *"Využíváte vy nebo jiní členové domácnosti – spolky, zájmová sdružení v obci?"*
>
> *"Do you, or any member of your household, use clubs or voluntary organizations in your village?"*

The response to this question could be "Yes" or "No". The verb *"využívat"* (to use) in the original Czech wording of the question is not commonly used in this context. We interpret it as a broader term covering not only a membership but also not binding interactions – such as playing football on a club field without being an organized club member. It, therefore, addresses the nature of social capital as it captures interactions with fellow villagers. We encoded positive responses with a value of 1 and negative ones with a value of zero.

The other questions used, by contrast, focus on general charitable and social activities not necessarily related to the municipality of residence. They aim at eliciting information on general activities such as social events attendance, donation on charitable activities and volunteering. In addition, the question on volunteering explicitly asks respondents to take into account activities outside the municipality of residence. The wording of these questions is the following:

> *"Využíváte vy nebo jiní členové domácnosti – společenské akce (disco, plesy, zábavy)?"*
>
> *"Do you, or any member of your household, attend social events (festivals, balls, etc.)?"*
>
> *"Zapojil(a) jste se v roce 2002 jako dobrovolník do nějaké akce (nejen v obci)?"*
>
> *"Did you volunteer for an event in 2002 (not only in your village)?"*
>
> *"Poskytl(a) jste v roce 2002 dar na dobročinné účely?"*
>
> *"Did you make a donation on charitable activities in 2002?"*

The first two questions could be answered by *"Yes"* or *"No"* only and were thus encoded by an indicator variable that takes on the value of 1 for positive and zero for negative responses. The last question could be answered by *"Yes, in kind and money"*, *"Yes, in kind"*, *"Yes in money"* and *"No"*. Here we encoded all positive responses (irrespective of whether they were in kind, in money or both) with one and a negative response with zero.

B.5.3 Representativity of Questionnaires

These questionnaires focus on a rather specific sample of municipalities (municipalities with 2,000 or less inhabitants). This raises obvious concerns about the representativity of this questionnaire. While we have no way to deal with this issue, we checked on the distribution of municipality size in 2001 among our matched municipalities, the municipalitiesin the proximity of the ethnic border, and our overall sample. This check suggests that municipalities with 2,000 inhabitants or less account for 90% of the municipalities in the Czech Republic and host 26% of the population. Similarly, such municipalities account for 98% of the matched municipalities in the Version 3 of the matching estimates and 90% of their population. Furthermore small municipalities account for 92% of the resettled and not resettled municipalities within 15 kilometers band around the ethnic border used in RD-estimates and 45% of the population of these municipalities. Thus, while clearly not representative of the Czech Republic, the Sudetenland or the matched population, the data is at least representative of a large share of the affected area.

References

Majerová, Věra, et al. 2003. *Český venkov 2003, Situace před vstupem do EU.* Prague: PEF CZU Praha.

———. 2004. *Český venkov 2004, Život mladých a starých lidí.* Prague: PEF CZU Praha.

Školl, Jaroslav. 1983. *Nové osídlení okresu Břeclav po roce 1945.* Mikulov: Okresní archív Břeclav.

Vaněk, Miroslav. 1996. *Nedalo se tady dýchat: Ekologie v českých zemích v letech 1968 až 1989.* Praha: Ústav pro soudobé dějiny AV ČR.

C Appendix: Robustness tests

This appendix presents the results of robustness tests presented in Table 5 as well as the additional robustness tests mentioned in the main part of the paper and an additional robustness test for regressions in Table 6. In these robustness tests we

1. Follow a recent suggestion by Kelly (2020) to calculate standard errors corrected for spatial auto-correlation and (as an alternative approach)
2. base inference on permutation tests rather than on clustered standard errors used in the baseline specification.
3. Consider specifications in which different functional forms are allowed for on different sides of the border.
4. Provide estimation results using local linear regression based on several optimal bandwidth selection methods.
5. Use an identification strategy based on spatial matching, that avoids having to define a border between resettled and not resettled municipalities and thus does not require defining this border based on the 50% population share.
6. We conduct the RD-analysis and define resettled municipalities as municipalities with pre-WWII a share of ethnic Germans of 50% or more and all others as municipalities that were not subjected to resettlement.
7. We exclude municipalities close to Bavaria (West Germany) from the estimation sample.
8. We control for potentially divergent regional trends by inclusion of linear trends for each region and period (before/after 1990) into the baseline specification.
9. We exclude the municipalities located in the northeast of the Czech Republic characterized by a large number of ethnic border irregularities.
10. We use a balanced panel dataset.
11. We include outliers omitted in the previous analysis.
12. We use the 1950 population as a control variable.
13. We estimate the baseline specification separately for three parts of the country.
14. We Estimate linear regressions of the social value data with this encoded on a scale from 1 to 4 to ensure that our change of coding has no effects on results

C.1 Robustness test: Standard errors adjusted for spatial auto-correlation

In a recent paper Kelly (2019) provides evidence that RD-design identification estimates based on geographic data may be severely biased in the presence of Spatial auto-correlation. He shows that standard errors could be severely affected by the presence of spatial autocorrelation in cases where the the Moran's Z scores are significant (i.e. exceed 2). In our application (as shown in Table C.1) this is the case for emigration and immigration rates.[33]

Kelly (2020) proposes this concern by explicitly considering spatial auto-correlation in the estimation of such RD-models. We therefore estimate standard errors using his suggested method using his R-code.[34]

In doing so, we closely follow the suggestions made by Kelly (2020): We iterate the procedure for ranges between 10 and 260 kilometers (10, 60, 110, 160, 210, 260) and smoothness parameters between 0.1 and 1 (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1). To stay on the conservative side we report the largest standard errors estimated in this parameter space.

Table C.1 shows that resulting standard errors are virtually identical for emigration rate and only marginally larger for immigration (by 0.006–0.013) and net immigration (by 0.006–0.016) rates.

C.2 Robustness test: Permutation test

Additionally, we test the significance of our point-estimates in a permutation test as an alternative way to treat potential spatial auto-correlation. In this test we obtain an empirical distribution of placebo estimates by artificially shifting ethnic border and use this distribution to test the statistical significance of baseline point-estimates at the actual border. Such a procedure should account for the spatial autocorrelation structure if this is identical at ethnic border and the nearby regions used in the permutation test.

To obtain the placebo estimates we use a sample of not resettled municipalities located from 5 to 50 kilometers from ethnic border toward inland regions. We then artificially shift ethnic border by approx. 15 meters at a time between 20 and 35 km away from the actual one. This ensures that the minimum band along the placebo border is 15 km - as it is in our estimation sample. Resulting distribution of 1,000 placebo estimates is used for testing. One limitation of this permutation test is that it is based only on placebo shifts away from

33. This table uses a Moran's I that is calculated by taking into account the five closest municipalities (as equally weighted neighbors).

34. For the R code developed by Kelly (2020) see https://github.com/morganwkelly/persistence (last accessed on March 15, 2021).

Table C.1: Standard errors adjusted for spatial autocorrelation

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.735***	0.690***	0.570***
	(0.062)	(0.091)	(0.131)
Moran's Z	3.203	3.238	3.249
Smoothness	0.200	0.200	0.200
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.504***	0.465***	0.405*
	(0.076)	(0.112)	(0.161)
Moran's Z	2.456	2.470	2.480
Smoothness	0.400	0.200	0.200
	Dependent variable: Net Immigration rate (%)		
Resettled municipality (=1)	−0.231***	−0.225*	−0.166
	(0.068)	(0.101)	(0.147)
Moran's Z	0.888	0.889	0.882
Smoothness	0.200	0.200	0.200

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are spatial autocorrelation-adjusted standard errors. Moran's Z scores are calculated by taking into account the five closest municipalities (as equally weighted neighbors).

the border toward the interior of the Czech Republic as the distance between the ethnic and country border is often too narrow to allow for shifting towards the country border.

Results reported in Table C.2 show that in the permutation test the effect on the net immigration rate is significant at 0.1% level for the specification with 2nd order RD-polynomial. Other results are identical to baseline estimates.

Table C.2: Permutation test

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.735***	0.690***	0.570***
	(0.021)	(0.033)	(0.055)
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.504***	0.465***	0.405***
	(0.069)	(0.069)	(0.085)
	Dependent variable: Net Immigration rate (%)		
Resettled municipality (=1)	−0.231***	−0.225***	−0.166
	(0.068)	(0.078)	(0.089)

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Values in brackets are standard deviations of the empirical distributions obtained in permutation tests. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities derived in a permutation test with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$.

C.3 Robustness test: RD-polynomials

The baseline specification (1) uses an RD-polynomial of the same degree on both sides of ethnic border. As a robustness check we relax this restriction and allow the order of an RD-polynomial to differ on either side of the border as such a specifications allows for a higher flexibility in modelling the running variable. Results reported in C.3 are in the line with the baseline results. As in our baseline results the effect varies between 0.6 and 0.7 for the emigration rate, between 0.4 and 0.5 for immigration rate, and it is around 0.2 for the net immigration rate.

Table C.3: RD-polynomials

	RD-polynomial (Not resettled municipalities)		
RD-polynomial	1st order	2nd order	3rd order
(Resettled municipalities)	(1)	(2)	(3)
	Dependent variable: Emigration rate (‰)		
1st order		0.731***	0.720***
		(0.068)	(0.078)
2nd order	0.694***		0.679***
	(0.088)		(0.099)
3rd order	0.585***	0.581***	
	(0.126)	(0.129)	
	Dependent variable: Immigration rate (‰)		
1st order		0.495***	0.528***
		(0.077)	(0.091)
2nd order	0.473***		0.498***
	(0.098)		(0.113)
3rd order	0.380**	0.372**	
	(0.137)	(0.140)	
	Dependent variable: Net Immigration rate (‰)		
1st order		−0.236***	−0.192*
		(0.068)	(0.080)
2nd order	−0.221*		−0.182†
	(0.088)		(0.101)
3rd order	−0.205†	−0.209†	
	(0.122)	(0.124)	

Note: Panel estimates from a specification equivalent to Equation (1) in the main text allowing for different RD-polynomials on each side of the cutoff. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

C.4 Robustness test: Local linear regressions

Further to relax our assumptions on the functional form even more we estimate local linear regressions with the optimum bandwidth selected by Mean Square Error (MSE) and Coverage Error Rate (CER) selectors implemented by Calonico et al. (2020). Selectors are used to choose one bandwidth for both sides of the border or for each side of the border individually.

Estimates for the full sample of observations reported in columns (1) to (3) in Table C.4 diverge from our baseline estimates. However, these results are affected by the definition of ethnic border which follows the administrative borders of municipalities. The municipality reference points, that are used for the calculation of the distance to ethnic border, are only rarely located at administrative borders (see an illustrative example in Figure C.1). As a result there are very few observations in the close proximity of ethnic border (see Figure 6) which makes non-parametric estimates unstable and sensitive to outliers. When we exclude 0.5 km band along the ethnic border, which contains 10 not resettled and 7 resettled municipalities, and adjust distances accordingly, the non-parametric estimates in columns (4) to (6) converge to parametric ones. Similarly doubling the bandwidth in columns (7) to (9), that reduces sensitivity to sparse observations in the proximity of ethnic border, leads to point estimates that are close to parametric ones.

Figure C.1: Calculation of distance to ethnic border (illustrative example)

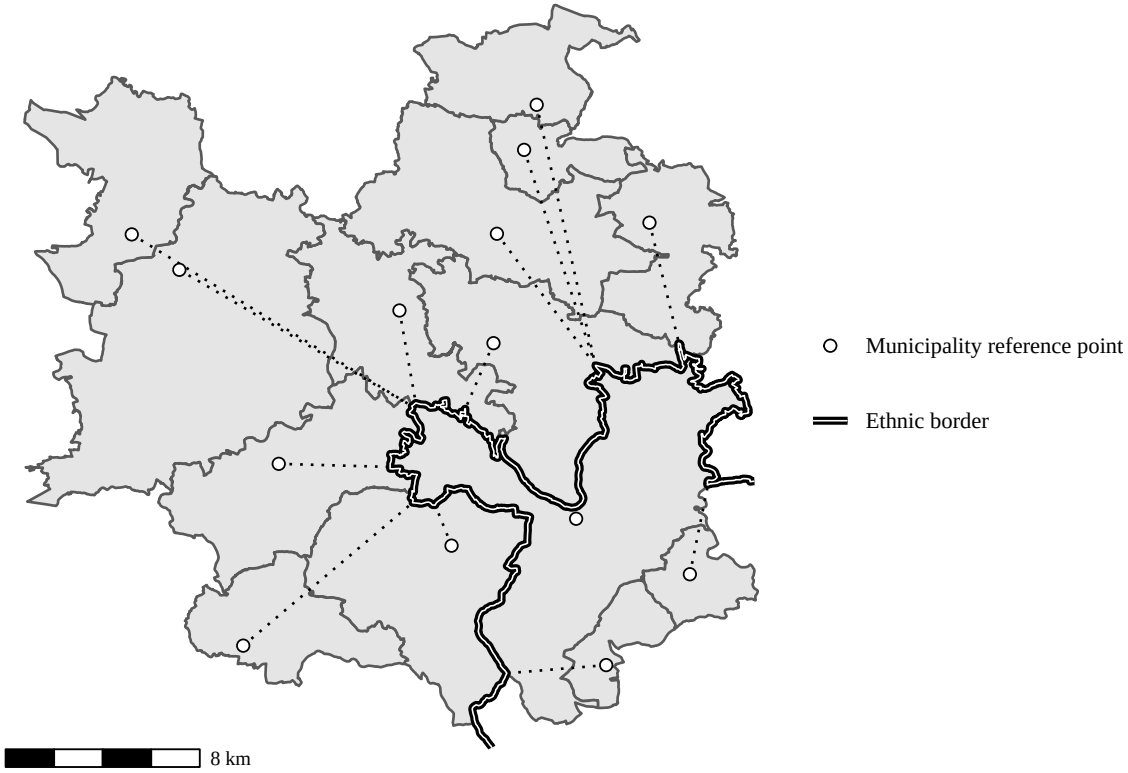


Table C.4: Non-parametric (local linear) estimates

	Full sample, optimum bandwidth			Exclusion of 0.5 km band			Double of optimum bandwidth		
	Dependent variable			Dependent variable			Dependent variable		
	Emigration rate (%)	Immigration rate (%)	Net immigration rate (%)	Emigration rate (%)	Immigration rate (%)	Net immigration rate (%)	Emigration rate (%)	Immigration rate (%)	Net immigration rate (%)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Bandwidth selection: MSE-optimal bandwidth									
Resettled municipality (=1)	0.209†	0.398***	−0.156†	0.778***	0.519***	−0.285***	0.568***	0.464***	−0.182**
	(0.107)	(0.093)	(0.091)	(0.060)	(0.064)	(0.068)	(0.061)	(0.054)	(0.056)
Bandiwdth (km)	1.843	2.765	3.624	2.511	2.855	3.650	3.687	5.529	7.249
Bandwidth selection: MSE-optimal bandwidths (below and above the cutoff)									
Resettled municipality (=1)	0.400***	0.387***	−0.144†	0.823***	0.526***	−0.288***	0.649***	0.478***	−0.184***
	(0.084)	(0.089)	(0.086)	(0.058)	(0.067)	(0.065)	(0.050)	(0.053)	(0.054)
Bandiwdth, Not resettled municipalities (km)	2.525	3.630	3.124	5.532	3.669	3.865	5.050	7.260	6.248
Bandiwdth, Resettled municipalities (km)	2.410	2.753	4.197	2.424	2.532	4.122	4.820	5.506	8.394
Bandwidth selection: CER-optimal bandwidth									
Resettled municipality (=1)	−0.557**	0.471**	0.136	0.789***	0.627***	−0.218*	0.309**	0.389***	−0.194*
	(0.202)	(0.161)	(0.149)	(0.081)	(0.088)	(0.094)	(0.097)	(0.084)	(0.082)
Bandiwdth (km)	1.047	1.570	2.058	1.426	1.622	2.073	2.093	3.139	4.116
Bandwidth selection: CER-optimal bandwidths (below and above the cutoff)									
Resettled municipality (=1)	−0.088	0.427**	0.152	0.795***	0.584***	−0.228*	0.457***	0.411***	−0.185*
	(0.145)	(0.151)	(0.144)	(0.077)	(0.088)	(0.089)	(0.076)	(0.080)	(0.078)
Bandiwdth, Not resettled municipalities (km)	1.434	2.061	1.774	3.142	2.084	2.195	2.867	4.122	3.548
Bandiwdth, Resettled municipalities (km)	1.368	1.563	2.383	1.377	1.438	2.341	2.737	3.127	4.766

Note: Panel estimates from local linear regressions. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, and log of the distance to the country border. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are robust standard errors.

C.5 Robustness test: Spatial matching strategy

The RD-strategy requires us to define the ethnic border, that never existed as a border of, for instance, an administrative unit. Every possible definition of ethnic border is, therefore, to some extent arbitrary and could potentially drive our results. Therefore, we also present estimates based on spatial matching strategy, which does not require any definition of ethnic border.

The matching estimates apply on nearest neighbor matching (with replacement) based on a Euclidean distance measure[35] using population in 1930 and geographic distance as matching variables. Population in 1930 is included as a pretreatment variable to control for important unobservable man-made amenities of agglomerations (e.g., provision of public services, infrastructure, hospitals, cinemas, or sports stadiums and others) that may affect migration. Distance, by contrast, is used to mitigate the effect of unobserved municipality characteristics such as the distance to regional capital, economic conditions, natural and man-made amenities, which also influence migration decisions.

We conduct three versions of the matching procedure. These differ in the maximum distance and population differences allowed for. In a first Version these maxima are 5 kilometers and 500 inhabitants. In a second Version they are 10 kilometers and 250 inhabitants. In the third and least restrictive Version they are 10 kilometers and 500 inhabitants. Depending on these restrictions the matching procedure provides between 46 and 179 matched pairs (Figure C.2).

For each matching Version we estimate the following specification by ordinary least squares:

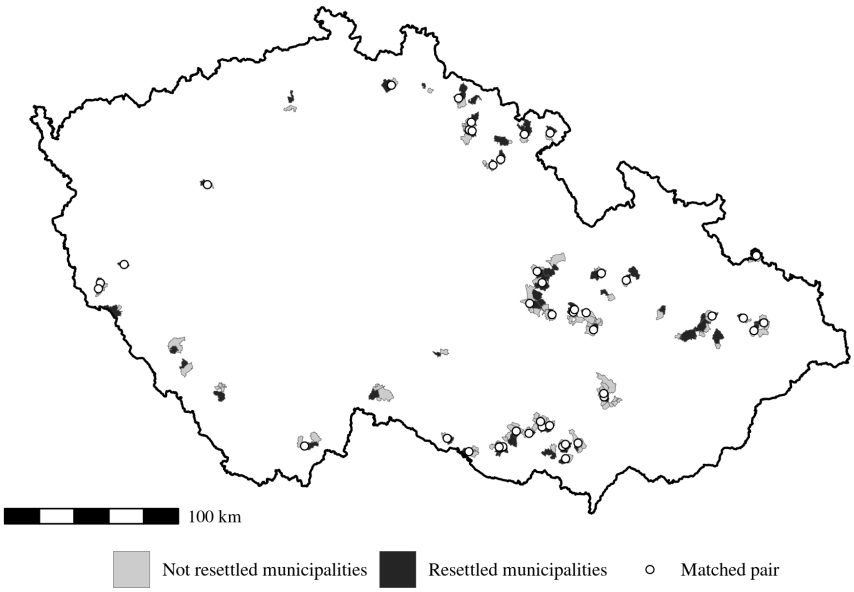
$$y_{it} = \phi_p + \gamma \mathrm{RM}_i + \beta \mathbf{Z}_i + \phi_t + \eta_{it} \tag{2}$$

where y_{it} is one of the three outcome variables (emigration, immigration, and net immigration rates) defined for municipality i in year t, ϕ_p is a match-specific intercept, RM_i is an indicator variable for a resettled municipality, and $\mathbf{Z}_i$ includes geographical variables (log of altitude, the terrain roughness, log of shortest distance to the Czech border, and 2^{nd} order polynomial of longitude and latitude). The year fixed effects ϕ_t are included to control for changes in institutions and economic developments affecting all municipalities alike, and η_{it} is the error term.

35. The standardized Euclidean distance d between two municipalities is calculated as $d_{\mathrm{AB}} = \sqrt{g_{AB}^2 + p_{AB}^2}$, where g is the geographical distance between municipalities adjusted for the earth's curvature, and p is the population difference. Both g and p are normalized variables with zero mean and unit variance across all municipality pairs eligible for matching. A resettled municipality is matched with an not resettled municipality that minimizes d.

Figure C.2: Matched pairs of municipalities by maximum permitted geographical distance and population differences *(First part)*

(a) Version 1



(b) Version 2

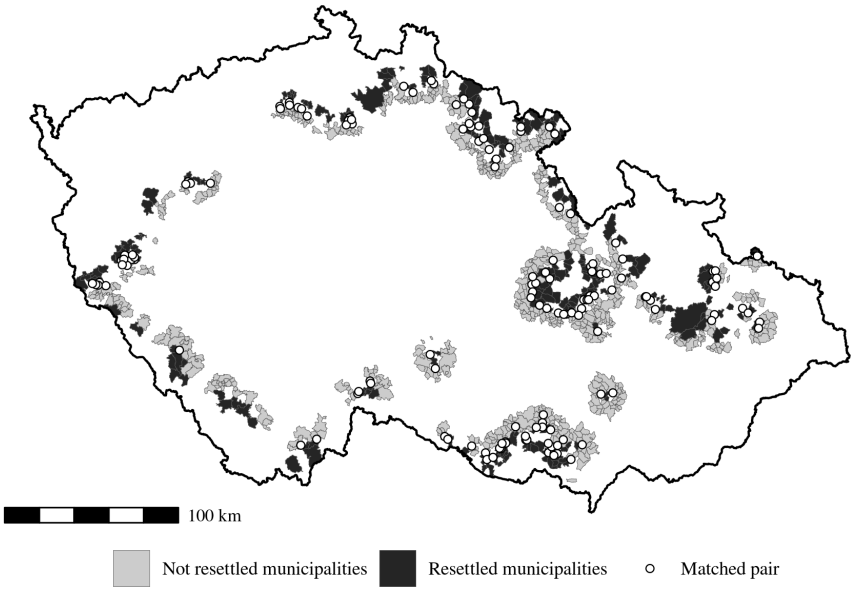
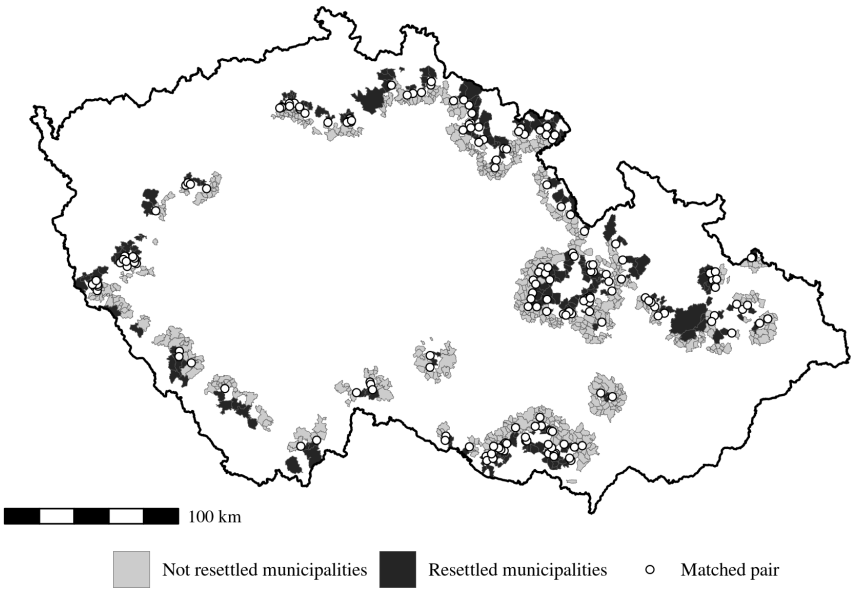


Figure C.3: Matched pairs of municipalities by maximum permitted geographical distance and population differences *(Continued)*

(c) Version 3



Source: CZSO, own calculations.
Note: In Version 1 municipalities are matched only if they are located less than 5 km from each other and their population difference is less than 500 inhabitants. In Versions 2 and 3 these maxima are 10 km and 250 inhabitants, and 10 km and 500 inhabitants, respectively.

Table C.5: Balance tests for matched pairs

	Matching version					
	Version 1		Version 2		Version 3	
	Not resettled municipalities	Resettled municipalities	Not resettled municipalities	Resettled municipalities	Not resettled municipalities	Resettled municipalities
	(1)	(2)	(3)	(4)	(5)	(6)
Altitude (meters)	348.304	353.598	362.827	385.405*	376.866	395.824
	(19.691)	(18.840)	(10.604)	(10.904)	(10.107)	(10.159)
Terrain roughness	2.699	2.574	2.679	2.800	2.798	2.910
	(0.188)	(0.203)	(0.112)	(0.112)	(0.099)	(0.102)
Distance to the country border (km)	27.280	25.901**	27.692	24.968***	26.622	24.005***
	(2.332)	(2.256)	(1.270)	(1.202)	(1.168)	(1.097)
Population[a]	832.718	831.254	821.441	819.232	882.305	921.772
	(9.914)	(12.011)	(7.527)	(7.809)	(7.746)	(8.351)
Maximum geographical distance (meters)	5000		10000		10000	
Maximum difference in population	500		250		500	
Mean geographical distance (meters)	4135.2		8112.6		7909.4	
Mean difference in population	206.4		118.6		214.6	
Number of matched pairs	46		147		179	

Note: In Version 1 municipalities are matched only if they are located less than 5 km from each other and their population difference is less than 500 inhabitants. In Versions 2 and 3 these maxima are 10 km and 250 inhabitants, and 10 km and 500 inhabitants, respectively. Table 2 reports means (and robust standard errors in parentheses) with statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. For geographical variables tests for differences in means control for matched pair fixed effects. Tests for difference in population additionally control for the log of altitude, terrain roughness, and log of distance to the country border. P-values are calculated using robust standard errors.

Table C.6 presents the baseline estimates of the parameter of interest (γ) of Equation (2) from all three versions of matching using standard errors that are clustered by municipality.

Table C.6: Robustness test: Spatial matching strategy

	Matching		
	1st version	2nd version	3rd version
	(1)	(2)	(3)
	Dependent variable: Emigration rate (‰)		
Resettled municipality (=1)	0.620***	0.782***	0.816***
	(0.088)	(0.053)	(0.040)
Observations	3342	10404	12788
Adjusted R2	0.178	0.219	0.219
	Dependent variable: Immigration rate (‰)		
Resettled municipality (=1)	0.516***	0.607***	0.546***
	(0.085)	(0.061)	(0.058)
Observations	3342	10404	12788
Adjusted R2	0.119	0.126	0.136
	Dependent variable: Net immigration rate (‰)		
Resettled municipality (=1)	−0.104*	−0.175**	−0.271***
	(0.051)	(0.053)	(0.055)
Observations	3342	10404	12788
Adjusted R2	0.119	0.136	0.140

Note: Estimates from Equation (2). In Version 1 settlements are matched only if they are located less than 5 km from each other and their population difference is less than 500 inhabitants. In Versions 2 and 3 these maxima are 10 km and 250 inhabitants, and 10 km and 500 inhabitants, respectively. All estimates control for pair fixed effects, year fixed effects, log of altitude, terrain roughness, and log of the distance to the country border. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

These estimates suggest that resettlement increased emigration rates from the resettled municipalities by 0.6 to 0.8 percentage points and immigration rates by approximately 0.5 to 0.6 percentage points on average in the years 1971 to 2015. In addition, according to the results net immigration decreased by 0.1 to 0.3 percentage points per year on average in the studied period. All estimates are significant at least at 5% level and are fully comparable with baseline results obtained with RD-strategy. These results suggest that our baseline results are not driven by a specific definition of ethnic border.

C.6 Robustness test: The 1950 population as a control variable

In the baseline specification we use 1930 population for population fixed effects. The 1930 population is clearly exogenous to the resettlement process, but it may be a bad proxy for the development infrastructure and other man-made amenities as municipality sizes of formerly German-dominated municipalities changed substantially between 1930 and the post-resettlement period. To address this concern we re-estimate the baseline specification (1) using 1950 population for population fixed effects.

In the estimates (Table C.7), coefficients for emigration and immigration as well as net immigration rates change by less than 0.1 percentage points relative to the baseline results reported in the main part of the paper and also maintain their significance levels.

C.7 Robustness test: Focusing on all municipalities

Another concern with the RD-strategy is that we focus only on municipalities which either had a 90% German speaking majority in the German part of the country or a 90% Czech speaking majority in the Czech speaking part of the country. This means we omit 466 (18%) municipalities from our data (see Figure 2 in the main part of the paper). This could imply that estimates may be influenced by a few outliers located directly at the border. To check on this issue, we added another robustness check where we define resettled municipalities as municipalities with a majority (more than 50%) of ethnic Germans. The rest of municipalities as not being resettled. The disadvantage of this approach is that this may lead to a comparison of municipalities at the border, that were only marginally differently affected by resettlement. The advantage is that now virtually all municipalities in the 15 kilometer band around the ethnic border are included in the analysis.

The results of this robustness check (reported in Table C.8) once more indicate a high consistency with the results reported in the main part of the paper. According to the results the impact of resettlement on emigration and immigration rates as well as net immigration rates changes by less than 0.2 throughout. The mild decline in the estimated effect is also consistent with inclusion of less-treated municipalities into the estimation sample.

Table C.7: Robustness test: The 1950 population as a control variable

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.710***	0.688***	0.597***
	(0.062)	(0.089)	(0.129)
Adjusted R^2	0.170	0.170	0.170
Observations	82388	82388	82388
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.536***	0.490***	0.448**
	(0.071)	(0.103)	(0.147)
Adjusted R^2	0.090	0.090	0.090
Observations	82388	82388	82388
	Dependent variable: Net Immigration rate (%)		
Resettled municipality (=1)	−0.173**	−0.198*	−0.150
	(0.062)	(0.089)	(0.125)
Adjusted R^2	0.097	0.097	0.097
Observations	82388	82388	82388

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect based on 1950 population, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Table C.8: Robustness test: Focusing on all municipalities

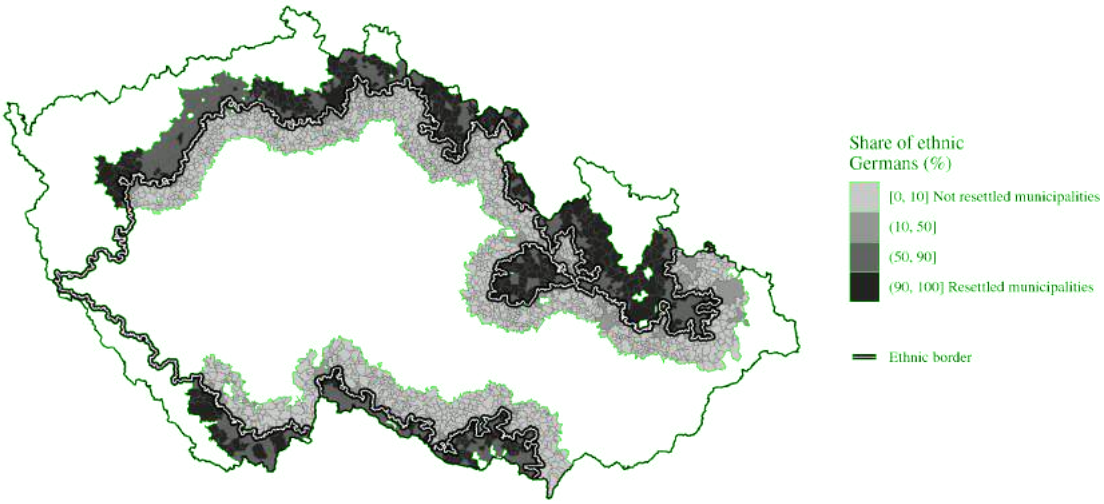
	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.593***	0.514***	0.491***
	(0.039)	(0.057)	(0.081)
Adjusted R^2	0.188	0.188	0.189
Observations	99109	99109	99109
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.350***	0.340***	0.353***
	(0.043)	(0.062)	(0.089)
Adjusted R^2	0.103	0.103	0.103
Observations	99109	99109	99109
	Dependent variable: Net Immigration rate (%)		
Resettled municipality (=1)	−0.242***	−0.175**	−0.138†
	(0.039)	(0.057)	(0.081)
Adjusted R^2	0.100	0.100	0.100
Observations	99109	99109	99109

Note: Panel estimates from Equation (1) in the main text with resettled municipalities being defined as municipalities with German-speaking majority in 1930. All other municipalities are considered as not being resettled. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

C.8 Robustness test: Exclusion of a part of ethnic border close to Bavaria (West Germany)

The major institutional shift related to fall of the Iron curtain and the communist regime in 1989 may have an asymmetric impact on municipalities in the estimation sample as it may asymmetrically increase the attractiveness of the Czech-German border region as a place of residence. In this robustness check we re-estimate regression (1) excluding the municipalities located close to Bavaria (West Germany) – see Figure C.4. The results reported in table C9 are once more highly consistent with those reported in the main text.

Figure C.4: Municipalities within 15 km to the ethnic border after omitting part of ethnic border closest to Bavaria (West Germany)



Results reported in Table C.9 are practically identical to the baseline estimates confirming high robustness of our results.

C.9 Robustness test: Exclusion of a part of ethnic border characterized by a large number of irregularities

One of the issues with using the RD-strategy is that throughout history (with the exception of the time from 1938 to 1945) a well-defined border between the Czech and German parts of today's Czech Republic did not exist and that the ethnic border between the two parts of the country is marked by a number of irregularities such as ethnic enclaves. To provide some evidence of the potential impact of these irregularities, we omitted municipalities

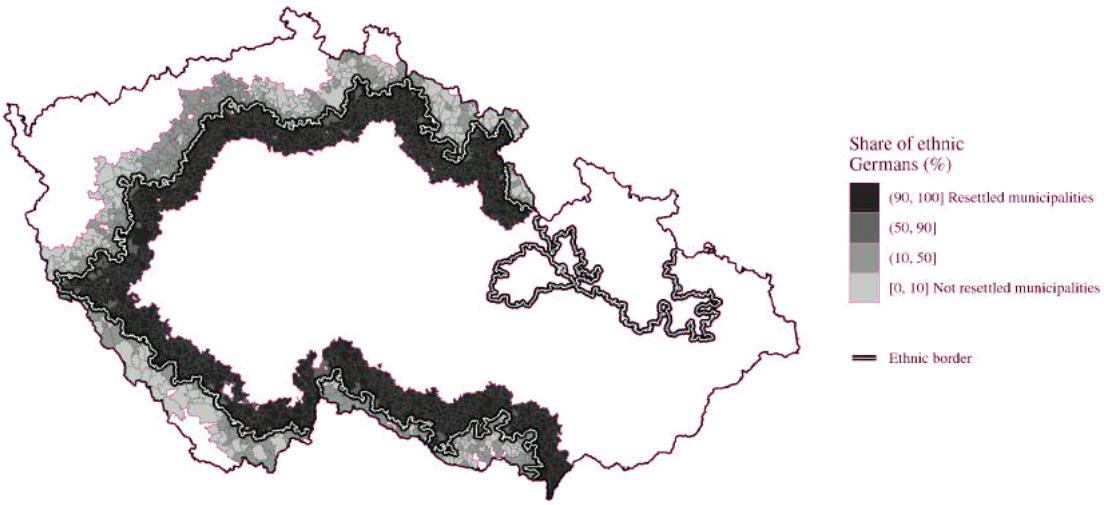
Table C.9: Robustness test: Exclusion of the part of the ethnic border closest to Bavaria (West Germany)

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.659***	0.605***	0.572***
	(0.064)	(0.091)	(0.131)
Adjusted R^2	0.172	0.172	0.172
Observations	72006	72006	72006
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.456***	0.409***	0.400**
	(0.072)	(0.103)	(0.149)
Adjusted R^2	0.098	0.098	0.098
Observations	72006	72006	72006
	Dependent variable: Net immigration rate (%)		
Resettled municipality (=1)	−0.203**	−0.196*	−0.172
	(0.065)	(0.094)	(0.134)
Adjusted R^2	0.101	0.101	0.101
Observations	72006	72006	72006

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Municipalities on the part of the ethnic border closest to Bavaria (West Germany) are excluded from the sample (see Figure C.4). Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

from the Northeast part of the Czech Republic characterized by a very ragged ethnic border.[36]

Figure C.5: Municipalities within 15 km to the ethnic border after omitting Northeast part of the Czech Republic



The exclusion of these municipalities from the sample has virtually no impact on the estimated effects in RD-analysis (see Table C.10). In the RD-regression the estimated impact of resettlement on the emigration rate is now marginally lower at 0.5 to 0.7 percentage points, as in the baseline specification, the estimated impact of the emigration rate also remains unchanged at 0.4 to 0.5 percentage points. The impact on net immigration is, however, insignificant for higher order RD-polynomials. This is in line with other results in the main part of the paper and suggests that results for net immigration rates are somewhat less robust (and limited only to the 1980s) than those pertaining to and net immigration rates.

36. For the sample of municipalities, included in this robustness check, after the exclusion of Northern Moravia see Figure C.5.

Table C.10: Robustness test: Exclusion of a part of ethnic border characterized by a large number of irregularities

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.734***	0.631***	0.546**
	(0.081)	(0.123)	(0.179)
Adjusted R^2	0.154	0.154	0.154
Observations	57763	57763	57763
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.516***	0.440**	0.494*
	(0.094)	(0.140)	(0.204)
Adjusted R^2	0.085	0.085	0.085
Observations	57763	57763	57763
	Dependent variable: Net immigration rate (%)		
Resettled municipality (=1)	−0.218**	−0.191	−0.052
	(0.080)	(0.120)	(0.170)
Adjusted R^2	0.092	0.092	0.092
Observations	57763	57763	57763

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Municipalities on the part of the ethnic border characterized by a large number of irregularities are excluded from the sample (see Figure C.5). Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

C.10 Robustness test: Divergent regional trends

One concern with our baseline specification may be that regions are following different time trends in the outcome variable. These divergent trends may be a results of the new settlers "settling in" to their new region of residence, could be due to longer run trends in economic development over our rather long period of interest (1971–2015) or may result from the many major institutional reforms in the period (such as e.g. transition) that could impact on the attractiveness of certain regions analyzed. Since such could bias our baseline results we also extended specification (1) to include linear time trends each region that are allowed to differ for the period period before and after 1989, to account for the major shift in the institutional environment and correspond to trends observable in outcome variables.

Results reported in Table C.11 are highly consistent with our baseline estimates as they do differ from those by less than 0.01 percentage points.

C.11 Robustness test: Using balanced panel data set

As noted in the main part of the paper and detailed in Appendix B, we use unbalanced panel data on the municipalities to include the maximum number of observations possible. Focusing on this unbalanced data may lead to biased results if there are systematic difference in missing observations of resettled and not resettled municipalities. Since such selectivity cannot be entirely precluded we repeat estimates using a balanced panel of municipalities. This was constructed by using only those municipalities for which annual observations are available for all years from 1971 to 2015.

Table C.12 shows the results for the causal impact of resettlement on residential migration. These results are highly consistent to results reported in the main part of the paper. The estimated treatment effects differ by less than 0.1 percentage points from those in the main part of the paper.

Table C.11: Robustness test: Divergent regional trends

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (‰)		
Resettled municipality (=1)	0.738***	0.687***	0.566***
	(0.063)	(0.092)	(0.133)
Adjusted R^2	0.172	0.172	0.172
Observations	82423	82423	82423
	Dependent variable: Immigration rate (‰)		
Resettled municipality (=1)	0.505***	0.466***	0.410**
	(0.070)	(0.103)	(0.148)
Adjusted R^2	0.096	0.096	0.096
Observations	82423	82423	82423
	Dependent variable: Net immigration rate (‰)		
Resettled municipality (=1)	−0.233***	−0.221*	−0.156
	(0.063)	(0.092)	(0.131)
Adjusted R^2	0.100	0.100	0.100
Observations	82423	82423	82423

Note: Panel estimates from extended Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, 2nd order polynomial of longitude and latitude, and linear time trends for each region and period before/after 1989. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Table C.12: Robustness test: Balanced panel dataset

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (‰)		
Resettled municipality (=1)	0.821***	0.780***	0.667***
	(0.078)	(0.113)	(0.164)
Adjusted R^2	0.214	0.214	0.214
Observations	57974	57974	57974
	Dependent variable: Immigration rate (‰)		
Resettled municipality (=1)	0.467***	0.478***	0.435**
	(0.080)	(0.113)	(0.168)
Adjusted R^2	0.098	0.097	0.098
Observations	57974	57974	57974
	Dependent variable: Net immigration rate (‰)		
Resettled municipality (=1)	−0.354***	−0.302**	−0.232†
	(0.068)	(0.101)	(0.140)
Adjusted R^2	0.101	0.101	0.101
Observations	57974	57974	57974

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border and with full set of observations are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

C.12 Robustness test: Inclusion of outliers

In the main part of the paper we excluded the top 1% of all observations in terms of emigration and immigration rates to avoid using implausible values caused by unobserved shocks.[37] Our motivation for this was that the RD-estimates (due to few observations directly at the border) may be very sensitive to such outliers. We therefore estimated RD-estimates with and without eliminating such outliers.

Results in Table C.13 indicate that depending on the functional form of the RD-polynomial, emigration rates were by 0.5 to 0.9 percentage points and statistically significantly higher in resettled than in not resettled municipalities. This is comparable or only slightly higher than the 0.6 to 0.7 percentage points in the baseline estimates. The results also suggest that immigration rates to the resettled municipalities from other regions were by 0.4 to 0.7 percentage points and statistically significantly higher in all estimates than in not resettled municipalities (rather than 0.4 to 0.5 percentage point in the baseline results). Finally, also the estimated effect of resettlement on net immigration rates are negative but lack statistical significance with the exception of the specification with first order RD-polynomial.

C.13 Robustness test: Impact of the resettlement by part of the country

Since resettled municipalities are located across different parts of the border of today's Czech the estimated overall impact of the resettlement could be driven by some unobserved local-specific characteristics of different border segments. In one robustness test we therefore split our estimation sample into three sub-samples by nearest foreign country of the municipality (Poland, Germany, and Austria) that coincide with the northern, western, and southern parts of the Czech Republic.

The descriptive evidence presented in Figure C.6 suggests that differences between resettled municipalities have higher migration rates than not resettled ones in all parts of the country. Results of the regression (1) estimated separately for each sub-sample corroborate this observation (see Table C.14). Overall the results follow the pattern of the pooled baseline estimates. The effect of the resettlement on emigration rate is positive and significant in all parts of the country. On the other hand the estimates for immigration and net immigration rate tend to be less statistically significant. Differences between regions also seem to minor and—if anything—suggest a moderately weaker impact if estimates are based solely on regressions focusing on the Austrian border.

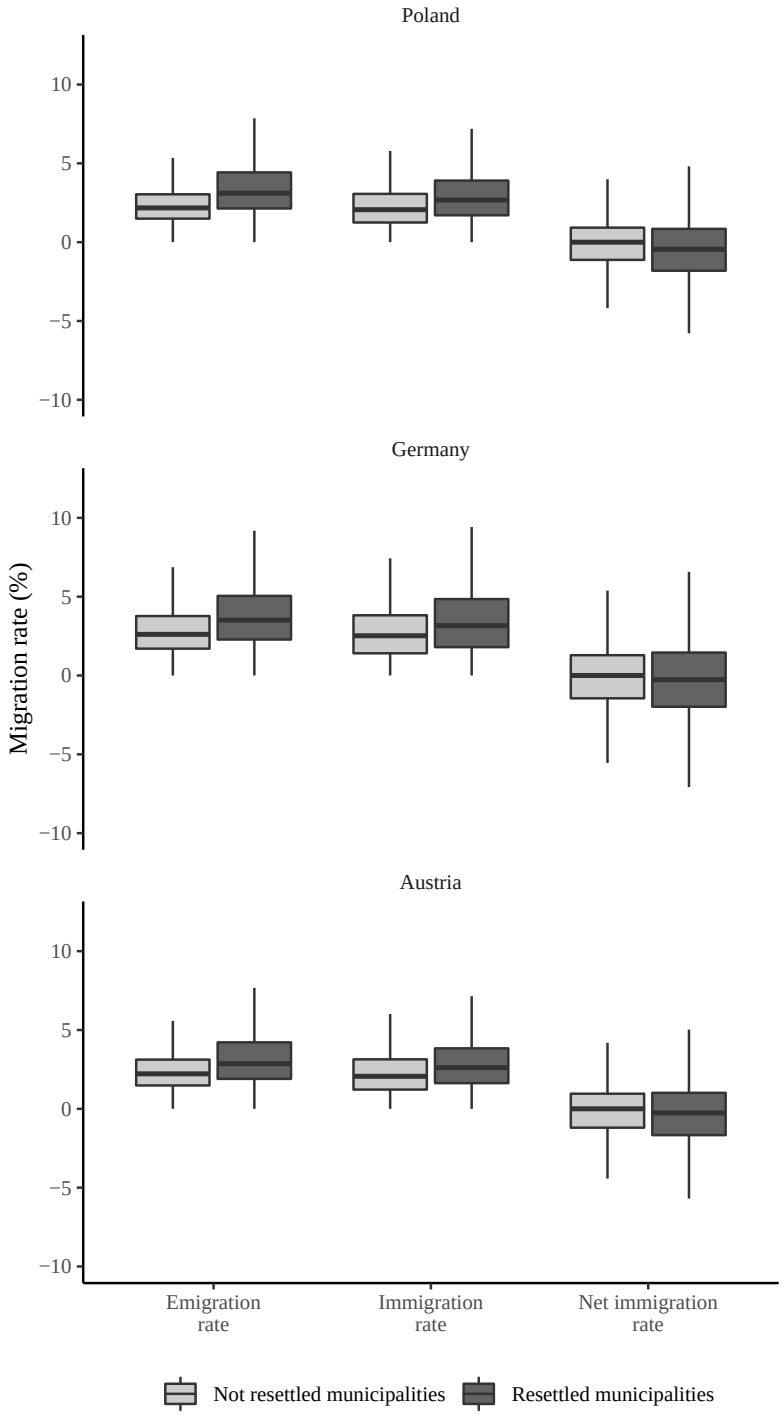
37. Vaněk (1996) discusses an example of such a when he states that 116 municipalities were destroyed due to lignite mining in Northwest part of the Czech Republic.

Table C.13: Robustness test: Inclusion of outliers

	RD-polynomial		
	1st order	2nd order	3rd order
	(1)	(2)	(3)
	Dependent variable: Emigration rate (%)		
Resettled municipality (=1)	0.878***	0.655***	0.543**
	(0.083)	(0.123)	(0.175)
Adjusted R^2	0.161	0.161	0.161
Observations	83849	83849	83849
	Dependent variable: Immigration rate (%)		
Resettled municipality (=1)	0.701***	0.530***	0.441*
	(0.100)	(0.139)	(0.197)
Adjusted R^2	0.069	0.069	0.069
Observations	83849	83849	83849
	Dependent variable: Net immigration rate (%)		
Resettled municipality (=1)	−0.177*	−0.125	−0.102
	(0.084)	(0.120)	(0.157)
Adjusted R^2	0.076	0.076	0.076
Observations	83849	83849	83849

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled municipalities with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

Figure C.6: Migration rates by part of the country



Source: CZSO, own calculations.
Note: Migration rates are reported for municipalities in the estimation sample by nearest foreign country.

Table C.14: Robustness test: Impact of the resettlement by part of the country

	Part of the country								
	Municipalities close to Poland			Municipalities close to Germany			Municipalities close to Austria		
	RD-polynomial			RD-polynomial			RD-polynomial		
	1st order	2nd order	3rd order	1st order	2nd order	3rd order	1st order	2nd order	3rd order
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Dependent variable: Emigration rate (‰)								
Resettled municipality (=1)	0.780***	0.802***	0.761***	0.992***	1.063***	0.767†	0.439***	0.427**	0.417*
	(0.083)	(0.121)	(0.170)	(0.155)	(0.278)	(0.445)	(0.111)	(0.150)	(0.199)
Adjusted R^2	0.179	0.179	0.180	0.149	0.149	0.149	0.153	0.154	0.154
Observations	40685	40685	40685	22897	22897	22897	18841	18841	18841
	Dependent variable: Immigration rate (‰)								
Resettled municipality (=1)	0.532***	0.577***	0.488**	0.767***	0.736*	0.659	0.266†	0.339†	0.486
	(0.087)	(0.127)	(0.182)	(0.171)	(0.301)	(0.469)	(0.148)	(0.194)	(0.301)
Adjusted R^2	0.091	0.091	0.091	0.079	0.079	0.079	0.082	0.082	0.082
Observations	40685	40685	40685	22897	22897	22897	18841	18841	18841
	Dependent variable: Net immigration rate (‰)								
Resettled municipality (=1)	−0.249**	−0.224†	−0.274†	−0.224	−0.327	−0.108	−0.172	−0.089	0.068
	(0.079)	(0.117)	(0.156)	(0.156)	(0.263)	(0.381)	(0.129)	(0.163)	(0.252)
Adjusted R^2	0.107	0.107	0.107	0.098	0.099	0.099	0.095	0.096	0.096
Observations	40685	40685	40685	22897	22897	22897	18841	18841	18841

Note: Panel estimates from Equation (1) in the main text. Only municipalities within a 15 km band around the ethnic border are considered. Results control for year fixed effects, an RD-polynomial, region fixed effects, population fixed effect, log of altitude, terrain roughness, log of the distance to the country border, and 2nd order polynomial of longitude and latitude. Symbols represent statistical significance of differences in means between resettled and not resettled with $^{\dagger} = p < 0.1$, $^{*} = p < 0.05$, $^{**} = p < 0.01$, $^{***} = p < 0.001$. Values in brackets are standard errors clustered by municipality.

C.14 Robustness Check 7: Estimation of values data using the full set of values for responses

As a final robustness check we conducted a linear regressions analysis of the values data used in the main part of the paper using the full set of values for responses. These robustness checks suggest that our choice of coding has few consequences for the quantitative results (see Table C.15). The estimated coefficients once more remain insignificant in most cases. The only exceptions being hobbies, health and material conditions in the old cohort. As also found in the main part of the paper the older cohort in resettled municipalities puts a higher value on the importance of hobbies. Furthermore, in contrast to the results in the main part of the paper, the same applies to the importance of material conditions, while the importance given to health is lower in the resettled municipalities.

Table C.15: Estimation of values data using the full set of values for responses

	Dependent variable: Things and values very important in personal life (Likert scale, 4 = "Very important", 1 = "Not important at all")									
	Nature, environment	Job, occupation	Relationships	Faith, spiritual values	Hobbies	Housing	Friendship	Health	Family life and children	Material conditions
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Young cohort (18–29)										
Resettled municipality (=1)	0.048	0.107	0.019	−0.120	0.130	0.058	0.039	−0.074	−0.150	0.143
	(0.106)	(0.091)	(0.117)	(0.163)	(0.117)	(0.079)	(0.098)	(0.113)	(0.149)	(0.102)
Adjusted R2	0.045	0.051	0.067	0.098	0.028	0.021	0.045	0.036	0.218	0.027
Observations	680	680	680	680	680	680	680	679	680	679
Panel A: Young cohort (18–29)										
Resettled municipality (=1)	−0.041	0.112	−0.212	−0.154	0.320*	−0.021	−0.030	−0.082†	−0.104	0.302*
	(0.114)	(0.235)	(0.137)	(0.181)	(0.153)	(0.099)	(0.107)	(0.046)	(0.103)	(0.141)
Adjusted R2	0.030	0.013	0.017	0.091	0.002	0.013	0.012	0.040	0.125	0.048
Observations	611	602	609	608	610	611	609	610	610	610

Table reports estimated coefficients on an indicator variable for resettled municipalities after controlling for municipality (log of altitude, terrain roughness, log of distance to the country border, region fixed effect, and population fixed effect) and personal (age group, education, labor market status, marital status, household size, and for being born in the municipality of residence) characteristics. Standard errors clustered by municipality are reported in parentheses are : $^{\dagger} = p < 0.1$, $^{**} = p < 0.05$, $^{***} = p < 0.01$, $^{***} = p < 0.001$. Descriptive statistics are the mean and the standard deviations (in parentheses).

References

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocio Titiunik. 2020. *rdrobust: Robust Data-Driven Statistical Inference in Regression-Discontinuity Designs.* R package version 0.99.9. https://CRAN.R-project.org/package=rdrobust.

Kelly, Morgan. 2019. "The standard errors of persistence."

———. 2020. "Understanding persistence."

Vaněk, Miroslav. 1996. *Nedalo se tady dýchat: Ekologie v českých zemích v letech 1968 až 1989.* Praha: Ústav pro soudobé dějiny AV ČR.