JOHANNES KEPLER UNIVERSITY LINZ

DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

On the Origin and Composition of the German East-West Population Gap

by

Christoph EDER

Martin HALLA

Philipp HILMBAUER

Working Paper No. 1817
This Version: January 2023
First Version: November 2018

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

martin.halla@jku.at

On the Origin and Composition of the German East-West Population Gap*

Christoph Eder[a], Martin Halla[a,b,c,d] and Philipp Hilmbauer[e]

[a] *Johannes Kepler University Linz, Austria*
[b] *Christian Doppler Laboratory Aging, Health and the Labor Market, Linz*
[c] *IZA, Institute for the Study of Labor, Bonn*
[d] *Austrian Public Health Institute, Vienna*
[e] *Central European University, Vienna*

January 19, 2023

Abstract

The East-West gap in the German population is believed to originate from migrants escaping the socialist regime in the *German Democratic Republic* (GDR). We use newly collected regional data and the combination of a regression discontinuity design in space with a difference-in-differences approach to document that the largest part of this gap is due to a massive internal migration wave 3 years *prior* to the establishment of the GDR. The timing and spatial pattern of this migration movement suggest that the dominant motive was escaping physical assault by the Soviet army and not avoiding the socialist regime. The skill composition of these migrants shows a strong positive selection. The gap in population has remained remarkably sharp in space and is growing. These patterns are equally evident when including *all* counties in the estimation sample.

JEL Classification: N44, N94, R23, R11, R12, J61.
Keywords: Institutions, wartime violence against civilians, selective migration, regional migration, World War II, Germany, spatial distribution, regional economic activity.

* *Corresponding author*: Martin Halla, Johannes Kepler University Linz, Department of Economics, Altenbergerstr. 69, 4040 Linz, Austria; email: martin.halla@jku.at. Philipp Henning and Rene Wiesinger provided excellent research assistance. The usual disclaimer applies.

1 Introduction

At the end of WW2 Germany was divided. The separation of this single nation proceeded in two steps. First, in 1945, Germany was divided by the victorious Allies into four occupation zones (see left map in Figure 1). Second, in 1949, it was divided in two countries. These two distinct political entities undertook the process of reconstruction, each guided by its own set of ideological principles (Witte and Wagner, 1995). Citizens in the 'East' were exposed to a centrally planned and largely state-owned economy. Citizens in the 'West' experienced the famous *Wirtschaftswunder* with economic aid provided by the US and the Marshall Plan in a free market economy. In 1990, the division of Germany came to an end.

This unique policy episode has attracted the attention of scholars across social sciences. A large number of studies in economics exploit it as a natural experiment to study the effect of institutions in post-reunification data. This approach rests on the assumption that the division of Germany was uncorrelated with confounding factors. Studies in contemporary German history, however, point to East-West migration in the aftermath of WW2 (see, e. g., Fassmann and Münz, 1994; Grundmann, 1998). These studies use exclusively aggregate statistics with time series starting in 1949.[1]

Thus, the exact timing and the spatial pattern of this migration wave is not documented. Nevertheless, this literature postulates that the origin of this migration movement is the emerging socialist regime in the East.

We provide evidence that the East-West gap in population originates from an internal migration wave of positively selected individuals that pre-dates the institutional shock. The bulk of the East-West migration took place during the period of the *Allied-occupied Germany*. Only a small fraction of the East-West migration in the aftermath of WW2 can be attributed to refugee flows after the establishment of the *German Democratic Republic* (GDR). The specific timing and spatial pattern of the initial migration wave suggest that the dominant motive of migrants was escaping physical assault by the Soviet army and not the emerging socialist regime. This finding is not only interesting from a cliometric perspective, but has also ramifications for the economic literature exploiting the division and reunification of Germany as a natural experiment to identify the impact of institutions. Our findings on the timing and composition of the East-West migration renders the interpretation of these studies to be difficult. We further highlight that the East-West population gap has remained remarkably sharp in space at the former demarcation line until today and is still growing. This result speaks to the debate on the existence of multiple equilibria in economic geography and suggests the presence of large agglomeration economies.

Our empirical analysis is based on newly compiled regional data from population

[1] See Appendix Table A.1 for a list of relevant references.

censuses spanning the period before, during and after WW2. We compare population levels between counties bordering the former demarcation line, established on 1 July 1945 as the boundary between the Western and Soviet occupation zones of former Nazi Germany. This line became the so-called *inner German border* in 1949 dividing East and West Germany until 1990. Our long series, dating back to 1900, allows us to demonstrate that the regions east and west to this border were following parallel trends in population development prior to WW2.

This suggests that the exact position of the demarcation line was exogenous. We combine ideas from a *Difference-in-Differences* (DiD) approach with those from a *Regression Discontinuity Design* (RDD) to show that the divergence in population levels between the East and the West can be traced back to the period between January 1945 and October 1946. In this short period of time, the population level in the East had dropped by almost 20 percent.[2] Over the subsequent 15 years, when escaping the East was still relatively easy, the population gap had only increased by another 5.1 percentage points. Thus, the largest part of the East-West migration was completed 3 years prior to the establishment of the GDR (in 1949), and 16 years before the construction of the Berlin Wall (in 1961). These patterns are equally evident when including *all* counties in the estimation sample.

We consider two competing explanations for the massive migration wave. First, migrants may have reacted to an *expected* institutional shock (and not as widely believed to an actual change in institutions). This explanation presumes that the vast majority of the German population had correct believes about the political future of the East already in the year 1945/46. The second explanation is that migrants reacted to the immediate threat of physical and sexual violence by the Soviet army, which successfully entered Germany in January 1945 on the Eastern Front. The specific timing of the migration wave and the second discontinuity in population density it created at the Line of Contact within the East, suggest that escaping the Soviet army (and not the socialist regime) was the main motive to migrate. More importantly, we find that migrants to the West were positively selected by their skills as compared to the stayers. Workers with a background in agriculture were about 30 percent more likely to stay in the East, whereas workers in manufacturing had a higher likelihood to migrate to the West.

Studies exploiting the division and reunification of Germany as a natural experiment aim to quantify the impact of the exposure to a socialist regime as compared to a democratic regime.[3] These studies provide overwhelming evidence for a significant East-West

[2]Eder and Halla (2016) document a comparable migration response from the Soviet to the non-Soviet zone in the case of the Allied Occupation of Austria. Using a comparable method, they find a reduction in the population in the Soviet zone of 11 percent between 1939 and October 1946. They do not provide any evidence on the migrants motive.

[3]Most papers studying micro-level outcomes focus on human capital and social behavior. The human capital outcomes studied are labor productivity (Fuchs-Schündeln and Izem, 2012) and educational attainment (Fuchs-Schündeln and Masella, 2016). The list of social behavior and attitudes comprises

gradient and attribute this difference exclusively to institutional differences. The vast majority of these studies share three methodological features: First, they exclusively analyze post-reunification data. This can be explained by the lack of historical data of the outcomes under consideration. Second, they follow a reduced-form approach. They do not aim to identify a specific causal channel, but quantify an East-West gap at a certain point in time after reunification. Third, they do not exploit the *local* randomness (in space) generated by the discontinuous border between East and West Germany. There are three exceptions. Redding and Sturm (2008a) analyze the effect of market access on the development of cities in West Germany using data covering the period from 1919 to 2002. Ahlfeldt *et al.* (2015) develop a quantitative model of internal city structure and provide structural estimates of agglomeration and dispersion forces using block-level data from Berlin for the years 1936, 1986, and 2006. Lichter *et al.* (2021) studies the effects of government surveillance in the GDR on social capital and economic performance after reunification.

Our finding adds to this literature by providing new evidence on the causal driver of the onset of the East-West population gap and the positive selection of migrants.[4] Our findings also add important context to the interpretation of previous studies that use the German division as a natural experiment to examine the effects of communism. Differences in post-reunification outcomes have *two potential sources*. They could either result from the exposure to the socialist regime and/or originate from the pre-existence of differences between stayers and movers and their offspring. While a detrimental causal effect of the exposure to the socialist regime is plausible, our evidence for the positive selection of East-West migrants in 1945/46 is also in line with worse post-reunification outcomes in the East. Given our evidence — that the dominant motive was escaping physical assault, and not avoiding the socialist regime — this positive selection into migration *should not even be interpreted as a reduced-form effect of socialism*. Becker *et al.* (2020) also argue that the division of Germany should not be interpreted lightly as a natural experiment. The authors main criticism is that East and West German populations were already different in the 1920s. They show that the population in the East was more likely working class (measured in 1925), less likely self-employed (in 1925), more likely to vote

political preferences (Alesina and Fuchs-Schündeln, 2007), trust (Rainer and Siedler, 2009), solidarity (Brosig-Koch *et al.*, 2011), gender attitudes (Bauernschuster and Rainer, 2012), self-reliance and entrepreneurship (Bauernschuster *et al.*, 2012), honesty (Ariely *et al.*, 2019), conspicuous consumption (Friehe and Mechtel, 2014), and tax morale (Möhlmann, 2014). Papers using more aggregated data study the effect on regional economic performance (Abadie *et al.*, 2015; Boltho *et al.*, 2018), and migration (see footnote 4).

[4] A number of quantitative papers study the East-West migration in the *post*-reunification period (Uhlig, 2008), when large parts of the German population moved from East to West leading to a widening in the population gap (see Figure 2). This literature analyzes the pattern and the composition of migration (Burda, 1993; Burda *et al.*, 1998; Hunt, 2006; Fuchs-Schündeln and Schündeln, 2009) and point to selective migration. Migrants to the West are comparably young, have above average education, and come from regions with low regional income. Workers recently laid-off have also a higher propensity to migrate.

for the communist party (in 1924), and less likely to attend church (in 1910). We add to this literature by showing that selective East-West migration in 1945/46 reinforced pre-existing differences between East and West German populations.[5]

In a second step, we use our research design to analyze the persistence of the East-West population gap in the post-reunification period. We find that the inner German border (i.e., the former demarcation line) left a remarkably sharp discontinuity in space with respect to population density. The East-West population gap is still increasing at this location. While we cannot disentangle to which degree this persistence is caused by the migration wave in 1945/46 and how important the past institutional differences are, this result speaks to the literature in economic geography, which tries to explain the (uneven) distribution of population across space. In models featuring locational fundamentals, high population density is the consequence of inherent productivity advantages of a specific location (such as topographical and climatic characteristics) that cannot be influenced by migration or an institutional shock. In contrast, scale economy models stress the importance of local interaction of economic agents, in which higher population density endogenously leads to higher productivity in a location (Henderson, 1974; Krugman, 1991). Scale economy models generally allow for multiple equilibria. A sufficiently large population shock — such as East-West migration — might shift the local economy to another spatial equilibrium, if the shock sets free agglomeration forces that outweigh dispersion forces. In this way a temporary shock can affect long-run outcomes. In contrast, models stressing locational fundamentals predict a convergence back to the initial spatial equilibrium after a temporary shock.[6] Our finding clearly corroborates scale economy models.

2 Research Design

Our research design is based on the idea that the onset of the East-West population gap can be identified by focusing on migration movements around the East-West border. This border was established in 1945 within the *Allied-occupied Germany* as a demarcation line. In 1949, it became the inner German border dividing the GDR in the East and the *Federal Republic of Germany* (FDR) in the West. We have to overcome two challenges to identify the onset of the East-West gap in population at this discontinuity in space.

[5] In response to Becker *et al.* (2020), some recent papers in this literature aim to address this pre-existing differences. For instance, Bondar and Fuchs-Schündeln (2022) provide suggestive evidence that the empirical approach by Alesina and Fuchs-Schündeln (2007) is immune to Becker et al.'s criticism.

[6] Existing empirical papers testing these two theories (i. e., exploiting exogenous variations in scale, while holding locational fundamentals constant) find mixed evidence. While the findings in Davis and Weinstein (2002, 2008), Brakman *et al.* (2004), and Miguel and Roland (2011) support models of locational fundamentals, the evidence shown in Bosker *et al.* (2007, 2008), Schumann (2014), and Peters (2019) favor scale economy models. Braun, Kramer, Kvasnicka and Meier (2020) aim to reconcile these findings with certain commuting streams.

First, we should allow for the possibility of unobserved differences between areas in the two occupation zones that were already in place before the demarcation line was decided. Becker *et al.* (2020) show that this is indeed a serious concern; neglected by almost all studies using the inner-German border as a natural experiment. To address this issue, we have collected a long data series starting in 1900. These data allow us to compare population levels and trends across regions in a period before separation. It turns out that the regions east and west to the demarcation line followed parallel trends in population development prior to WW2. This suggests that the exact position of the demarcation line was exogenous with respect to population growth trajectories. Motivated by these parallel trends in the pre-occupation period, we assume in our analysis that the population trends would have been parallel in the absence of the separation later on.[7] Second, we have to be careful to rule out other time-varying confounding factors, such as differences in the proximity to Western markets in the post-WW2 period. To address this, we exploit the demarcation line as a discontinuity in space. Small geographic units bordering the demarcation line have the same geographic features and equal access to markets. More generally speaking, we assume that there are no confounding factors, which change discontinuously at the demarcation line.

2.1 Data

The history of the German nation over the last 100 years includes multiple shifts of borders at all administrative levels. The resulting lack of traceable administrative units makes it hard to connect data from population censuses over time. So far, economic scholars interested in longer series of German data have focused on larger cities (Brakman *et al.*, 2004; Bosker *et al.*, 2007; Redding and Sturm, 2008a), which are relatively easy to organize in a longitudinal data set. Since we are particularly interested in the development of population levels in a smaller geographic area around the former inner German border, a focus on cities is not conducive. Fortunately, we were successful in reconstructing a panel dataset at the county (*Kreis*) level.

The construction of our data set comprises three main steps. First, we collect data from historic population censuses covering the time period between 1900 and 2009 and calculate the population density of each historic county. We have six data points before WW2 (1900, 1910, 1919, 1925, 1933, 1939), three during WW2 (1943, 1944, 1945), one during the occupation period (1946), six during the division of Germany (1950, 1961, 1964, 1971, 1981, 1987), and three after reunification (1991, 2001, 2009). For the years during WW2 we use data on the number of issued food stamps to approximate popula-

[7] Our estimation procedure is equal to a DiD approach. In contrast, to a standard DiD approach — which assumes that only one group was affected by the treatment — we recognize that both the East and West have been affected by the events after WW2 and we aim to estimate the relative difference in population.

tion.[8] Second, we connect these historic county data with GIS-shapefiles of the respective historic county borders. Third, we superimpose the GIS-shapefile of the current county borders on historic population densities and calculate the average historic population density of each current county. This procedure provides us with an approximation of the true historic population levels in every current county. We can assess the quality of our approximation by using special data from the state of Bavaria. Bavaria is the only region, which constructed official statistics on the historic population based on current country borders by taking account of all border changes since 1840. This allows us to compare our approximation with exact data. It turns out that our approach works exceptionally well. We obtain a correlation coefficient between the log population variables in cross-sections of these two data sets covering the whole of Bavaria between 0.94 and 0.99.[9]

The census of 1946 is of crucial importance in our analysis, but also contains the biggest concern about data quality as Germany was split into occupation zones at the time. The census was entirely conducted by German officials, except for the number of camp inmates. A committee of German statisticians (*Ausschuß der Deutschen Statistiker für die Volks- und Berufszählung 1946*), appointed by the four occupation forces, were instructed to present population numbers that are comparable across all states of Germany. Nevertheless, there have been differences in the treatment of camp inmates in the four zones. While the Western allies included camp inmates to the resident population, the Soviets did not (Ausschuß der Deutschen Statistiker für die Volks- und Berufszählung 1946, 1949). However, this affects only 1.7 percent of the total population in 1946 and can not be the main explanation of our results.

Figure 2 shows the mean population development in the East and West for different samples. In all panels there is, relative to the year 1939, a diverging trend in the population levels in the East versus the West. In the East, the total population, the population in states along the border, the population in counties along the border, and the urban population has decreased. Note, information on population in 1943, 1944, and 1945 is not available in the latter sample. The increase in the total German population between 1939 and 1946 is driven by the immigration of ethnic Germans from Eastern Europe, which overcompensate population losses due to WW2.

[8] The number of issued food stamps was published by the West German statistical agency (*Statistisches Bundesamt*) as the only statistical documentation on the fluctuation of population during WWII. Because the results of the census of 1939 were not available during the war, food stamps constitute an independent source of information on the civilian population.

[9] Another way to assess the quality of our approximation is to plot the series for each county and to check for sudden jumps or drops. The vast majority of our series are very smooth. Appendix Table A.2 provides descriptive statistics for different sample definitions. Detailed data sources are listed in Appendix Table A.3.

2.2 Estimation Strategy

The core idea of our estimation strategy is to exploit the demarcation line (i. e., the latter inner German border) as a discontinuity in space while accounting for pre-WWII differences. This lends itself to a conventional RD approach, in which the distance to the demarcation line serves as the running variable, with population growth since 1939 as the dependent variable. A drawback of this approach is the mismatch between a one-dimensional running variable in a two-dimensional plane. Our preferred approach accounts for the two-dimensionality of space in a simple but effective way. We focus on the sample of counties that border the demarcation line highlighted in the right map of Figure 1. Among these, we form pairs of areas that share a common border (which is the demarcation line). For each of these pairs we calculate the difference in the population level for each year and compare the mean of the differences over time. This approach translates into the following estimation model:

$$P_{i,j,t} \quad = \quad \alpha + \sum_{t} \beta_t \cdot East_{i,j} \cdot Year_t + \sum_{j,t} \phi_{j,t} \cdot Pair_j \cdot Year_t + \varepsilon_{i,j,t}, \tag{1}$$

where $P_{i,j,t}$ is the log population in county i, belonging to pair j, measured in year t. The binary variable $East_{i,j}$ is equal to one if the county is in the East (and zero otherwise), and the binary variables $Year_t$ denote years. The estimate of $\phi_{j,t}$ denotes a time-varying fixed-effect for county-pair j in year t. These are quite powerful controls, since they account for all time-varying factors that affect the population levels of bordering counties on both sides of the former demarcation line.

The parameters of primary interest are the β_t. These parameters provide the average difference between the population of a county in the West to one in the East in a given year t relative to the baseline year of 1939. Estimates of β_t for years before WW2 test for differential pre-occupation trends and provide suggestive evidence for the parallel-trend assumption. Estimates of β_t post WW2 show at what point in time the East-West population gap arises and how it has developed over time. The estimate β_{1939} is the average difference in the outcome variable between counties in the East and the West in 1939.

By construction, many counties along the demarcation line appear in several area-pairs. Therefore, we cluster standard errors by county within a pair. In Section 3.4, we will demonstrate the robustness of our estimation results with respect to different approaches of inference and alternative RDD-setups. In Section 4, we use alternative (broader) samples to highlight the generalizability of our findings.

3 Estimation Results

Figure 3 depicts our estimation results of equation (1). More detailed estimation output is available in Column (I) of Table 1. The first important result is that all six estimates in the period before WW2 are statistically and economically insignificant. Thus, the population levels developed in the period from 1900 to 1939 almost identically on both sides of the demarcation line. Note that the statistically significant level difference in 1939 is due to larger counties in the East. During WW2 the relative distribution of the population between East and West has also remained constant. We do not find any statistically significant differences for the years 1943 to 1945 as compared to 1939. This supports the assumption that the *exact* position of demarcation line was exogenous.

The estimates for later years will inform us about the onset of the East-West population gap and its development. Below, we first discuss in Section 3.1 the onset, which can be located during the short period of the Allied-occupied Germany lasting from the end of WW2 until the division of Germany in 1949. We then describe the development of the gap from 1949 until the construction of the Berlin Wall in 1961. This period can be characterized by a very modest increase in the gap despite relatively open borders. In Section 3.2, we briefly comment on the development after the construction of the Berlin Wall until 1991. During this period East-West migration was strictly monitored and regulated. In Section 3.3, we discuss the development of the gap in the post-reunification period. Finally, in Section 3.4 we demonstrate the robustness of our findings with respect the specific sample and estimation method used.

3.1 The Origin of the East-West Population Gap

The most remarkable feature of Figure 3 is the sudden drop in population between January 1945 and October 1946. In this short period of time, the population level had dropped in the East by 18.4 percent. The most plausible explanation for this drastic change is a large migration wave from East towards West between January 1945 and October 1946. We see two potential causes for this migration movement.

First, the German population wanted to avoid an encounter with the Soviet army, which successfully entered Germany in January 1945 during the *Vistula-Oder-Offensive* on the Eastern Front, and escaped westbound. Due to Nazi propaganda demonizing communists, as well as factual reports on misconduct of the Soviet Army in Hungary, the German population was terrified by the Soviet Army. Sadly, the seeking of revenge and craving for booty indeed led to assaults on the local population. In particular, there is evidence for mass rapes taking place in connection with combat operations, but also during the subsequent occupation (Dack, 2008).[10] By contrast, the reputation of

[10]The best available evidence is for Berlin. Using information from hospital records, Johr (1992) estimates that in the period between April 1945 and September 1945 about 7 percent of all women of

the troops of the Western Allies, who crossed the German border in the West about one month later, was much better. While there are also documented cases of rape, the incidence seems much lower.

Second, the German population may have had already formed correct expectation about the political future of the different occupation zones and their primary motivation for migration was to avoid living in a (Soviet led) communist country. Already starting with the *Tehran Conference* in 1943 the Allies started discussing a post-war division of Germany. However, only at the *Yalta Conference* in February 1945, the major Allies agreed on the boundaries of post-war occupation zones for Germany. The international press discussed, shortly before and after the *Tehran Conference*, the division of Germany and even included some maps. However, these maps provided only a vague idea of the different zones. It is hard to reconstruct whether the German population, which had very limited access to international media, was aware of these plans.[11]

While the onset of East-West population gap can clearly be traced back to the period between January 1945 and October 1946, it is impossible to unambiguously uncover migrants' motivation. Below, we provide two findings, that provide at least suggestive evidence.

3.1.1 Timing of the Migration Wave

Between 1946 and our next data point in the year 1950, the East-West population gap increased only by about 3 percentage points. Until 1961, the gap increased further; but only modestly by another 1.9 percentage points. Thus, the bulk of the East-West migration was completed 3 years prior to the establishment of the GDR (in 1949), and 16 years before the construction of the Berlin Wall (in 1961). That means that about 80 percent of the migrants, leaving the East between 1945 and 1961, were actually not exposed to the socialist regime.

This specific timing of the East-West migration is more in line with the first motive (escaping the Soviet army) and less consistent with the second motive (escaping the socialist regime). After the establishment of the GDR there was no remaining uncertainty about the political future of the East. However, it was still possible to slip from East to West. Between 1950 and 1952, it was relatively easy to cross the inner German border at any location.[12]

childbearing age were raped at least once by members of the Soviet army.

[11] A complementary cause could be forced migration. There are some reports that Soviet officials ordered people to move to the US zone in order to avoid caring for them. However, this phenomenon was quantitatively less relevant.

[12] Two months after Germany's unconditional surrender in May 1945, the *Allied Control Council* (ACC) was founded. This military occupation governing body of the Allied Occupation Zones was initially in charge of all border control operating procedures. The ACC officially closed all zonal borders per 30 June 1946. Thereafter, all persons had to obtain an interzonal pass to visit another zone. It is however documented that regulations and border control elements were circumvented regularly. While the Soviets

In 1952, the GDR erected barbed-wire fence along the whole inner German border. Between 1952 and 1961, it was still possible to escape to the West through Berlin, which had a loose border between its Western and Eastern sectors. This last loophole was finally closed in 1961, when the construction of the Berlin wall started. It seems implausible that such a large number of people migrated because of an *expected* institutional shock, while only a small number react to the *actual* exposure to the less favorable political regime.

3.1.2 Spatial Pattern of the Migration Wave

To provide further evidence that the progress of the Soviet army into German territory and the ongoing (sexual) assault on civilian population was the dominant migration motive, we examine the spatial pattern of the migration wave in more detail. We exploit that the farthest advance of Soviet Armies into German controlled territory (also known as the *Line of Contact*) is located east to the demarcation line. Thus, there is a region, which was conquered by Western Allies, but became part of the Soviet occupation zone later (see Figure 4). The British and US troops withdrew from there in July 1945 and passed the territory on to the Soviets.

Given that the majority of assaults took place in connection with combat, this area should have witnessed a lower number of (sexual) assaults as compared to the rest of the East. On the other hand, the areas on both sides of the Line of Contact were within the pre-determined Soviet occupation zone and faced the same expectations about a future socialist regime. If migrants were escaping the socialist regime, they should have proceeded past the demarcation line. In contrast, if migrants were escaping the Soviet army, it would have been sufficient to cross the Line of Contact. If the latter is true, we should find a discontinuity in population density at the Line of Contact.

To test this hypothesis, we estimate the effect of the Soviet forces within county pairs along the Line of Contact. We employ the same estimation strategy of a within county pair comparison as we did along the demarcation line described by equation (1). Column (I) of Table 2 summarizes estimation results. As in the case of the demarcation line, we do not find any significant difference in population levels at the Line of Contact before and during WW2. However, there is a significant drop in population levels by almost 13 percent in 1946. This population gap vanishes gradually over time. By 1971, the population gap is gone. This suggests that the significant force of the East-West migration wave in 1945/46 was escaping the physical assault by the Soviet army and not the expectation about differences in the future institutional framework.

began to apply strict border control procedures in September 1947, with an increased number of border guards and help from the newly established *East German Volkspolizei*, it was still fairly easy to cross the border (Stacy, 1984). The situation continued even after the declaration of the GDR in October 1949 until 1952.

The assignment of todays counties to the East or West side of the Line of Contact is in some instances ambiguous. We perform a sensitivity check for these cases in the remaining columns of Table 2. In column (II), we drop all pairs including the city of Dessau-Roßlau of which at least a small part was captured by US forces. The same is true for the city of Magdeburg, which we drop in column (III). In column (IV), we drop the pairs including the county of Nordsachsen, where the US forces pushed forward to first meet Soviet forces on German ground. Reassuringly, the results are not sensitive to these sample modifications, even when we drop all of these pairs in column (V). We also perform a combined regression for the population development around the demarcation line and the the Line of Contact. This regression provides equivalent results (see Appendix Table A.4).

We consider the timing of the East-West migration wave and the second discontinuity in space at the Line of Contact as suggestive evidence for the supposition that the migrants predominantly escaped from the Soviet Army fearing assaults. Clearly, the emerging socialist regime – which lead to subsequent East-West migration, however, at a much lower scale – might have been a factor for migrants of the first wave to stay in the West.

3.1.3 Composition of the Migration Wave

In our analysis above, we were only concerned with the size of migration flows, and ignored the characteristics of migrants. We now provide evidence on the composition of the East-West migration wave in 1945/46. In particular, we are interested in the distribution of sex, skills and regions of origin.

Sex and skill distribution An important aspect of migration movements is the selection of immigrants with respect to their skills (Borjas, 1999). The theoretical literature, mostly building upon the Roy model, highlights the relative earnings potential of low versus high skilled workers in the sending and receiving region as the main determinant of selection (Chiswick, 1978; Borjas, 1987). Empirically, one can examine the observable characteristics of stayers and movers to provide evidence on the type of selection into migration. We observe population by sex and occupation between 1939 and 1961 in county-level data. We use these data to estimate a model equivalent to eq. (1), where the dependent variable is equal to the share of population of a certain sex or occupation. Our estimation results are summarized in Table 3. Column (I) shows that there is no evidence of selection into immigration by sex. One might have expected that women had been comparably more terrified by the Soviet Army, and therefore more prone to leave. However, even if this supposition is true, migration could still have been a family decision resulting in sex-proportional migration. Column (II) and (III) provide evidence for selection by occupation. We find that the share of workers in agriculture has increased

in the East by almost 9 percentage points between 1939 and 1950. This effect is (relative to share of workers in agriculture in the East in 1939) equivalent to an increase by more than 29 percent. Inversely, the share of workers in manufacturing has decreased by 5 and 10 percentage points in 1950 and 1961, respectively. This means that individuals with an occupational background in manufacturing were substantially more likely to migrate to the West. This pattern is in line with higher migration cost for farmers, who could not transfer their landholdings.An alternative explanation is, that after WW2, workers in the West were more likely to change their occupation from agricultural work towards industry. This could be explained by higher economic success in the West as compared to the East. However, since we observe already economically relevant changes in 1950 (and not only in 1961), we consider this alternative explanation of secondary importance. Under the assumption that the skills of workers in manufacturing and services were (compared to those in agriculture) more productive in the post WW2 economy, we interpret this as a positive selection into migration to the West.

Unfortunately, data restrictions do not allow us to analyze other outcomes. We are not aware of any other outcomes, which had been consistently measured before and after WW2, and in both, FDR and GDR.

Region of origin After WW2 there were several significant movements into and out of Germany. Therefore, it is useful to distinguish between three population groups, depending on their place of residence prior to WW2. First, there are German residents who were residing on today's German territory (henceforth natives). Second, there are German citizens who were residing in prewar German territory east of the Oder and Neisse Rivers (henceforth expellees). These areas belong, according todays borders, to several Eastern and Central European countries (such as Czech Republic, Poland, Slovakia, and Romania). Third, there are people who were residing outside prewar Germany before the war (henceforth foreigners). The latter group comprises ethnic Germans who were residing outside the prewar German territory (e. g. Sudeten Germans from Czechoslovakia.), former prisoners of war, and other displaced persons. The group of expellees is quantitatively very important.[13] In 1946 (our first post-WW2 data point), 9.7 million expellees lived in Allied-occupied Germany, a number that grew to 11.3 million by 1950. Thus, the first group could have contributed to the onset of the East-West population

[13]Their migration movement occurred in three overlapping phases. First, from mid-1944 to early 1945 there were some organized evacuations by the Nazi government in the face of the advancing Soviet Army. Second, following the Wehrmacht's defeat in January 1945, many ethnic Germans (not covered by previous organized evacuation) escaped on their own initiative and spontaneously. These formed kilometers-long refugee treks pushing their carts through snow trying to stay ahead of the advancing Soviet Army, with many of them eventually targeted by low-flying aircraft and some crushed by tanks. The third phase was a more organized expulsion following the Potsdam Conference later that year. In the period between mid 1945 and 1950 all remaining ethnic Germans in prewar German territory east of the Oder-Neisse line were transferred to Germany.

gap (i. e., the spike in Figure 3), if they have predominantly moved to the Western zone as compared to the Soviet zone.

We have access to state-level population data for the years 1946 and 1950 by location of residence in 1939. We distinguish the three groups defined above and a residual group with unknown origin. We use these data to estimate a model equivalent to eq. (1), where the dependent variable is equal to the ratio of population group g in year t to the total population in 1939. Columns (I) to (III) of Table 4 summarize the estimation results. We see that the East-West population gap is driven by natives. For this group, we observe an economically and statistically significant drop. The estimated effects for expellees and foreigners are much smaller and statistically not significant. The residual group exhibits a larger presence in the East in the year 1950, but the effect is quantitatively negligible.[14] We conclude that East-West migration wave in 1945/46, and the resulting onset of the East-West population gap, was driven by the group of natives. Thus, the event can be described as an internal migration phenomenon. However, one has to bear in mind that this finding is based on state-level data with a substantially lower number of observations.

3.2 Period of the Berlin Wall

Between 1961 and 1989, essentially no migration between East and West Germany took place.[15] During this period the population gap increased in our sample from minus 23.5 to minus 28.4 (see column I of Table 1). This development can be attributed to differences in the 'normal' demographic processes of birth, death and external migration. The East had, probably due to a set of pronatalist policies, higher fertility as compared to the West (Büttner and Lutz, 1990). Mortality was also comparably higher in the East, since the West experienced higher gains in life expectancy during this period (Heilig *et al.*, 1990). Finally, the West attracted much higher number of external immigrants. Most importantly, the strong economy in the West attracted large numbers of guestworkers from Southern Europe and Turkey. The East experienced a much smaller influx of workers from socialist nations (such as Vietnam or Mozambique).

3.3 The East-West Population Gap after Reunification

After the fall of the iron curtain and the reunification, inner German mobility was restored immediately. Figure 3 shows that this led to further East-West migration. Over

[14] Column (V) of Table 4 lists results for the total population (i.e., sum of all groups). This estimation provide the same qualitative result as the specification in Column (I) of Table 1. However, the estimated effect size differs. This difference can be explained by the differential geographical coverage (bordering states vs. bordering counties) and the differential unit of observations (states vs. county).

[15] During the summer of 1989 Hungary reduced its border patrol to Austria. A large number of East German citizens exploited this situation and traveled to Hungary as a tourist to enter West Germany via Austria (Heiland, 2004).

the entire post-reunification period the East-West population gap has grown. Column (I) of Table 1 shows that the gap has increased from 28.4 percent in 1987, to 37.3 percent in 1991, and to 51 percent in 2009. While this general pattern is well-known in aggregate numbers, our results reveal, based on disaggregated units, how large and sharp the difference in population has remained at the former demarcation line. Clearly, our estimated effects have to be interpreted as reduced form estimates. They capture the sum of all differential treatments the East has experienced (as compared to the West) since the end of WW2. Nevertheless, our estimations results are informative to discriminate among the two leading explanations for the (uneven) distribution of population and economic activity in space.

The literature in economic geography discusses locational fundamentals vs. economies of scale as the determinants of the spatial distribution of economic activity. Given the increasing difference in the East-West population difference after reunification, we can reject that locational fundamentals produce this pattern. After the removal of migration restrictions, locational fundamentals theory would predict a reversal to the population difference towards its pre-WW2 level. Our findings show no evidence that this has happened in the almost 20 years after reunification.[16] Scale economy models, on the other hand, suggest that the population shock increases productivity or consumption amenities in the West and hence attracts additional workers/households from the East after reunification. Our findings are perfectly consistent with this prediction. It is surprising, however, that the population difference between counties along the former demarcation line remains so distinctively sharp over time. Firms and households in the West would face lower land prices and lower wages by relocating just across the former demarcation line. The sharp discontinuity suggests that productivity and amenity spillovers originating from the higher population density in the West decline relatively fast in space.[17]

Other factors that could affect the population distribution in space include differences in labor market regulations after reunification and public infrastructure investment in the East. While labor market regulations favored the West with, for instance, higher collective bargaining wages, the massive investment in infrastructure would have increased the population in the East.

[16]Of course, we cannot rule out that the pre-shock spatial distribution will eventually be restored. In comparison, Davis and Weinstein (2002), who use the Allied bombing of Japanese cities in WWII as a shock to relative city sizes, find that already after 15 years most cities have returned to their relative position in the distribution of city sizes. Using data from the so-called coordinated population projection for the year 2017, provided by the Federal Statistical Office of Germany (*Statistisches Bundesamt*), we can extend our estimation period by another ten years. These population data are less reliable, since these are not census based. Appendix Figure A.1 shows that there is no trend reversal; if anything, we see a further widening in the gap. Thus, we tend to interpret our result for up to 27 years after the shock as convergence towards a new steady state.

[17]One form of productivity increases in the West could have been a direct result of the reunification as well. A complementarity between high-skilled West-German and low-skilled East-German workers would lead to an increase in productivity if these two groups come together in a single labor market. Higher bargaining power of western high-skilled workers would let East-Germans move to the West.

3.4 Sensitivity Analysis

We now briefly discuss results based on alternative approaches for inference and from a conventional one-dimensional RDD. We also explore treatment effect heterogeneity along the North-South dimension of the demarcation line.

3.4.1 North versus South

It is interesting to explore whether the development of the population gap varies along the demarcation lines as one goes from north to south. To explore this potential dimension of heterogeneity, we split the sample in two samples of equal size. We find the same pattern across samples, with somewhat larger differences in the North as compared to the South (see Appendix Table A.5).

3.4.2 Calculation of Standard Errors

In our estimation strategy, we form pairs of bordering counties along the demarcation line. By construction, many counties appear in several of these county-pairs. This feature of the estimation strategy increases our sample size and gives rise to several different ways to cluster standard errors. In our baseline specification in column (I) of Table 1, we cluster standard errors at the county level within a pair.

We now investigate other options of clustering and calculating standard errors. Column (I) of Table 5 repeats our baseline estimates for reference. In columns (II)-(V) we cluster standard errors at the county level, at the pair level, at an East-German county level including all bordering West-German counties, and at the West-German county level including all bordering East-German countries, respectively. In each case, we obtain statistically significant effects. In the next two columns, we address that some counties enter several pairs. In column (VI), we transform the dataset such that for each East-German county, there is only one synthetic control county. That synthetic control country consists of the average of all bordering West-German counties. Column (VII) performs the same exercise for each West-German county. Reassuringly, the point estimates and standard errors are very similar to our baseline estimates.

3.4.3 Conventional Regression Discontinuity Design

We also perform a one-dimensional RDD. Therefore, we restrict the sample to counties with a distance of at most 40 kilometers to the demarcation line and use this distance as the running variable. We calculate the distance as the minimum distance of the centroid of a county to the centroid of a county on the other side of the demarcation line and subtract the lowest distance of a county within each zone. Figure 5 summarizes estimates based on different functional form assumption for zone-specific distance functions. Across

specifications, we observe the same qualitative (and very comparable quantitative) results for the origin and development of the East-West population gap.

4 Generalizability of our Findings

The focus on the area along the demarcation line (or the Line of Contact) bears the risk of missing the larger picture. Our estimated effect may be only a local phenomenon that is specific to the counties along these lines. First, with respect to the origin of the population gap, we have to consider that migration cost vary across different points of departure (and destinations). This variation may generate discontinuous migration patterns. For instance, migrants may have predominantly settled in areas of West Germany, which are close to the demarcation line. Or, population from the far East may have migrated with a lower probability. Second, with respect to the development of the population gap, we have to consider regional planning policies. West Germany promoted settlements along the demarcation line with reduced taxes and investment subsidies in special economic zones (*Zonenrandgebiet*). These temporary place-based subsidies have had a persistent positive impact on economic density, which is driven by higher local public investment levels (von Ehrlich and Seidel, 2018). East Germany, in contrast, prevented settlements close to the demarcation line. There was no trespassing within 5 kilometers of the demarcation line and forced relocation of the population to impede escape towards the West.

Given these specific factors in a narrow corridor around the demarcation line, we explore now the generalizability of our results. Therefore, we replicate our analysis in different estimation samples comprising regions further away from the demarcation line. This procedure involves a trade-off between internal and external validity. While the identifying assumptions of a RDD are generally stronger, the further one moves away from the discontinuity (since differences in unobserved factors will increase), the generalizability is higher in broader samples. We start by defining different samples of non-bordering counties. In a second step, we use all counties in the nine states bordering the demarcation line, and then the full sample of German counties.

4.1 Distribution of Migrants within the West

So far, we have analyzed East-West-migration in regions with a maximum distance to the (former) demarcation line of 125 km. In some sense, we have focused on border crossings and disregard the spatial distribution of migrants within the (considerably larger) West. To account for the latter dimension, we define now four regions within the West, which differ in their distance to the demarcation line. We estimate for each region the difference in the log population relative to 1939. Our estimates are depicted in Figure 6 by six maps capturing the changes in population in 1943, 1944, 1945, 1946, 1950 and 1961. These

year-specific maps inform us about the influx of external migrants (mainly expellees), the internal East-West migration wave, and how the migrants dispersed over time *within* the West.

Between 1939 and 1943 the population had decreased throughout Germany. Starting with 1944/45 we observe that, relative to 1939, population increases. These are predominantly in the East and capture the influx of expellees from former German territory east of the Oder and Neisse Rivers. In 1946, we see the massive inner-German East-West migration crossing the newly established demarcation line. A comparison of the 1946 map with the ones for 1950 and 1961 provides two important insights. First, for border regions in the West the bulk of the East-West migration had taken place already in 1946. This is what we also find in our baseline estimates (see Column I of Table 1). Second, this analysis provides the additional insight that regions further in the West also experienced a significant influx of migrants, but later. This migration movement is not captured by our baseline estimates, since the dispersion of migrants to the West followed a different time pattern across regions. Notably, by 1961, the East-West migrants were equally dispersed across the West. This additional analysis solves to some degree the discrepancy between our baseline estimates and the finding put forward by the historical literature that there was substantial East-West migration between 1949 and 1961 in aggregate statistics.

4.2 Full Sample of all German Counties

Now, we expand our sample stepwise. First, we include all counties of the nine states bordering the demarcation line in our estimation sample (see Panel b of Figure 1). This increases our number of counties/observations substantially. In a second step, we use the entirety of all 411 German counties. In both samples, the concept of county pairs cannot be meaningfully maintained. To account for the spatial dimension, we substitute the county-pair fixed-effects in our estimation with controls for counties' latitude and longitude. In Columns (II) of Table 1, we apply this alternative estimation strategy to our baseline sample of counties that border the demarcation line. A comparison with our baseline results in Column (I) of Table 1 shows that we obtain very similar results based on this alternative specification forgoing the pairing of counties. Columns (III) and (IV) summarize then the respective results for the two larger samples. Across all samples, we find a very comparable development of the development of the East-West population gap. In line with our analysis in the previous section, we find for the overall sample a larger difference in the East-West Population gap between the year 1946 and 1950 (Column IV: $0.323 - 0.200 = 0.123$) as compared to the sample of bordering counties (Column II: $0.216 - 0.193 = 0.023$).[18]

[18]Note that the estimated East-West Population gap in the year 1946 (see columns (III) and (IV) of Table 1), is not visible in the raw data (see upper panels of Figure 2). It is important to control for the location of a county within East- and West-Germany. This accounts for the lag in the East-West

This set of estimations suggests that our results are not specific to the area along the demarcation line, but apply by and large to Germany as a whole.

4.3 Urban Sample

Both, the demarcation line and the Line of Contact run through mainly rural areas. We now check, whether the origin and development of the population gap is comparable in urban areas, where economic activity is most concentrated. For this analysis, we rely on the sample of cities from Redding and Sturm (2008a).[19] The lower-right panel of Figure 2 provides descriptive evidence. It plots the population development of cities in the East and West relative to the year 1939. Interestingly, with find again very comparable trends across the two groups for the period before WW2, and diverging trends thereafter.

Clearly, an important limitation of this sample definition is that other factors, like market access and geographic factors, might affect cities in the West differently than in the East. To mitigate these concerns, we estimate different models with varying bandwidths and covariates. While the quantitative results vary across models, all specifications confirm the general pattern found in our sample of counties along the demarcation line also for the urban sample (see Table 6).

5 Conclusions

We study the development of population levels east and west of the inner German border between 1900 and 2009. We show that the origin of the German East-West population difference was the advancing Soviet army and not, as generally believed, the socialist regime in the former GDR. In fact, population differences along the inner German border remained relatively stable during the GDR period compared to the period of the Soviet occupation and the period after reunification. These patterns are equally evident when including *all* counties in the estimation sample. In addition, we find evidence for strong positive selection in skills of these early East-West migrants. These findings challenge the dominant interpretation of this episode among contemporary historians and put the German division as a valid natural experiment for institutional differences into question.

Our second finding is the surprisingly sharp and increasing discontinuity of population levels along the former inner German border after reunification. We suggest persistent local productivity differences between West and East-Germany, possibly generated by agglomeration economies, as the reason for this pattern. This finding speaks to the policy debate on market integration through economic and political cooperation. Large productivity differences between integrating countries could lead to a large out-migration

migration to the Western part of West-Germany (as shown in the maps of Figure 6).

[19] Redding and Sturm (2008a) collected data for West- and East-German cities, but did not test differences between population development between the two parts of Germany.

of workers instead of an alignment of productivity levels even in geographically close regions.

An obvious limitation of our work is the focus on population levels as compared to more welfare-related measures like income or consumption. Population levels is one of only few comparable measures between the statistical agencies of West and East Germany. Future work could inquire our explanation of the continuously sharp discontinuity of population levels at the inner German border after reunification. We speculate that local productivity differences are at work, but fall short of showing direct evidence. This could be investigated with detailed firm or labor market data of the post-reunification period.

References

Abadie, A., Diamond, A. and Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, **59** (2), 495–510.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015). The Economics of Density: Evidence from the Berlin Wall. *Econometrica*, **83** (6), 2127–2189.

Alesina, A. and Fuchs-Schündeln, N. (2007). Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences. *American Economic Review*, **97** (4), 1507–1528.

Ariely, D., Garcia-Rada, X., Gödker, K., Hornuf, L. and Mann, H. (2019). The Impact of Two Different Economic Systems on Dishonesty. *European Journal of Political Economy*, **59**, 179–195.

Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949). Volkszählung — Tabellenteil. In *Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Großberlin*, Duncker & Humbolt.

Bauernschuster, S., Falck, O., Gold, R. and Heblich, S. (2012). The Shadows of the Socialist Past: Lack of Self-reliance Hinders Entrepreneurship. *European Journal of Political Economy*, **28** (4), 485–497.

— and Rainer, H. (2012). Political Regimes and the Family: How Sex-role Attitudes Continue to Differ in Reunified Germany. *Journal of Population Economics*, **25** (1), 5–27.

Becker, S. O., Mergele, L. and Woessmann, L. (2020). The Separation and Reunification of Germany: Rethinking a Natural Experiment Interpretation of the Enduring Effects of Communism. *Journal of Economic Perspectives*, **34** (2), 143–71.

Boltho, A., Carlin, W. and Scaramozzino, P. (2018). Why East Germany Did Not Become a New Mezzogiorno. *Journal of Comparative Economics*, **46** (1), 308–325.

Bondar, M. and Fuchs-Schündeln, N. (2022). *Good Bye Lenin Revisited: East-West Preferences Three Decades After Reunification*. Unpublished manuscript, Goethe University Frankfurt.

Borjas, G. J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, **77** (4), 531–553.

— (1999). The Economic Analysis of Immigration. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, *28*, North-Holland, pp. 1697–1760.

Bosker, M., Brakman, S., Garretsen, H. and Schramm, M. (2007). Looking for Multiple Equilibria when Geography Matters: German City Growth and the WWII Shock. *Journal of Urban Economics*, **61** (1), 152–169.

—, —, — and — (2008). A Century of Shocks: The Evolution of the German City Size Distribution 1925–1999. *Regional Science and Urban Economics*, **38** (4), 330–347.

Brakman, S., Garretsen, H. and Schramm, M. (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.

Braun, S. T., Kramer, A., Kvasnicka, M. and Meier, P. (2020). Local labor markets and the persistence of population shocks: evidence from West Germany, 1939–1970. *Journal of Economic Geography*, **21** (2), 231–260.

Brosig-Koch, J., Helbach, C., Ockenfels, A. and Weimann, J. (2011). Still Different after All these Years: Solidarity Behavior in East and West Germany. *Journal of Public Economics*, **95** (11), 1373–1376.

Burda, M. C. (1993). The Determinants of East-West German Migration: Some First Results. *European Economic Review*, **37** (2–3), 452–461.

—, Härdle, W., Müller, M. and Werwatz, A. (1998). Semiparametric Analysis of German East-West Migration Intentions: Facts and Theory. *Journal of Applied Econometrics*, **13** (5), 525–541.

Büttner, T. and Lutz, W. (1990). Estimating Fertility Responses to Policy Measures in the German Democratic Republic. *Population and Development Review*, **16** (3), 539–555.

Chiswick, B. R. (1978). The Effect of Americanization on the Earnings of Foreign-born Men. *Journal of Political Economy*, **86** (5), 897–921.

Dack, M. (2008). Crimes Committed by Soviet Soldiers Against German Civilians, 1944-1945: A Historiographical Analysis. *Journal of Military and Strategic Studies*, **10** (4).

Davis, D. R. and Weinstein, D. E. (2002). Bones, Bombs, and Break Points: The Geograohy of Economic Activity. *American Economic Review*, **92** (5), 1269–1289.

— and — (2008). A Search for Multiple Equilibria in Urban Industrial Structure. *Journal of Regional Science*, **48** (1), 29–65.

Eder, C. and Halla, M. (2016). *The Long-lasting Shadow of the Allied Occupation of Austria on its Spatial Equilibrium*. IZA Discussion Paper 10095, Institute for the Study of Labor (IZA), Bonn.

Fassmann, H. and Münz, R. (1994). European East-West Migration, 1945-1992. *International Migration Review*, **28** (3), 520–538.

Friehe, T. and Mechtel, M. (2014). Conspicuous Consumption and Political Regimes: Evidence from East and West Germany. *European Economic Review*, **67**, 62–81.

Fuchs-Schündeln, N. and Izem, R. (2012). Explaining the Low Labor Productivity in East Germany – A Spatial Analysis. *Journal of Comparative Economics*, **40** (1), 1–21.

— and Masella, P. (2016). Long-lasting Effects of Socialist Education. *The Review of Economics and Statistics*, **98** (3), 428–441.

— and Schündeln, M. (2009). Who Stays, Who Goes, Who Returns? East–West Migration within Germany since Reunification. *Economics of Transition*, **17** (4), 703–738.

Grundmann, S. (1998). *Bevölkerungsentwicklung in Ostdeutschland*. Opladen, Germany: Leske + Budrich.

Heiland, F. (2004). Trends in East-West German Migration from 1989 to 2002. *Demographic Research*, **11** (7), 173–194.

Heilig, G., Büttner, T. and Lutz, W. (1990). Germany's Population: Turbulent Past, Uncertain Future. *Population Bulletin*, **45** (4).

Henderson, J. V. (1974). The Types and Size of Cities. *American Economic Review*, **64** (4), 640–656.

Hunt, J. (2006). Staunching Emigration from East Germany: Age and the Determinants of Migration. *Journal of the European Economic Association*, **4** (5), 1014–1037.

Johr, B. (1992). Die Ereignisse in Zahlen. In H. Sander and B. Johr (eds.), *BeFreier und Befreiter*, Frankfurt am Main, Germany: Fischer Taschenbuch Verlag, pp. 46–73.

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

Lichter, A., Löffler, M. and Siegloch, S. (2021). The Long-Term Costs of Government Surveillance: Insights from Stasi Spying in East Germany. *Journal of the European Economic Association*, **19** (2), 741–789.

Miguel, E. and Roland, G. (2011). The Long-run Impact of Bombing Vietnam. *Journal of Development Economics*, **96** (1), 1–15.

Möhlmann, A. (2014). Persistence or Convergence? The East–West Tax-Morale Gap in Germany. *FinanzArchiv/Public Finance Analysis*, **70** (1), 3–30.

Peters, M. (2019). *Refugees and Endogenous Local Productivity – Evidence from Germany's Post-War Population Expulsions*. Unpublished manuscript, Princton University.

Rainer, H. and Siedler, T. (2009). Does Democracy Foster Trust? *Journal of Comparative Economics*, **37** (2), 251–269.

Redding, S. J. and Sturm, D. M. (2008a). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review*, **98** (5), 1766–1797.

— and — (2008b). The Costs of Remoteness: Evidence from German Division and Reunification: Dataset. *American Economic Review*.

Schumann, A. (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.

Stacy, W. E. (1984). *US Army Border Operations in Germany, 1945–1983*. Tech. rep., US Army Military History Office.

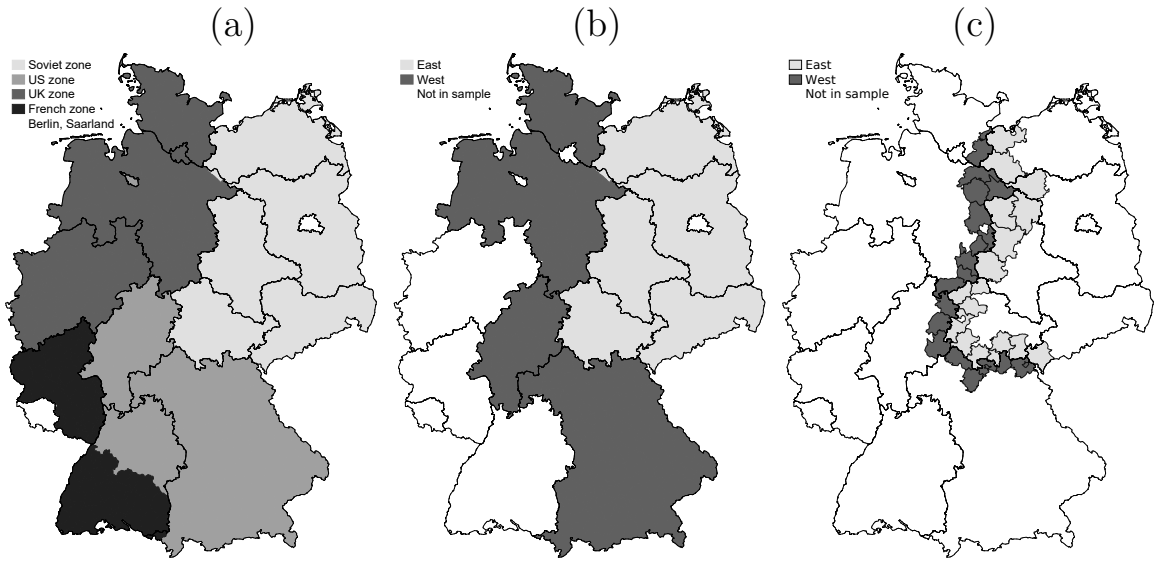
Uhlig, H. (2008). The Slow Decline of East Germany. *Journal of Comparative Economics*, **36** (4), 517–541.

von Ehrlich, M. and Seidel, T. (2018). The Persistent Effects of Place-Based Policy: Evidence from the West-German Zonenrandgebiet. *American Economic Journal: Economic Policy*, **10** (4), 344–374.

Witte, J. C. and Wagner, G. G. (1995). Declining Fertility in East Germany After Unification: A Demographic Response to Socioeconomic Change. *Population and Development Review*, **21** (2), 387–397.

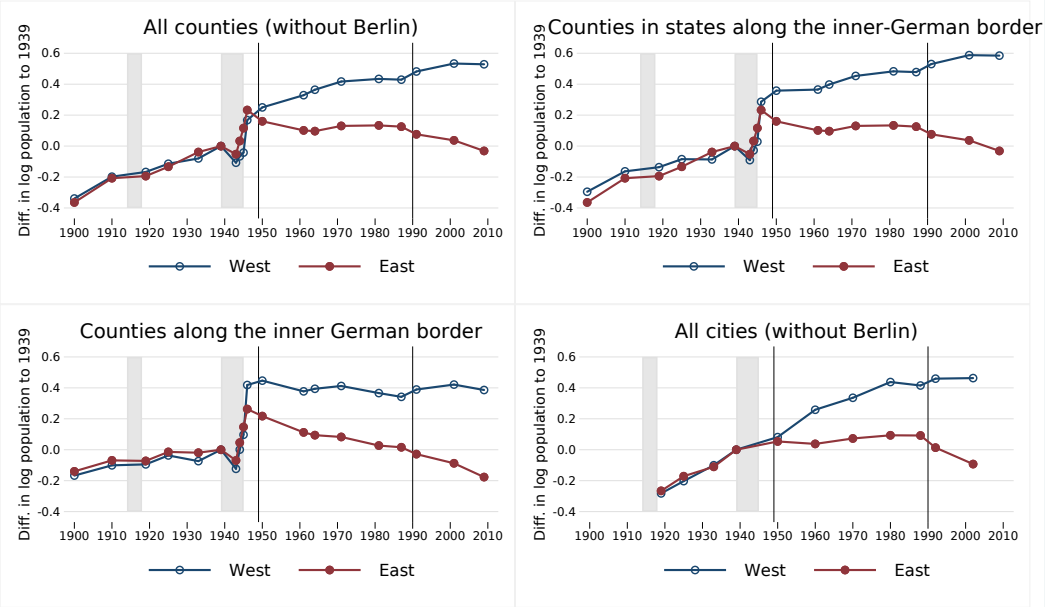
6 Figures (to be placed in article)

Figure 1: Zones of Allied-occupied Germany and Bordering States/Counties



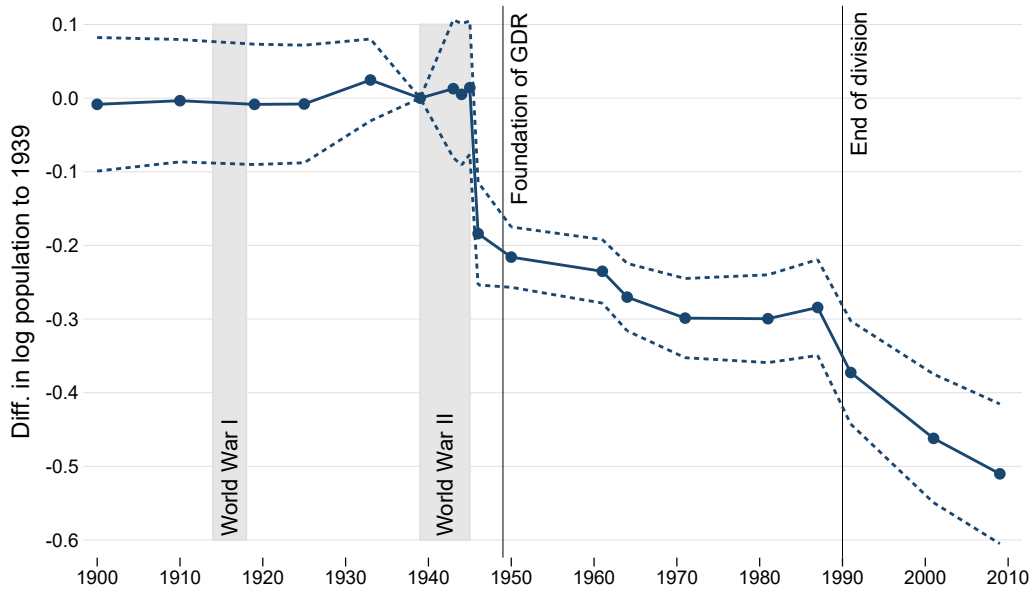
Notes: The left map shows the different zones of the *Allied-occupied Germany*, which existed between 1945 and 1949. The map in the middle highlights the bordering states (*Länder*) along the demarcation line dividing the American/British zones and the Soviet zone. The figure on the right shows the bordering counties (*Kreise*) along the demarcation line dividing the American/British zones and the Soviet zone. The demarcation line became the inner German border in 1949 dividing the newly established *German Democratic Republic* (GDR) in the East and the *Federal Republic of Germany* (FDR) in the West. After the reunification of Germany in 1990, this border became obsolete.

Figure 2: Population Development in East and West Germany



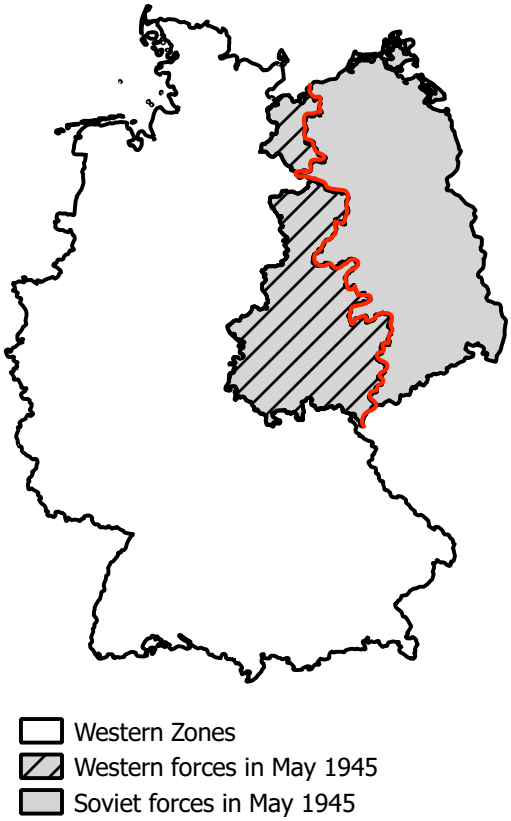
Notes: These figures show the development of population levels in the 'East' and 'West' for different sample definitions. In the period from 1945 to 1949, 'East' refers to the Soviet zone of occupation, and 'West' to the territory of the American, British and French zones of occupation. In 1949, the *German Democratic Republic* was established in the Soviet zone, while the *Federal Republic of Germany* was established in the three western zones. Grey bars mark World War I and II. The first vertical line indicates the foundation of the GDR, while the second vertical line indicates the reunification of Germany. The upper-left panel shows the population levels in all counties (except Berlin), the upper-right panel in all countries in states along the inner German border, and the lower-left panel in counties located at the border (i. e., along the demarcation line or the inner German border.) The lower-right panel shows population levels of all German cities (except Berlin). In the city-sample, there is no information for the years 1943, 1944, and 1945 available.

Figure 3: The Origin and Evolution of the East-West Population Gap along the Demarcation Line



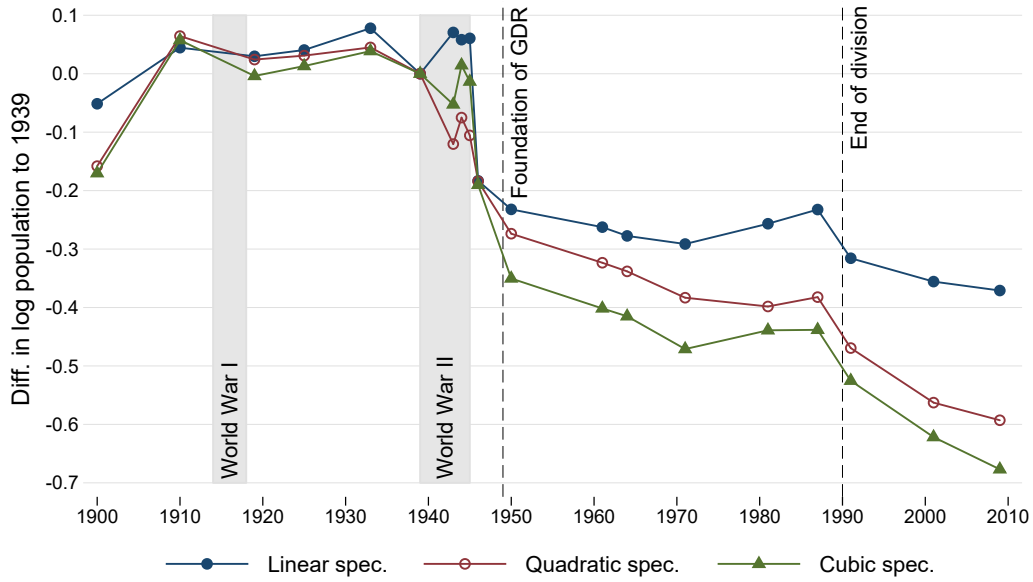
Notes: This figure shows the estimated East-West population gap in bordering counties along the demarcation line. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Dashed lines show 95% confidence intervals. Column (I) of Table 1 provides detailed estimation output.

Figure 4: The Line of Contact within Germany



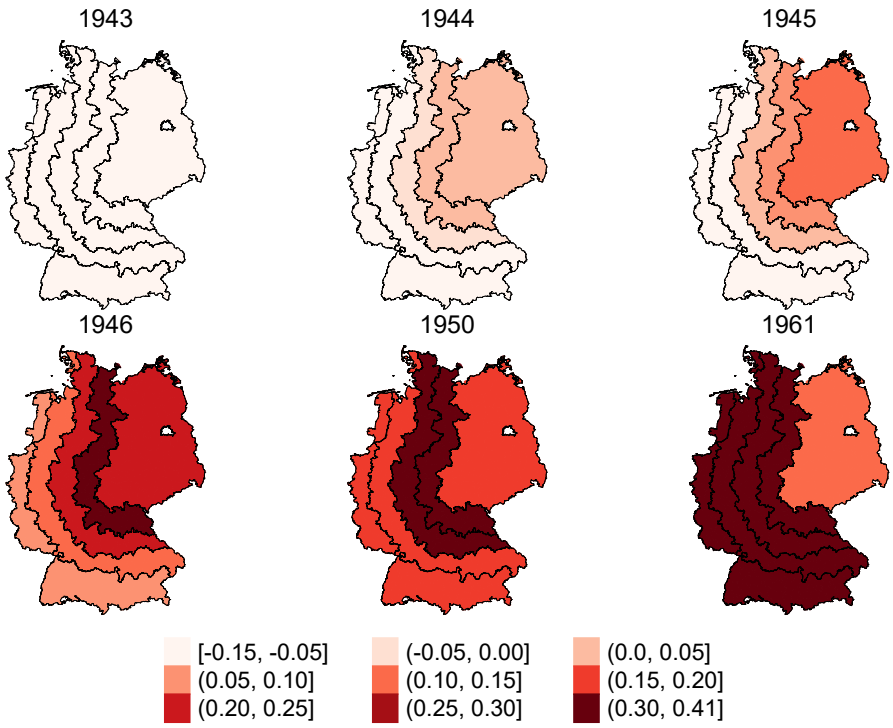
Notes: The so-called *Line of Contact* (printed in red) marks the farthest advance of Canadian, American, British and Soviet Armies into German controlled territory at the end of WW2. Notably, the Line of Contact is located East to the demarcation line (printed in black) within the later Soviet zone. The gray area shows the Soviet occupation zone after July 1945. The territory marked by the diagonally shaded area was conquered by Western Allies, but became part of the Soviet occupation zone 2 months later. The British and US troops withdrew from this territory in July 1945 and passed it on to the Soviets.

Figure 5: The Origin and Evolution of the East-West Population Gap along the Demarcation Line – One-dimensional RDD Estimates



Notes: This figure summarize the estimated East-West population gap based on a one-dimensional RDD approach with a 40 KM band around the demarcation line for different points in time. Different markers indicate estimates of specifications with a linear, quadratic, and cubic zone-specific distance functions around the demarcation line. None of the estimates of 1945 or before are statistically significantly different from zero and all of the post-WW2 are statistically significant different from zero at the 1% level (except the point estimates for 1946 in the linear and quadratic specifications, which are statistically significant at the 5% level).

Figure 6: Distribution of Migrants within the West



Notes: This figure shows the estimated difference in log population to 1939 for East Germany and West Germany (which is split into four distance groups from the demarcation line) at different points in time (1943, 1944, 1945, 1946, 1950 and 1961). Darker colors indicate higher population growth. Detailed estimation output is available upon request.

7 Tables (to be placed in article)

Table 1: The Origin and Evolution of the East-West Population Gap

	Log population			
Estimation method:	RDD-DiD with year-specific pair fixed-effects	DID with year-specific spatial control variables		
Sample definition:	Bordering counties	Bordering counties	All counties in bordering states	All counties
	(I)	(II)	(III)	(IV)
Pre-WWII differences				
1900 × East	−0.008	−0.050	−0.131	−0.129
	(0.045)	(0.065)	(0.096)	(0.091)
1910 × East	−0.003	−0.026	−0.078	−0.107
	(0.042)	(0.060)	(0.078)	(0.073)
1919 × East	−0.008	−0.031	−0.086	−0.110
	(0.041)	(0.056)	(0.079)	(0.073)
1925 × East	−0.008	−0.023	−0.064	−0.083
	(0.040)	(0.053)	(0.078)	(0.072)
1933 × East	0.025	0.011	0.014	0.012
	(0.028)	(0.052)	(0.054)	(0.050)
Base-year (1939) differences				
East	0.592***			
	(0.087)			
Differences during WWII				
1943 × East	0.013	0.032	−0.021	−0.026
	(0.047)	(0.085)	(0.058)	(0.052)
1944 × East	0.005	0.047	−0.009	−0.030
	(0.048)	(0.087)	(0.062)	(0.058)
1945 × East	0.014	0.041	0.005	−0.069
	(0.045)	(0.084)	(0.068)	(0.073)
Differences during division				
1946 × East	−0.184***	−0.193***	−0.150**	−0.200***
	(0.035)	(0.054)	(0.074)	(0.069)
1950 × East	−0.216***	−0.216***	−0.287***	−0.323***
	(0.021)	(0.025)	(0.032)	(0.035)
1961 × East	−0.235***	−0.232***	−0.338***	−0.351***
	(0.022)	(0.037)	(0.033)	(0.033)
1964 × East	−0.270***	−0.256***	−0.364***	−0.379***
	(0.023)	(0.037)	(0.034)	(0.035)
1971 × East	−0.299***	−0.280***	−0.396***	−0.417***
	(0.027)	(0.041)	(0.046)	(0.046)
1981 × East	−0.300***	−0.263***	−0.482***	−0.501***
	(0.030)	(0.040)	(0.063)	(0.059)
1987 × East	−0.284***	−0.252***	−0.485***	−0.505***
	(0.033)	(0.044)	(0.065)	(0.061)
Differences after reunification				
1991 × East	−0.373***	−0.341***	−0.591***	−0.614***
	(0.035)	(0.047)	(0.068)	(0.064)
2001 × East	−0.462***	−0.422***	−0.642***	−0.677***
	(0.044)	(0.064)	(0.081)	(0.077)
2009 × East	−0.510***	−0.460***	−0.702***	−0.740***
	(0.048)	(0.070)	(0.083)	(0.079)
Pair-Year FE	Yes			
County & Year FE		Yes	Yes	Yes
Flexible latitude/longitude controls		Yes	Yes	Yes
Number of observations	1,292	684	5,111	7,767
Number of pairs	34			
Number of unique counties	36	36	269	411
Number of periods	19	19	19	19
Mean of dep. var.	11.61	11.65	11.53	11.64

Notes: This table summarizes estimation results based on German county-level data. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach (column I) or a difference-in-differences (DiD) approach (columns II–IV). Flexible latitude/longitude controls include the latitude of the centroid of a county, the longitude and its interaction, each interacted with year dummies. Clustered standard errors (at the county level within a pair (column I) or at the county level (columns II–IV)) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Table 2: The Origin and Evolution of the East-West Population Gap along the Line of Contact

	Log population				
	All counties along Line of Contact	Dessau-Roßlau dropped	Magdeburg dropped	Nordsachsen dropped	All three dropped
	(I)	(II)	(III)	(IV)	(V)
Pre-WWII differences					
1900 × East of Line of Contact	−0.042	0.003	−0.029	−0.057	0.016
	(0.102)	(0.095)	(0.110)	(0.119)	(0.120)
1910 × East of Line of Contact	−0.052	−0.008	−0.023	−0.080	0.008
	(0.106)	(0.100)	(0.112)	(0.121)	(0.122)
1919 × East of Line of Contact	−0.044	0.001	−0.018	−0.069	0.017
	(0.106)	(0.099)	(0.112)	(0.120)	(0.120)
1925 × East of Line of Contact	−0.042	0.003	−0.013	−0.064	0.025
	(0.106)	(0.099)	(0.111)	(0.120)	(0.120)
1933 × East of Line of Contact	−0.043	−0.030	−0.019	−0.066	−0.021
	(0.040)	(0.039)	(0.039)	(0.043)	(0.040)
Base-year (1939) differences					
East of Line of Contact	−0.045	−0.037	−0.081	0.012	−0.015
	(0.093)	(0.096)	(0.096)	(0.093)	(0.101)
Differences during WWII					
1943 × East of Line of Contact	−0.028	−0.022	−0.013	−0.035	−0.010
	(0.024)	(0.024)	(0.023)	(0.026)	(0.025)
1944 × East of Line of Contact	−0.032	−0.023	−0.003	−0.047	−0.002
	(0.033)	(0.032)	(0.027)	(0.034)	(0.025)
1945 × East of Line of Contact	−0.066	−0.051	−0.013	−0.097*	−0.018
	(0.050)	(0.049)	(0.036)	(0.051)	(0.027)
Differences during division					
1946 × East of Line of Contact	−0.128***	−0.128***	−0.099**	−0.161***	−0.129***
	(0.043)	(0.045)	(0.041)	(0.045)	(0.044)
1950 × East of Line of Contact	−0.090**	−0.073*	−0.055	−0.113**	−0.050
	(0.041)	(0.038)	(0.035)	(0.045)	(0.031)
1961 × East of Line of Contact	−0.075**	−0.061**	−0.050*	−0.092***	−0.044**
	(0.031)	(0.028)	(0.027)	(0.033)	(0.022)
1964 × East of Line of Contact	−0.058**	−0.044*	−0.036	−0.072**	−0.028
	(0.029)	(0.026)	(0.026)	(0.031)	(0.020)
1971 × East of Line of Contact	−0.015	−0.001	0.007	−0.026	0.019
	(0.049)	(0.049)	(0.050)	(0.056)	(0.057)
1981 × East of Line of Contact	−0.043	−0.033	−0.030	−0.063**	−0.038
	(0.029)	(0.028)	(0.029)	(0.029)	(0.029)
1987 × East of Line of Contact	−0.035	−0.026	−0.024	−0.059*	−0.038
	(0.030)	(0.030)	(0.031)	(0.030)	(0.032)
Differences after reunification					
1991 × East of Line of Contact	−0.041	−0.032	−0.031	−0.070**	−0.050
	(0.032)	(0.032)	(0.034)	(0.032)	(0.034)
2001 × East of Line of Contact	−0.051	−0.038	−0.027	−0.089*	−0.048
	(0.045)	(0.045)	(0.045)	(0.045)	(0.045)
2009 × East of Line of Contact	−0.060	−0.046	−0.043	−0.095*	−0.061
	(0.048)	(0.047)	(0.050)	(0.051)	(0.054)
Pair-Year FE	Yes	Yes	Yes	Yes	Yes
Number of observations	1,102	1,064	1,026	950	836
Number of pairs	29	28	27	25	22
Number of unique counties	29	28	28	26	24
Number of periods	19	19	19	19	19
R-squared	0.80	0.82	0.80	0.81	0.82
Mean of dep. var.	12.16	12.17	12.14	12.11	12.08

Notes: This table summarizes estimation results based on German county-level data. In columns (II)-(V), all pairs comprising a county mentioned in the column header are dropped. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Clustered standard errors (at the county level within a pair) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Table 3: The Composition of the East-West Migration Wave in 1945/46: Sex and Skill Distribution

	Share of		
	male population	workers in agriculture	workers in manufacturing
	(I)	(II)	(III)
Base-year (1939) differences			
East	−0.028**	−0.055**	0.037**
	(0.014)	(0.025)	(0.018)
Differences during division			
1946 × East	0.012		
	(0.014)		
1950 × East	−0.018	0.085***	−0.048***
	(0.016)	(0.013)	(0.011)
1961 × East	0.015	0.082***	−0.104***
	(0.014)	(0.023)	(0.019)
Pair-Year FE	Yes	Yes	Yes
Number of observations	272	194	194
Number of pairs	34	34	34
Number of unique counties	36	36	36
Number of periods	4	3	3
R-squared	0.68	0.70	0.81
Mean of dep. var.	0.47	0.32	0.37

Notes: This table summarizes estimation results based on German county-level data. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Clustered standard errors (at the county level within a pair) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Table 4: The Composition of the East-West Migration Wave in 1945/46: Region of Origin

	Ratio of pop. of group g in t to total pop. in 1939				
Group:	Natives	Expellees	Foreigners	Unknown	All
	(I)	(II)	(III)	(IV)	(V)
Differences during division					
1946 × East	−0.124***	−0.003	0.014	0.001	−0.112*
	(0.030)	(0.040)	(0.034)	(0.001)	(0.056)
1950 × East	−0.185***	−0.025	−0.027	0.002***	−0.235***
	(0.024)	(0.041)	(0.027)	(0.001)	(0.045)
Pair-Year FE	Yes	Yes	Yes	Yes	Yes
Number of observations	48	48	48	48	48
Number of pairs	8	8	8	8	8
Number of unique states	9	9	9	9	9
Number of periods	3	3	3	3	3
R-squared	0.85	0.92	0.77	0.80	0.93
Mean of dep. var.	0.98	0.15	0.08	0.00	1.20

Notes: This table summarizes estimation results based on German state-level data, split by region of residence in 1939 of the population. The census of 1950 in the GDR reports these numbers only for the population born before September 1, 1939. We correct the respective data points by the average share of the population of age 11 years or older in the entire GDR (84.5%). The dependent variable is equal to the ratio of population group g in year t to the total population in 1939. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Clustered standard errors (at the state level within a pair) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Table 5: The Origin and Evolution of the East-West Population Gap – Different Standard Errors

	Log population						
	Base	Different standard error clustering				Synthetic control	
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)
Pre-WWII differences							
1900 × East	−0.008	−0.008	−0.008	−0.008	−0.008	0.044	−0.023
	(0.045)	(0.047)	(0.065)	(0.075)	(0.060)	(0.072)	(0.042)
1910 × East	−0.003	−0.003	−0.003	−0.003	−0.003	0.034	−0.003
	(0.042)	(0.043)	(0.059)	(0.061)	(0.061)	(0.066)	(0.037)
1919 × East	−0.008	−0.008	−0.008	−0.008	−0.008	0.027	−0.008
	(0.041)	(0.042)	(0.058)	(0.060)	(0.060)	(0.065)	(0.035)
1925 × East	−0.008	−0.008	−0.008	−0.008	−0.008	0.027	−0.007
	(0.040)	(0.040)	(0.057)	(0.059)	(0.056)	(0.064)	(0.034)
1933 × East	0.025	0.025	0.025	0.025	0.025	0.048	0.025
	(0.028)	(0.034)	(0.040)	(0.036)	(0.059)	(0.038)	(0.036)
Base-year (1939) differences							
East	0.592***	0.592***	0.592***	0.592***	0.592***	0.533***	0.650***
	(0.087)	(0.109)	(0.123)	(0.150)	(0.162)	(0.118)	(0.118)
Differences during WWII							
1944 × East	0.005	0.005	0.005	0.005	0.005	0.038	0.008
	(0.048)	(0.058)	(0.069)	(0.070)	(0.094)	(0.074)	(0.048)
Differences during division							
1946 × East	−0.184***	−0.184***	−0.184***	−0.184**	−0.184***	−0.145***	−0.201***
	(0.035)	(0.045)	(0.050)	(0.072)	(0.056)	(0.045)	(0.048)
1950 × East	−0.216***	−0.216***	−0.216***	−0.216***	−0.216***	−0.210***	−0.220***
	(0.021)	(0.024)	(0.029)	(0.032)	(0.036)	(0.029)	(0.028)
1961 × East	−0.235***	−0.235***	−0.235***	−0.235***	−0.235***	−0.243***	−0.230***
	(0.022)	(0.026)	(0.031)	(0.040)	(0.034)	(0.027)	(0.027)
1964 × East	−0.270***	−0.270***	−0.270***	−0.270***	−0.270***	−0.281***	−0.260***
	(0.023)	(0.028)	(0.033)	(0.045)	(0.035)	(0.028)	(0.028)
1971 × East	−0.299***	−0.299***	−0.299***	−0.299***	−0.299***	−0.313***	−0.282***
	(0.027)	(0.033)	(0.038)	(0.056)	(0.038)	(0.032)	(0.031)
1981 × East	−0.300***	−0.300***	−0.300***	−0.300***	−0.300***	−0.319***	−0.283***
	(0.030)	(0.035)	(0.043)	(0.060)	(0.039)	(0.037)	(0.032)
1987 × East	−0.284***	−0.284***	−0.284***	−0.284***	−0.284***	−0.303***	−0.268***
	(0.033)	(0.040)	(0.047)	(0.066)	(0.047)	(0.039)	(0.036)
Differences after reunification							
1991 × East	−0.373***	−0.373***	−0.373***	−0.373***	−0.373***	−0.390***	−0.358***
	(0.035)	(0.043)	(0.050)	(0.071)	(0.051)	(0.042)	(0.038)
2001 × East	−0.462***	−0.462***	−0.462***	−0.462***	−0.462***	−0.478***	−0.441***
	(0.044)	(0.053)	(0.062)	(0.082)	(0.068)	(0.054)	(0.045)
2009 × East	−0.510***	−0.510***	−0.510***	−0.510***	−0.510***	−0.530***	−0.486***
	(0.048)	(0.057)	(0.068)	(0.090)	(0.073)	(0.059)	(0.048)
Pair-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1,292	1,292	1,292	1,292	1,292	722	646
Number of pairs	34	34	34	34	34	19	17
Number of unique counties	36	36	36	36	36	38	34
Number of periods	19	19	19	19	19	19	19
R-squared	0.73	0.73	0.73	0.73	0.73	0.72	0.77
Mean of dep. var.	11.61	11.61	11.61	11.61	11.61	11.67	11.60

Notes: This table summarizes estimation results based on German county-level data. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively. Clustered standard errors (different levels) are in parentheses below: (I) baseline approach (see column (I) in Table 1), (II) clustered within each county, irrespective of the pair of the county, (III) clustered within each pair, (IV) clustered within each county in the East plus all bordering counties in the West, (V) clustered within each county in the West plus all bordering counties in the East, (VI) dataset is transformed so that for each county in the East, there is a synthetic control county in the West. The synthetic control county is the mean of all bordering counties in the West, (VII) dataset is transformed so that for each county in the West, there is a synthetic control county in the East. The synthetic control county is the mean of all bordering counties in the East.

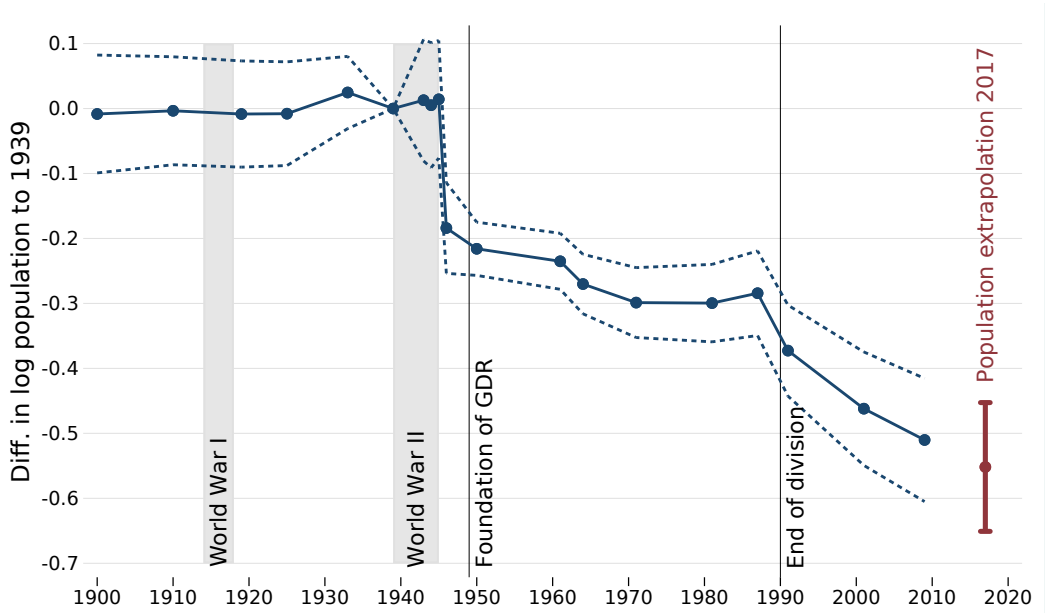
Table 6: The Origin and Evolution of the East-West Population Gap in an Urban Sample

	Log population				
	(I)	(II)	(III)	(IV)	(V)
Pre-WWII differences					
1919 × East	0.017	0.017	0.003	0.117*	−0.056
	(0.026)	(0.048)	(0.033)	(0.068)	(0.098)
1925 × East	0.032	0.067*	0.043	0.157***	−0.004
	(0.022)	(0.033)	(0.027)	(0.053)	(0.081)
1933 × East	−0.010	0.048	0.034	0.111**	−0.016
	(0.019)	(0.030)	(0.025)	(0.047)	(0.070)
Differences during division					
1950 × East	−0.028	−0.172*	−0.068	−0.284***	−0.550***
	(0.031)	(0.086)	(0.054)	(0.102)	(0.137)
1960 × East	−0.221***	−0.257***	−0.211***	−0.320***	−0.506***
	(0.029)	(0.075)	(0.043)	(0.094)	(0.142)
1970 × East	−0.263***	−0.249***	−0.226***	−0.317***	−0.470***
	(0.034)	(0.085)	(0.051)	(0.111)	(0.171)
1980 × East	−0.344***	−0.290**	−0.319***	−0.332**	−0.595***
	(0.045)	(0.112)	(0.071)	(0.145)	(0.216)
1988 × East	−0.323***	−0.246**	−0.294***	−0.265*	−0.544**
	(0.050)	(0.112)	(0.073)	(0.145)	(0.213)
Differences after reunification					
1992 × East	−0.446***	−0.362***	−0.416***	−0.380**	−0.677***
	(0.051)	(0.116)	(0.075)	(0.150)	(0.221)
2002 × East	−0.556***	−0.466***	−0.528***	−0.436***	−0.783***
	(0.050)	(0.115)	(0.076)	(0.150)	(0.222)
City FE	Yes	Yes	Yes	Yes	Yes
Distance cutoff in KM		50	100	100	100
Linear year-spec. distance				Yes	Yes
Quadratic year-spec. distance					Yes
Number of observations	1,936	308	759	759	759
Number of unique cities	176	28	69	69	69
Number of periods	11	11	11	11	11
R-squared	0.97	0.97	0.97	0.97	0.97
Mean of dep. var.	11.27	11.12	11.08	11.08	11.08

Notes: This table summarizes estimation results based on German city data from Redding and Sturm (2008b). The city of Berlin is excluded, since the demarcation disunited the city. The dependent variable is equal to the log of population. Method of estimation is a one-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Clustered standard errors (at the city level) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Appendix A: Additional Figures and Tables

Figure A.1: The Origin and Evolution of the East-West Population Gap along the Demarcation Line – Longer Series



Notes: This figure is equivalent to Figure 3, but is based on an estimation sample with a longer series. It uses in addition data from the the so-called coordinated population projection for the year 2017. This data is provided by the Federal Statistical Office of Germany (*Statistisches Bundesamt*).

Table A.1: Relevant Studies in Contemporary German History

Gehler, M. (2010). *Deutschland: Von der Teilung zur Einigung. 1945 bis heute.* Wien – Köln – Weimar: Böhlau Verlag.

Horn, K. (1970). *Die Berlin-Krise 1958/61. Zur Funktion der Krise in der internationalen Politik.* No. 6 in Modelle für den politischen und sozialwissenschaftlichen Unterricht, Frankfurt a. M.: Europäische Verlagsanstalt.

Koch, H. R. (1986). Die Massen-Migration aus der SBZ und der DDR. *Zeitschrift für Soziologie*, **15** (1), 37–40.

Lemke, M. (2011). *Vor der Mauer. Berlin in der Ost-West-Konkurrenz 1948 bis 1961.* No. 48 in Zeithistorische Studien, Wien – Köln – Weimar: Böhlau Verlag.

Mirrow, J. (2004). *Die Geschichte des deutschen Volkes. Von den Anfängen bis zur Gegenwart: Die Deutschen seit dem Zweiten Weltkrieg.* Band 4, Gernsbach: Casimir Katz Verlag.

Ploetz, K. J. (2002). *Der große Ploetz: Die Daten-Enzyklopädie der Weltgeschichte. Daten, Fakten, Zusammenhänge.* 33. neu bearb. Auflage, Köln.

Pötzsch, H. and Haider, W. (2015). *Deutsche Geschichte von 1945 bis zur Gegenwart: Die Entwicklung der beiden deutschen Staaten und das vereinte Deutschland.* München: Lau Verlag.

Rühle, J. and Holzweissig, G. (1981). *13. August 1961. Die Mauer von Berlin.* Köln: Edition Deutschland Archiv im Verlag Wissenschaft und Politik.

Schwab-Felisch, H. (1953). Die Flucht vor der Angst. In G. Birkenfeld, M. Dor, R. Federmann, D. Levine, J. Leblond and H. Schwab-Felisch (eds.), *Sprung in die Freiheit. Berichte über die Ursachen, Begleitumstände und Folgen der Massenflucht au der Sowjetischen Besatzungszone Deutschlands*, Ulm-Donau: Knorr & Hirth Verlag, pp. 15–37.

Steinert, J. (1995). Die große Flucht und die Jahre danach. Flüchtlinge und Vertriebene in den vier Besatzungszonen. In H. Volksmann (ed.), *Ende des dritten Reiches – Ende des Zweiten Weltkrieges. Eine perspektivische Rückschau*, München: Piper Verlag, pp. 557–579.

Storbeck, D. (1963). Flucht oder Wanderung? Eine Rückschau auf Motive, Folgen und Beurteilung der Bevölkerungsabwanderung aus Mitteldeutschland seit dem Kriege. *Soziale Welt*, **14** (2), 153–171.

Streibel, R. (1994). Vorwort. In R. Streibel (ed.), *Flucht und Vertreibung. Zwischen Aufrechnung und Verdrängung*, Wien: Picus Verlag, pp. 9–19.

Ueberschär, G. R. (1994). Die Vertreibung der Deutschen Bevölkerung aus dem Osten und die Alliierten Grundsätze von der 'Besseren Welt'. In R. Streibel (ed.), *Flucht und Vertreibung*, Wien: Picus Verlag, pp. 20–41.

van Melis, D. and Bispinck, H. (eds.) (2006). *'Republikflucht': Flucht und Abwanderung aus der SBZ/DDR 1945 bis 1961.* Schriftenreihe der Vierteljahreshefte für Zeitgeschichte, München: Oldenbourg Verlag.

Zentrum für Zeithistorische Forschung e.V.,Bundeszentrale für politische Bildung and Deutschlandradio (). Chronik der Mauer. Fluchtbewegung aus der DDR und dem Ostsektor von Berlin 1949-1961. (abgerufen am 28.02.2019).

Table A.2: Descriptive Statistics

	Both zones		West	East
	Mean	Std. dev.	Mean	Mean
All counties (without Berlin) (N=411)				
Population				
in 1900	100,135	102,306	89,913	138,767
in 1939	135,095	143,951	123,987	177,071
in 1950	160,474	138,789	149,834	200,682
in 2009	190,656	164,356	201,298	150,437
Counties in bordering states (N=269)				
Population				
in 1900	91,543	85,084	69,350	138,767
in 1939	118,740	117,839	91,327	177,071
in 1950	149,363	118,836	125,246	200,682
in 2009	157,140	127,878	160,290	150,437
Bordering counties (N=36)				
Population				
in 1900	99,828	60,613	70,490	132,617
in 1939	117,415	71,795	85,542	153,037
in 1950	158,522	75,316	133,400	186,599
in 2009	126,936	54,756	129,730	123,814
Cities (N=176)				
Population				
in 1919	94,386	149,854	104,781	70,902
in 1939	123,384	193,736	138,606	88,995
in 1950	122,810	178,234	138,741	86,819
in 2002	151,093	205,428	183,016	78,972
Counties within 40KM of demarcation line (N=58)				
Population				
in 1900	87,939	66,478	67,119	117,434
in 1939	107,532	77,007	85,419	138,858
in 1950	140,733	76,515	125,698	162,033
in 2009	121,269	60,046	126,343	114,082

Notes: German city data comes from Redding and Sturm (2008b). All other data are from collections and calculations of the authors as described in Section 2.1.

Table A.3: Date Sources of German Population Data at the County Level

Variable	Year	Entity	Source
Population	1900	German Empire	"Volkszählung 1900 1. Dezember 1900 im Deutschen Reich." Statistisches Reichsamt.
Population	1910	German Empire	"Vorläufige Ergebnisse der Volkszählung im Deutschen Reich vom 16. Juni 1925." Statistisches Reichsamt.
Population	1919	German Empire	"Vorläufige Ergebnisse der Volkszählung im Deutschen Reich vom 16. Juni 1925." Statistisches Reichsamt.
Population	1925	German Empire	"Vorläufige Ergebnisse der Volkszählung im Deutschen Reich vom 16. Juni 1925." Statistisches Reichsamt.
Population	1933	German Empire	"Ergebnisse der Volks-, Berufsund landwirtschaftlichen Betriebszählung 1939." Statistik des Deutschen Reichs.
Male population; employment: total, agriculture, manufacturing	1939	German Empire	"Ergebnisse der Volks-, Berufsund landwirtschaftlichen Betriebszählung 1939." Statistik des Deutschen Reichs.
Population	1939	FDR	"Statistisches Jahrbuch für die Bundesrepublik Deutschland 1962." Statistisches Bundesamt. Wiesbaden.
Population	1939	GDR	"Statistisches Jahrbuch der Deutschen Demokratischen Republik 1965." 10. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Population	1943–45	German Empire	"Statistische Berichte." Arb.-Nr. VIII/19/1. Statistisches Bundesamt. Wiesbaden. (based on food stamps for the periods 8.2.-7.3.1943, 1.2.-5.3.1944, and 11.12.1944-7.1.1945)
Population: total, male	1946	Allied-occupied Germany	"Volks- und Berufszählung vom 29. Oktober 1946 in den vier Besatzungszonen und Groß-Berlin." Ausschuß der Deutschen Statistiker für die Volks- und Berufszählung 1946.
Population: total, male	1950	FDR	"Statistisches Jahrbuch für die Bundesrepublik Deutschland 1962." Statistisches Bundesamt. Wiesbaden.
Employment: total, agriculture, manufacturing	1950	FDR	"Kreisdaten (Volkszählungen 1950-1987). GESIS Datenarchiv, Köln.
Population: total, male	1950	GDR	"Statistisches Jahrbuch der Deutschen Demokratischen Republik 1965." 10. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Employment: total, agriculture, manufacturing	1950	GDR	"Sonderreihe mit Beiträgen für das Gebiet der ehemaligen DDR." 1994. Heft 15. Statistisches Bundesamt. Wiesbaden.
Population: total, male	1961	FDR	"Statistisches Jahrbuch für die Bundesrepublik Deutschland 1962." Statistisches Bundesamt. Wiesbaden.
Employment: total, agriculture, manufacturing	1961	FDR	"Regionaldaten VZ 1961 (Kreise)." GESIS Datenarchiv, Köln.
Population: total, male	1961	GDR	"Statistisches Jahrbuch der Deutschen Demokratischen Republik 1963." 8. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Population	1964	FDR	"Kreisbericht 1964." Statistisches Bundesamt. Wiesbaden.
Population	1964	GDR	"Statistisches Jahrbuch der Deutschen Demokratischen Republik 1965." 10. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Employment: total, agriculture, manufacturing	1964	GDR	"Sonderreihe mit Beiträgen für das Gebiet der ehemaligen DDR." 1994. Heft 15. Statistisches Bundesamt. Wiesbaden.
Population	1971	FDR	"Statistisches Jahrbuch für die Bundesrepublik Deutschland 1972." Statistisches Bundesamt. Wiesbaden.
Population	1971	GDR	"Statistisches Jahrbuch der Deutschen Demokratischen Republik 1974." 19. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Population	1981	FDR	"Statistisches Jahrbuch für die Bundesrepublik Deutschland 1982." Statistisches Bundesamt. Wiesbaden.
Population	1981	GDR	"Statistisches Jahrbuch 1985 der Deutschen Demokratischen Republik." 30. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Population	1987	FDR	"Kreisbericht 1987." Received from the Statistisches Bundesamt
Population	1987	GDR	"Statistisches Jahrbuch 1988 der Deutschen Demokratischen Republik." 33. Jahrgang. Staatliche Zentralverwaltung für Statistik.
Population	1991	Germany	"Kreisbericht 1991." Received from the Statistisches Bundesamt. Wiesbaden.
Population	2001	Germany	"Kreisbericht 2001." Received from the Statistisches Bundesamt. Wiesbaden.
Population	2009	Germany	"Kreisbericht 2009." Received from the Statistisches Bundesamt. Wiesbaden.
Population	2017	Germany	"Bevölkerungsfortschreibung auf Grundlage des Zensus 2011." Statistisches Bundesamt. Wiesbaden. (`www-genesis.destatis.de/genesis/online/link/tabellen/12411*`)
Historic GIS-shapefiles			`http://www.censusmosaic.org`

Table A.4: The Origin and Evolution of the East-West Population Gap – Combinded Difference-in-Differences Approach

	Log population		
	All	Bordering states	
	states	rural and urban	rural only
	(I)	(II)	(III)
1900 × East of inner-German border	−0.096	−0.113	−0.010
	(0.107)	(0.103)	(0.035)
1910 × East of inner-German border	−0.090	−0.083	−0.026
	(0.089)	(0.084)	(0.037)
1919 × East of inner-German border	−0.109	−0.097	−0.046
	(0.088)	(0.083)	(0.036)
1925 × East of inner-German border	−0.093	−0.080	−0.033
	(0.089)	(0.083)	(0.035)
1933 × East of inner-German border	0.045	0.044	0.017
	(0.042)	(0.046)	(0.032)
1943 × East of inner-German border	−0.012	−0.004	−0.035
	(0.048)	(0.052)	(0.038)
1944 × East of inner-German border	0.001	0.011	−0.029
	(0.053)	(0.056)	(0.039)
1945 × East of inner-German border	−0.009	0.029	−0.025
	(0.064)	(0.062)	(0.040)
1946 × East of inner-German border	−0.134**	−0.116*	−0.197***
	(0.058)	(0.062)	(0.027)
1950 × East of inner-German border	−0.277***	−0.255***	−0.266***
	(0.025)	(0.025)	(0.020)
1961 × East of inner-German border	−0.299***	−0.280***	−0.286***
	(0.026)	(0.028)	(0.030)
1964 × East of inner-German border	−0.331***	−0.308***	−0.324***
	(0.028)	(0.030)	(0.032)
1971 × East of inner-German border	−0.368***	−0.336***	−0.361***
	(0.037)	(0.040)	(0.042)
1981 × East of inner-German border	−0.440***	−0.409***	−0.437***
	(0.046)	(0.051)	(0.050)
1987 × East of inner-German border	−0.446***	−0.415***	−0.457***
	(0.049)	(0.053)	(0.051)
1991 × East of inner-German border	−0.549***	−0.516***	−0.555***
	(0.051)	(0.055)	(0.054)
2001 × East of inner-German border	−0.622***	−0.578***	−0.653***
	(0.065)	(0.070)	(0.067)
2009 × East of inner-German border	−0.692***	−0.644***	−0.735***
	(0.067)	(0.072)	(0.068)
1900 × East of Line of Contact	0.004	0.010	−0.161***
	(0.147)	(0.173)	(0.056)
1910 × East of Line of Contact	0.006	0.044	−0.056
	(0.146)	(0.169)	(0.051)
1919 × East of Line of Contact	0.028	0.075	−0.015
	(0.147)	(0.169)	(0.050)
1925 × East of Line of Contact	0.029	0.078	−0.008
	(0.146)	(0.168)	(0.048)
1933 × East of Line of Contact	0.019	0.014	−0.005
	(0.067)	(0.062)	(0.040)
1943 × East of Line of Contact	0.010	0.028	0.047
	(0.067)	(0.067)	(0.043)
1944 × East of Line of Contact	−0.038	−0.007	0.027
	(0.073)	(0.074)	(0.043)
1945 × East of Line of Contact	−0.131	−0.058	0.012
	(0.090)	(0.090)	(0.044)
1946 × East of Line of Contact	−0.359***	−0.292***	−0.166***
	(0.080)	(0.073)	(0.047)
1950 × East of Line of Contact	−0.254***	−0.181***	−0.108***
	(0.041)	(0.047)	(0.035)
1961 × East of Line of Contact	−0.093**	−0.054	0.004
	(0.038)	(0.046)	(0.041)
1964 × East of Line of Contact	−0.077**	−0.028	0.029
	(0.038)	(0.047)	(0.042)
1971 × East of Line of Contact	−0.036	0.032	0.111*
	(0.070)	(0.077)	(0.067)
1981 × East of Line of Contact	0.017	0.089	0.101*
	(0.087)	(0.087)	(0.061)
1987 × East of Line of Contact	0.033	0.105	0.116*
	(0.090)	(0.090)	(0.063)
1991 × East of Line of Contact	0.040	0.114	0.118*
	(0.094)	(0.094)	(0.065)
2001 × East of Line of Contact	0.011	0.111	0.152**
	(0.098)	(0.098)	(0.075)
2009 × East of Line of Contact	0.028	0.137	0.178**
	(0.100)	(0.102)	(0.082)
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Flexible latitude/longitude controls	Yes	Yes	Yes
Number of observations	7,767	5,111	3,952
Number of unique counties	412	269	269
Number of periods	19	19	19
R-squared	0.90	0.89	0.92
Mean of dep. var.	11.64	11.53	11.62

Notes: This table summarizes estimation results based on German county-level data. Method of estimation is a difference-in-differences (DiD) approach. Flexible latitude/longitude controls include the latitude of the centroid of a county, the longitude and its interaction, each interacted with year dummies. Clustered standard errors (at the county level) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.

Table A.5: The Origin and Evolution of the East-West Population Gap along the Demarcation Line – Split into North and South

	Log population	
	North	South
	(I)	(II)
Pre-WWII differences		
1900 × East	0.064	−0.081*
	(0.078)	(0.043)
1910 × East	0.028	−0.035
	(0.074)	(0.040)
1919 × East	0.038	−0.055
	(0.073)	(0.038)
1925 × East	0.026	−0.042
	(0.072)	(0.036)
1933 × East	0.077*	−0.028
	(0.042)	(0.033)
Base-year (1939) differences		
East	0.747***	0.436***
	(0.134)	(0.102)
Differences during WWII		
1943 × East	0.101	−0.075
	(0.072)	(0.056)
1944 × East	0.076	−0.066
	(0.076)	(0.058)
1945 × East	0.102	−0.074
	(0.071)	(0.050)
Differences during division		
1946 × East	−0.223***	−0.145***
	(0.063)	(0.030)
1950 × East	−0.236***	−0.195***
	(0.035)	(0.021)
1961 × East	−0.242***	−0.229***
	(0.037)	(0.024)
1964 × East	−0.299***	−0.242***
	(0.038)	(0.025)
1971 × East	−0.324***	−0.274***
	(0.047)	(0.027)
1981 × East	−0.339***	−0.260***
	(0.049)	(0.033)
1987 × East	−0.326***	−0.243***
	(0.054)	(0.037)
Differences after reunification		
1991 × East	−0.409***	−0.336***
	(0.059)	(0.039)
2001 × East	−0.475***	−0.449***
	(0.073)	(0.052)
2009 × East	−0.534***	−0.486***
	(0.079)	(0.056)
Pair-Year FE	Yes	Yes
Number of observations	646	646
Number of pairs	34	34
Number of unique counties	36	36
Number of periods	19	19
R-squared	0.77	0.75
Mean of dep. var.	11.65	11.58

Notes: This table summarizes estimation results based on German county-level data. Method of estimation is a two-dimensional regression discontinuity design (RDD) in space combined with a difference-in-differences (DiD) approach. Clustered standard errors (at the county level within a pair) are in parentheses below. *, ** and *** indicate statistical significance at the 10% level, 5% level, and 1% level, respectively.