Local Labor Markets and the Persistence of Population Shocks

Sebastian Till Braun, Anica Kramer and Michael Kvasnicka

University of St Andrews

School of Economics and Finance Discussion Papers

School of Economics and Finance Discussion Paper No. 1715
30 Sep 2017
JEL Classification: J61, R12, R23, N34
Keywords: Population shock, locational fundamentals, agglomeration economies, regional migration, post-war Germany

School of Economics and Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

Local Labor Markets and the Persistence of Population Shocks *

Sebastian Till Braun†
University of St Andrews and Kiel Institute for the World Economy

Anica Kramer‡
RWI and Ruhr-University of Bochum (RUB)

Michael Kvasnicka§
Otto von Guericke University Magdeburg, RWI, IZA

Abstract. This paper studies the persistence of a large, unexpected, and regionally very unevenly distributed population shock, the inflow of eight million ethnic Germans from Eastern Europe to West Germany after World War II. Using detailed census data from 1939 to 1970, we show that the shock had a persistent effect on the distribution of population *within* local labor markets, but only a temporary effect on the distribution *between* labor markets. These results suggest that locational fundamentals determine population patterns across but not within local labor markets, and they can help to explain why previous studies on the persistence of population shocks reached such different conclusions.

Keywords: Population shock, locational fundamentals, agglomeration economies, regional migration, post-war Germany.
JEL Classification: J61, R12, R23, N34.

*This paper has benefited from comments by Thomas Bauer, Dávid Krisztián Nagy and participants at RWI, the 2017 EALE Conference, the 2016 RGS/RWI Workshop on the Economics of Migration, and the 2016 Spring Meeting of Young Economists. Kathleen Kürschner, Yue Huang, and Anja Rösner have provided valuable research assistance in collecting and processing various historical data for different sub-national administrative regions in Germany. The research in this paper was funded by Deutsche Forschungsgemeinschaft (grant no. BR 4979/1-1, "Die volkswirtschaftlichen Effekte der Vertriebenen und ihre Integration in Westdeutschland, 1945-70"). Any remaining errors are our own.

†University of St Andrews, Castlecliffe, The Scores, Fife, KY16 9AR, Scotland, UK. Email: stb2@st-andrews.ac.uk. Phone: +44-1334-462421.

‡Corresponding author. RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany. Email: anica.kramer@rwi-essen.de. Phone: +49-201-8149-238.

§Otto von Guericke University Magdeburg, Universitätsplatz 2, 39016 Magdeburg, Germany. Email Michael.Kvasnicka@ovgu.de. Phone: +49-391-67-58739.

1 Introduction

We study the persistence of a very large population shock, the inflow of eight million displaced Germans (expellees) from Eastern Europe to West Germany after World War II. This population shock hit West German counties very unequally, with expellee inflow rates ranging from 1.4% of the pre-war population to as much as 83%. We show that this migration-induced regional population shock had a persistent effect on the distribution of population *within* inter-connected labor markets, but only a temporary effect on the distribution of population *between* labor markets.

Our findings can help to explain the disparate results in the growing empirical literature on the persistence of population shocks. This literature exploits population shocks to gauge the relative importance of the two main explanations put forward for the spatial distribution of economic activity, locational fundamentals and increasing returns.[1] The locational fundamentals' theory holds that long-lasting geographic conditions, such as access to a river, determine the spatial distribution of economic activity. Consequently, shocks to the spatial distribution of population should have only temporary effects on regional population patterns. Policy makers, as a consequence, have little scope to affect the spatial distribution of economic activity, as locational fundamentals are typically hard to change (see e.g. Head and Mayer, 2003). The increasing returns theory, in contrast, suggests that population density itself may enhance productivity because of agglomeration economies. According to this second theory, therefore, policies and shocks to the distribution of economic activity could well have long-run consequences if they are large enough to move the economy from one equilibrium to another (see Henderson, 1974; Krugman, 1991, for seminal theoretical contributions).

Empirical studies exploring these explanations have produced diverging results. A first set of studies shows that bombings during World War II had no persistent effect on city size in Japan (Davis and Weinstein, 2002) and West Germany (Brakman et al., 2004). Furthermore, Davis and Weinstein (2008) find that the industrial structure of Japanese cities also recovered quickly to its pre-war pattern. The findings of this first set of studies provide empirical support for the locational fundamentals theory, which predicts that temporary shocks have only temporary effects. A distinctive feature of these studies is that they typically use larger cities as their unit of observation.[2] This is of importance for the argument developed in this paper, since larger cities are usually located in different regional labor markets.

A second set of studies, in contrast, finds that migration-induced population shocks during and after World War II were highly persistent. Sarvimäki (2011) shows that the inflow of forced migrants into rural areas of Finland had a re-inforcing effect on post-war population growth, and Schumann (2014), focusing on the West German state of Baden-Württemberg, shows that expellee inflows had a persistent effect on municipality size. Similarly, Eder and Halla (2016) find that inner-Austrian migration out of the (temporary) Soviet occupation zone still affects the spatial distribution of population in Austria today. The findings of this second set of studies hence suggest that locational fundamentals do not determine long-run population

[1] See Redding (2010) for a general overview of the existing empirical literature on new economic geography, including the empirical approaches to distinguish between locational fundamentals and increasing returns.

[2] Miguel and Roland (2011) is an exception in this regard. The authors use district-level data to show that US bombing during the Vietnam War had no long-run effect on later economic development in Vietnam.

patterns.

We contribute to this empirical literature by studying the persistence of a major population shock, the inflow to West Germany of German expellees from Eastern Europe after World War II.[3] Two features make the historical episode particularly well suited for our analysis. First, the inflow was not only large, increasing West Germany's population from 39 million in 1939 to 48 million in 1950, but also very unequally distributed across West German counties. Second, the initial allocation of expellees was driven by the availability of housing and the geographic distance between origin and destination regions, not by economic fundamentals. In particular, we show that conditional on control variables for the local housing supply, the distribution of expellees was unrelated to pre-war trends in population growth.[4]

We show that the choice of the regional unit of observation and the type of variation exploited, so far largely ignored in the literature, are vital for the estimated persistence of the population shock. Specifically, we find that expellee inflows had a persistent effect on the spatial distribution of population within but not between interconnected local labor markets. A likely explanation for our finding is that an individual's choice of location depends largely on labor-market-wide variables, such as wages. If so, a population shock will only cause re-adjustments between but not within labor markets. Whether the former adjustments between labor markets are detectable empirically will depend, as noted, on the unit of observation and the type of variation exploited.

This general insight can help to explain the diverging results in the existing literature on the persistence of population shocks.[5] To illustrate, consider the aforementioned study by Schumann (2014) who also focuses on the inflow of expellees to West Germany after World War II. Schumann (2014) restricts the analysis to one federal state, Baden-Württemberg. After the war, Baden-Württemberg was temporarily divided into two occupation zones, a French and an American zone. Expellees were initially not resettled into the French zone of occupation, which created a sharp discontinuity at the border to the American zone of occupation. Schumann (2014) shows that this discontinuity is still visible 25 years after the war. Importantly, however, municipalities along the occupation zone border often belonged to the same local labor market. Schumann thus effectively exploits only variation *within* local labor markets.

Unlike Schumann, our analysis considers the whole of West Germany and exploits variation in expellee inflows not only within, but also between local labor markets. When we do exploit only variation within local labor markets, we confirm the results Schumann obtained for municipalities in Baden-Württemberg. However, and importantly, we also show that his results do not carry over to population patterns *between* local labor markets. At this more aggregated regional level, population patterns quickly revert back to their

[3] Previous studies have exploited regional variation in expellee inflow rates to analyze the short-run effect on native employment (Braun and Omar Mahmoud, 2014) and structural change (Braun and Kvasnicka, 2014), the dynamic response of local labor markets (Braun and Weber, 2016), and the effect on productivity and regional economic development (Peters, 2017).

[4] Note also that expellees and natives were very similar in several important aspects. Expellees and natives both spoke German as their mother tongue and had both been educated in German schools. In fact, their pre-war level of education was almost identical (Bauer et al., 2013). Expellees were also not selected based on their skills or other characteristics, as virtually all Germans living east to the post-war border between East Germany and Poland were displaced (see Section 2 for details).

[5] Our findings complement previous arguments by Schumann (2014) who suggests that locational fundamentals might be particularly important for geographically diverse countries and for urban areas. Likewise, Sarvimäki (2011) suggests that a population shock may be large enough to change the equilibrium of rural areas "at the brink of becoming a local manufacturing center" (p. 3) but not the equilibrium of well established cities.

pre-war level. Our preferred estimate suggests that as much as 83% of the initial shock is dissipated 25 years after the war. This finding provides evidence for the importance of locational fundamentals in determining the spatial distribution of population between local labor markets. And it highlights, more generally, the crucial relevance of the choice of the regional unit[6] and the type of variation exploited in the analysis for the estimated persistence of a population shock.

Our findings are also relevant for the literature that studies the effect of immigrant inflows on population outflows. This literature has not yet reached a definite conclusion: Some studies find that immigrant inflows lead to native outflows (Borjas, 2006; Boustan et al., 2010), whereas other studies find no such link (Card and DiNardo, 2000; Card, 2001). Using net migration as an additional outcome variable, we show that variation in expellee inflows between but not within local labor markets is negatively associated with net population flows, mirroring our results for population growth. Since expellees were more likely to migrate than natives (Bauer et al., 2013; Braun and Weber, 2016), they are likely to have contributed disproportionally to these migration flows.

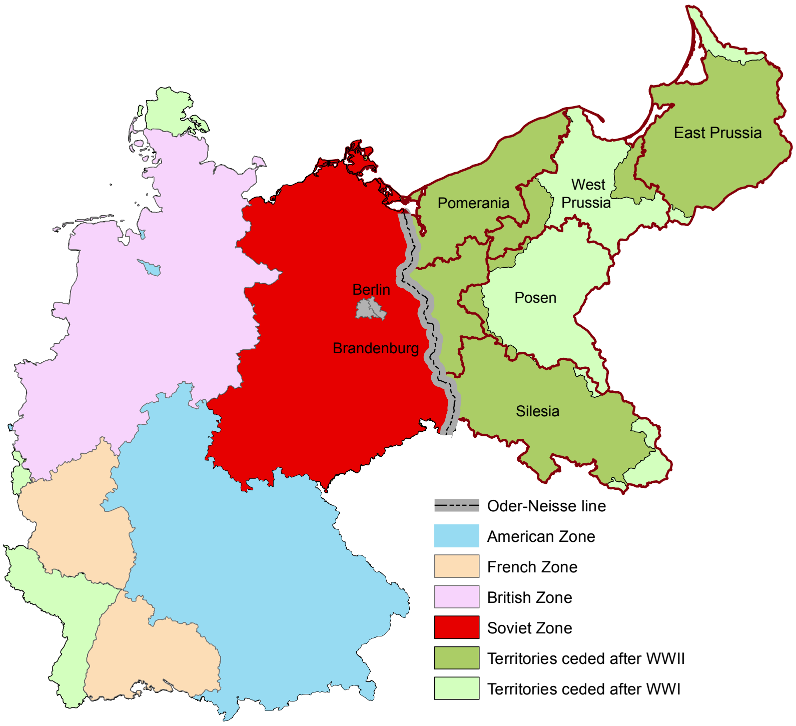
The paper proceeds as follows. Section 2 provides background information on the expellee inflow to West Germany after World War II and develops our hypotheses. Section 3 describes the various data sources and the identification strategy we use in our empirical analysis. Section 4 presents and discusses our regression results. Finally, Section 5 summarizes our main findings and concludes.

2 Historical Background and Hypotheses

After World War II, West Germany experienced the inflow of eight million expellees (*Heimatvertriebene*), most of them from the ceded eastern provinces of the defeated German Reich. The displacement of Germans took place from 1944 to 1950 and occurred in three distinct phases (for further details see, e.g., Connor (2007), Douglas (2012), and Schulze (2011)). The first phase began in 1944, when hundreds of thousands of Germans from the eastern provinces of the German Reich fled from the approaching Red Army. Most of these refugees planned to return home after the end of the war, and therefore fled to the nearest West German regions. After Nazi Germany's unconditional surrender in May 1945, Polish and Czech authorities began to drive their remaining German populations out. These so-called wild expulsions, which constituted the second phase of the displacement, were not yet sanctioned by an international agreement. The third phase began after the Soviet Union, the United Kingdom, and the United States signed the Potsdam Agreement in August 1945. The Potsdam Agreement shifted Germany's eastern border westwards to the Oder-Neisse line. The former eastern provinces of the German Reich were placed under Polish or Russian control (see Figure 1). Germans remaining east to the new border were brought to post-war Germany in compulsory and organized transfers. The German territory west to the Oder-Neisse line was divided into four occupation zones: a British, a French, an American, and a Soviet zone.

[6]The choice of regional unit also conciliates the findings of Schumann (2014) and the results on internal migration in Braun and Weber (2016). The latter develop a two-region search and matching model to analyze how regional labor markets adjusted to the expellee inflow, and show that migration from high- to low-inflow regions was an important channel of adjustment. The result appears to contradict Schumann who finds no evidence for major outflows from the high-inflow American occupation zone. The different units of observations can explain these seemingly disparate findings: While Schumann (2014) studies small municipalities located close to each other, Braun and Weber (2016) divide West Germany in their analysis in only two large regions.

FIG. 1: The Division of Germany and German Territorial Losses after World War I and II



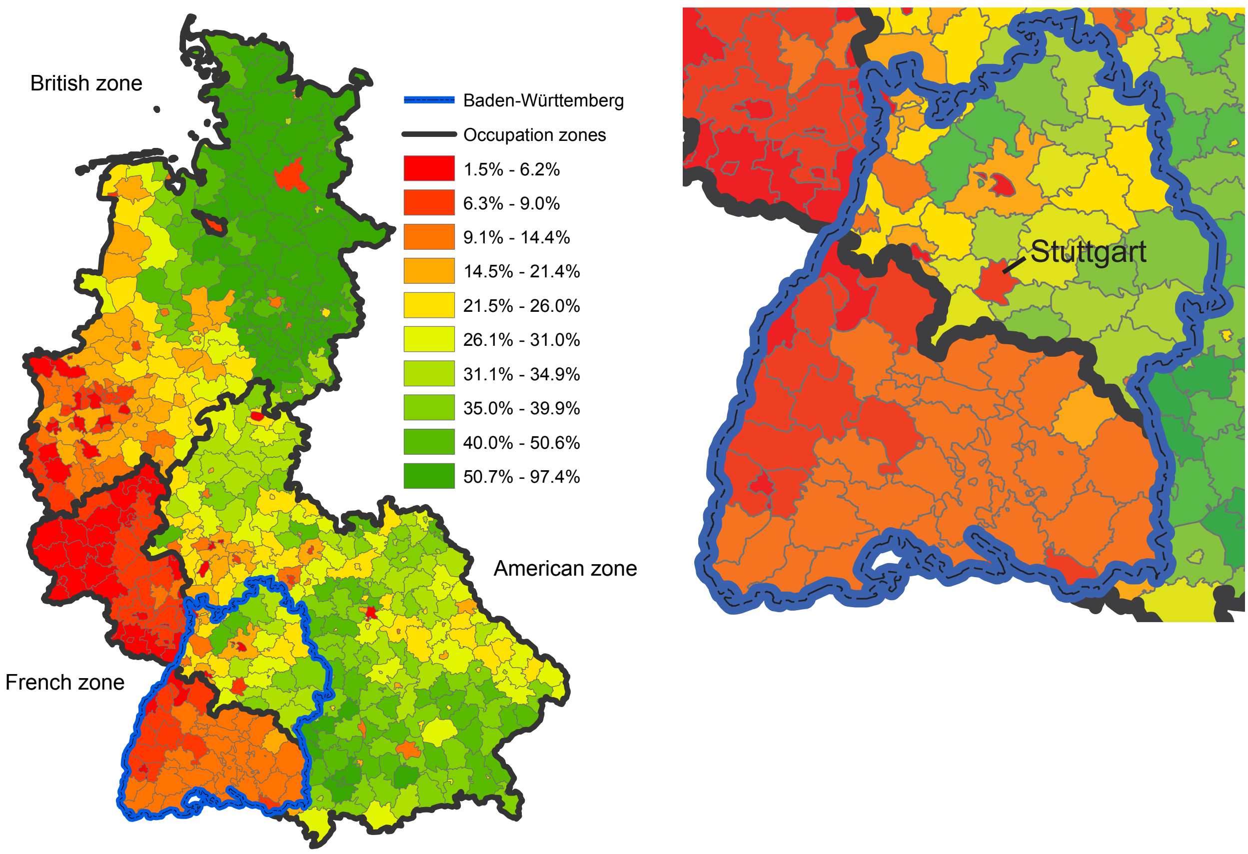
Base maps: MPIDR (2011).

Overall, the mass exodus of Germans from East and Central Europe involved at least 12 million people. Most expellees re-settled in West Germany. By September 1950, expellees accounted for 16.5% of the West German population. However, the population share of expellees differed greatly across West German counties, ranging from 1.8% in Pirmasens to 41.4% in Goslar. In our empirical analysis, we will exploit this pronounced regional variation. We document the regional structure of the expellee inflow and its underlying reasons below, before outlining the hypotheses that we will test later in the empirical analysis.

Regional Distribution: Figure 2a illustrates the immigration-induced increase in population across counties, as measured by the number of expellees in 1950 over the population in 1939 (henceforth, expellee inflow rate). This figure provides three main insights. First, there were large differences in the expellee inflow rate *between* occupation zones. In particular, the rate was much higher in the American zone (30.2 %) and British zone (31.4%) than in the French zone (7.5%). This is because the French initially refused to accept any expellees in their occupation zone. The French felt not bound by the Potsdam Agreement, as they had not been invited to the Potsdam conference. As a result of the French refusal, expellees were initially transferred only to the American and British occupation zones in the third phase of the displacement. This created a sharp discontinuity in expellee inflow rates at the border between the American and French zones of occupation, as illustrated in greater detail in Figure 2b. It is this sharp discontinuity that Schumann (2014) exploits to estimate the persistence of the expellee inflow on the spatial distribution of population in parts of Baden-Württemberg.

FIG. 2: Expellee Inflow Rates

(a) West Germany (b) Baden-Württemberg



Notes: The figures depicts the number of expellees per county on 13 September 1950 over the population per county on 1 September 1939 in West Germany (panel 2a) and the state of Baden-Württemberg (panel 2b). The black line depicts the border of the three occupation zones. The blue line, which partly overlaps with with the black line, depicts the border of the West German state of Baden-Württemberg.
Source: Statistisches Bundesamt (1952). *Basemap*: MPIDR (2011).

Second, the population share of expellees also differed greatly *within* occupation zones. This is particularly evident for the British zone where the expellee inflow rate ranged from 4.0% in the western county of Bocholt to 83.5% in the north-eastern county of Eckernförde. This west-east divide was a result of the largely undirected flight of Germans during the final stages of the war (the first phase of the displacement). As the Soviet troops pushed westwards, Germans residing in the eastern provinces of the German Reich were forced to seek shelter further west. The refugees thus crowded in the most accessible regions in the west and north-west of West Germany. Refugees from East Prussia, for instance, mostly ended up in the northern state of Schleswig-Holstein, as East Prussia and Schleswig-Holstein were connected via the Baltic Sea. The wild expulsions (second phase of the displacement) only added to the regional imbalance between counties in the west and east, as Polish and Czech authorities often just drove Germans across the border into occupied Germany. Many Germans from the Sudetenland, for instance, were forced into neighboring Bavaria.

Third, the population share of expellees also differed systematically between cities and surrounding rural areas. Figure 2b highlights the example of the city of Stuttgart. While the expellee inflow rate was only 8.5% in Stuttgart, it ranged from 27.3% to 31.7% in the five immediately neighboring rural counties. Similar patterns can be observed for other cities such as Hamburg in the north, Kassel in the center, and Munich in the south of Germany. Expellees were generally more likely to be placed in rural areas, where the housing stock had remained largely intact during the war (Connor, 2007).

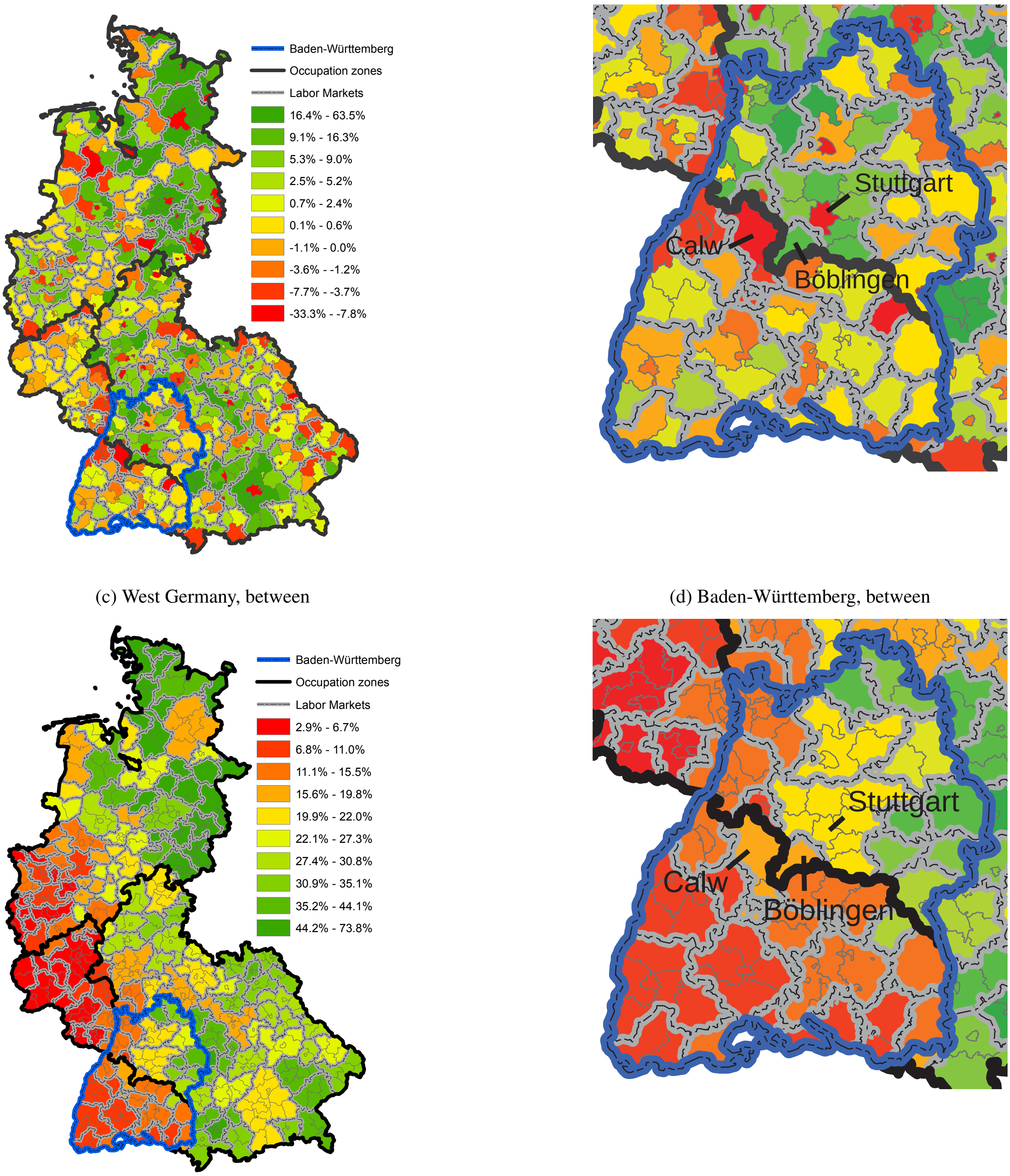
Recapitulating the above, the historical setting we explore provides rich spatial variation in expellee inflows rates. Expellee inflow rates differed both *between* counties far away from each other–for instance, between counties located in the west and the north of Germany–and between neighboring counties–for instance, between neighboring counties at either side of the French occupation zone border. The average inflow rate across all counties was 0.270, with a standard deviation of 0.176.

Variation Between and Within Local Labor Markets: The labor markets of neighboring counties are often well connected through commuting flows, and several counties typically form one local labor market. Based on commuting flows, IfW (1974) defines 164 labor market regions, each consisting of an average of 3.4 counties.[7] Expellee inflow rates in our setting differ greatly both *within* and *between* these local labor markets. To show this, we decompose the overall variation in expellee inflow rates. Let I_{ij} be the expellee inflow rate for county i located in labor market j. We decompose I_{ij} into a between component, $\bar{I}_j$, and a within component, $I_{ij} - \bar{I}_j$. The between component is simply the expellee inflow rate measured at the level of local labor market j, while the within component is the difference between the inflow rate of a particular county i in labor market j and the inflow rate of labor market j.

Figure 3a illustrates for West German counties the within component, i.e., the variation in expellee inflow rates across counties located in the same labor market region. The within component ranges from -0.333 to 0.635 with a standard deviation of 0.112. Zooming in to the state of Baden-Württemberg, Figure 3b illustrates that the within-labor-market variation comes from three sources. First, the borders of local

[7] To the best of our knowledge, the definition in IfW (1974) is the earliest definition of local labor markets in West Germany. A few counties belong to more than just one local labor market. In this case, we assign the county to the labor market with which it shares the larger area.

FIG. 3: Variation in Expellee Inflow Rates Within and Between Labor Markets



(a) West Germany, within

(b) Baden-Württemberg, within

(c) West Germany, between

(d) Baden-Württemberg, between

Notes: The figures depict the number of expellees per county on 13 September 1950 over the population on 1 September 1939 in West Germany (panels 3a and 3c) and in the state of Baden-Württemberg (panels 3b and 3d). The upper two panels calculate figures at the level of counties, the lower two panels at the level of local labor markets. The solid black line depicts the borders of the three occupation zones, the dashed black line on grey ground depicts the borders of local labor markets, and the blue line, which partly overlaps with the black line, depicts the border of the West German state of Baden-Württemberg.
Source: Statistisches Bundesamt (1952). *Basemap*: MPIDR (2011).

labor markets (the dashed black line on grey ground in the figure) frequently spanned counties from both sides of the French occupation zone border, and these counties typically experienced very different inflow rates. The counties of Calw and Böblingen, for instance, were both part of the same labor market but their inflow rates differed greatly. Whereas the inflow rate of Calw stood at 8.7% in 1950, the inflow rate of Böblingen was 30.5% (see Figure 3b). The inflow rate in Calw, therefore, was significantly below the inflow rate of the local labor market in which it was situated. Second, local labor markets frequently consisted of both a larger city, typically with low expellee inflow rates, and surrounding hinterlands, with larger inflow rates. The city of Stuttgart is a case in point (see again Figure 3b). Third, variation in expellee inflow rates within local labor markets also reflected the east-west or north-south gradient in inflow rates, although this variation was typically more modest between neighboring counties.

In addition to this variation within local labor markets, there was also sizeable variation in expellee inflow rates between local labor markets. Figure 3c illustrates this between component of the total variation in expellee inflow rates for West Germany. The between component varies between 0.029 and 0.738, with a mean of 0.257 and a standard deviation of 0.162. The figure shows that much of the variation in the between component came from the stark difference between local labor markets in the north and east of the country and those in the west and south-west. As noted before, this east-west divide was mostly the result of the largely undirected flight to the most accessible West German regions at the end of World War II; and it was reinforced by the French refusal to allow any expellees into their occupation zone in the south-west of Germany. Importantly, however, Figure 3d, which zooms in to the state of Baden-Württemberg, shows that the sharp discontinuity of expellee inflow rates at the French occupation zone largely disappears when inflow rates are calculated at the level of local labor markets. This is mainly because some labor markets spanned counties from both sides of the occupation zone border. Moreover, the low inflow rate into Stuttgart counter-balanced the high inflow rates of counties in its hinterland, including those at the occupation zone border.

Hypotheses: We hypothesize that the persistence of the population shock will differ, depending on the type of variation exploited in the empirical analysis. The key idea is simple. Theories of spatial equilibrium typically postulate that wages (or labor market conditions more generally) are specific to a location, and that the location decision of workers depends on wages across locations. The relevant units of analysis for testing the predictions of these theories are thus local labor markets, not administrative units. Furthermore, if the location decision depends only on labor-market-wide variables, movements within labor markets are irrelevant for the spatial equilibrium. Consequently, a shock to the spatial distribution of population within local labor markets will not cause adjustment within but only across labor markets.

What does this simple insight imply for our specific historical setting? As we have shown, expellee inflows implied a shock to population both between and within local labor markets. We hypothesize that population shocks within local labor markets are persistent. To illustrate, consider again the counties of Calw and Böblingen which are located in the same labor market. If location decisions *only* depend on labor-market-wide conditions, the differential inflows into the two counties should not have led to relocations between these two counties (i.e., within the local labor market of which the two counties are part). Consequently, the expellee-induced shock to population should prove persistent when using only the variation within local labor markets. This could explain why Schumann (2014), who exploits the sharp discontinuity in expellee inflows at the French occupation zone border, finds expellee inflows to be very persistent. Between nearby municipalities and counties that are located at opposite sides of the occupation zone border but belong to the same labor market, there simply should not occur sizeable migration flows that significantly attenuate the initial population shock.

Inflows into Calw and Böblingen had an effect on the spatial equilibrium only in so far as they affected the population of the local labor market they were located in. It is this variation between, and not the one within, local labor markets that allows us to test theories of spatial labor market equilibrium. If locational fundamentals determine population patterns, population shocks should not be persistent. We would thus expect migration from local labor markets with high inflows to local labor markets with low inflows.[8] If, instead, individuals migrate from local labor markets with low inflows to local labor markets with high inflows, this would point towards the importance of agglomeration economies, and be inconsistent with the locational fundamentals hypothesis.

3 Empirical Strategy

We exploit regional variation in expellee-induced population increases across West German counties to test our hypothesis. We use West German counties in their 1970 borders.[9] Our main data sources are the population and occupation censuses of 1939, 1946, 1950, 1961 and 1970 which we have digitalized for our analysis. Appendix D provides a detailed overview of the data sources for all variables.

Within and Between Regressions: We begin by estimating the following OLS regression:

$$G_{ij}^{70,50} = \alpha_1 + \beta_1 I_{ij}^{50,39} + X_{ij}\gamma_1 + u_{ij}, \tag{1}$$

where $G_{ij}^{70,50}$ is the population change in 1950-70 over the population in 1939 of county i in labor market j (henceforth: population growth in 1950-70), $I_{ij}^{50,39}$ is the expellee inflow rate of county i between 1939-50, X_{ij} is a vector of covariates, and u_{ij} is an error term.[10] The regression tests whether expellee-induced population growth in 1939-50 reduced or reinforced population growth in 1950-70. The former case would suggest that locational fundamentals are of importance, the latter instead that increasing returns matter. Specification (1) mimics the conventional approach in the literature (see, for instance, Sarvimäki, 2011; Davis and

[8] This outcome is not necessarily inconsistent with the increasing returns hypothesis. If the population shock is not large enough to move local labor markets from one equilibrium to another, we would expect mean reversion even under the increasing returns hypothesis.

[9] There are 548 counties in 1970. However, a few of them experienced changes in their administrative borders between 1939 and 1970. While population data for 1939, 1950 and 1970 are available for counties in their 1970 borders, some of our control variables refer to counties in their 1939 or 1950 borders. We account for border changes between 1939 and 1970 by merging counties so that county borders are generally comparable over time (see Appendix C for the details). This leaves us with 511 counties. Counties located in the states of Rhineland-Palatinate and Schleswig-Holstein saw major border changes in 1969/70. For counties located in these two states, we use the administrative borders immediately before the major border changes.

[10] We normalize both population change in 1950-70 and expellee inflows by population in 1939 to simplify the interpretation of β_1. In particular, $\beta_1 = -1$ indicates that the expellee-induced population shock is completely reversed by 1970.

Weinstein, 2002) to test whether shock-induced population growth in one period affects population growth in subsequent (post-shock) periods.

As discussed in Section 2, we expect the persistence of expellee-induced population growth to differ depending on the type of variation we exploit in the empirical analysis. We thus run two additional specifications in which we only exploit variation within or between local labor markets:

$$\text{Within:} \qquad (G_{ij}^{70,50} - \bar{G}_{j}^{70,50}) = \beta_2 (I_{ij}^{50,39} - \bar{I}_{ij}^{50,39}) + (X_{ij} - \bar{X}_j)\gamma_2 + (u_{ij} - \bar{u}_j) \tag{2}$$

$$\text{Between:} \qquad \bar{G}_{j}^{70,50} = \alpha_3 + \beta_3 \bar{I}_{jt}^{50,39} + \bar{X}_j \gamma_3 + \bar{u}_j, \tag{3}$$

where $\bar{X}_j$ denotes the value of variable X for local labor market j. Specification (2) considers deviations from labor-market-wide levels, and thus exploits only variation between (nearby) counties *within* the same local labor market. Specification (3) aggregates the county-level data to the level of local labor markets, and only uses the variation *between* (more distant) local labor markets in West Germany. The between specification differs from Specification (1) in the choice of the regional unit considered: The former studies local labor markets, the latter focuses on counties.

Hypotheses: Our hypothesis, which we developed informally in Section 2, can be decomposed into three parts. Our first hypothesis states that $\beta_2 \neq \beta_3$, i.e., the persistence of population growth will differ depending on whether we exploit variation within or between local labor markets.[11] The second hypothesis is that $\beta_2 = 0$, i.e., variation in expellee-induced population growth within labor markets has no bearing on subsequent population growth. This is simply because variation in population growth within local labor markets has no effect on the spatial equilibrium if economic agents base their location decisions only on labor-market-wide variables. The third hypothesis concerns β_3. If people move as a response to a population shock, they will do so by moving between local labor markets so as to exploit variation in labor market conditions. It is thus β_3 that allows us to test theories of spatial equilibrium. Specifically, $\beta_3 < 0$ provides support for the locational fundamentals theory, with $\beta_3 = -1$ indicating a complete reversal of the initial population shock, whereas $\beta_3 > 0$ is only consistent with the increasing returns theory.

Identification: Identifying the *causal* effect of population growth on subsequent population growth is challenging because confounding factors may drive population growth in both periods (Davis and Weinstein, 2002; Sarvimäki, 2011). Our empirical exercise isolates variation in wartime population growth that is due to the inflow of expellees. The key identifying assumption for a causal interpretation of β_1, β_2, and β_3 is that there is no unobserved factor that drives both the expellee inflow rate and population growth in 1950-70. In particular, estimates will be upward (downward) biased if expellees systematically selected, based on unobservable characteristics, into West German regions with a higher (lower) underlying potential for population growth.

For several reasons, self-selection of expellees was arguably a minor problem until 1950, when we

[11] This hypothesis implies that regression equation (1) is misspecified. In particular, we postulate a regression model in which labor-market wide expellee inflows have a different effect on post-war population growth than deviations from this average, i.e., $G_{ij}^{70,50} = \alpha_1 + \beta_2 (I_{ij}^{50,39} - \bar{I}_{j}^{50,39}) + \beta_3 \bar{I}_{j}^{50,39} + X_{ij}\gamma_1 + u_{ij}$.

measure expellee inflows. First, expellees did not choose their initial destination in West Germany based on local economic conditions (which, in turn, are likely to correlate with potential population growth). Expellees initially fled to the most accessible regions in West Germany and were later forcibly transferred to a destination (see Section 2). Second, the military governments of the occupation powers, overburdened by the mass inflow of millions of expellees, did not redistribute expellees according to local economic conditions (Braun and Omar Mahmoud, 2014; Braun and Kvasnicka, 2014). Finally, once expellees were resettled in a destination, they could not just move on by their own choice. The occupying powers enacted severe moving restrictions (Müller and Simon, 1959), so that the initial distribution of expellees proved very persistent in the first years after the war.

Our specific historical context thus limits concerns of endogenous self-selection. However, there are still two main threats to identification. First, while military governments did not allocate expellees according to local economic conditions, the distribution of expellees was not altogether random. Since the main objective of military authorities at the time was to find accommodation for all expellees, expellees were under-represented in urban areas that were devastated by the war and offered only limited housing capacity. If war destruction and urbanization rates had an effect on post-war population growth, coefficients on expellee inflow rates will be biased in unconditional OLS regressions. Second, moving restrictions were gradually phased out by 1949. Some expellees, as a consequence, might have moved endogenously by 1950.

We deal with these threats to identification in two main ways. First, we control for war destruction and urbanization, and for other local characteristics that might have affected population growth. We then show that conditional on these covariates, expellee inflow rates are unrelated to regional population growth between 1871 and 1939. This corroborates our argument that once we condition on urbanization and measures of war destruction, expellee inflows were unrelated to potential population growth. Second, we use the expellee inflow rate between 1939 and 1946 as an instrument for the expellee inflow rate between 1939 and 1950. Since strict restrictions on relocations were still in place in 1946, this IV regression exploits only variation in expellee inflow rates that is attributable to the initial inflow of expellees, and not to subsequent, and potentially endogenous, relocations within West Germany.

Controls: We control for regional characteristics that might have affected expellee settlement patterns and influenced potential population growth. First and foremost, we include various measures of war destruction. War destruction correlates–through the availability of housing–with local expellee inflow rates and might have affected also post-war population growth.[12] We use three different measures of war destruction. As our baseline measure, we consider the share of dwellings built until 1945 that were damaged in the war, using information from the 1950 housing census. Unfortunately, the housing census did not count dwellings that were completely destroyed in the war. The share of damaged dwellings is thus calculated only relative to residential housing that could still accommodate residents in 1950. Our second measure is rubble at the end of the war per capita in 1939, as also used in previous work by Brakman et al. (2004), Burchardi and Hassan (2013) and Braun and Kvasnicka (2014). Unfortunately, data on rubble are only available for the

[12]Heavily destroyed cities, in fact, grew faster after the war (Brakman et al., 2004).

199 largest West German cities. We aggregated the city-level data to the county level, assuming that smaller municipalities did not suffer any war destruction. The rubble indicator will thus underestimate the extent of war destruction in counties with smaller municipalities. The third measure classifies the loss in housing space in four categories, ranging from 'no losses' (1) to 'very substantial losses' (4). This indicator variable is based on various administrative sources at the national and federal state level. We include dummies for three categories, with 'negligible' (2) as the baseline category.

Second, concerning measures of urbanization, we control for a county's population density in 1939. Urban areas offered less potential for housing expellees, and thus received lower expellee inflows. At the same time, population growth may have systematically differed between rural and urban areas. We also use, as an alternative measure of urbanization, the population share living in cities with at least 10,000 inhabitants.

A third set of covariates includes variables that proxy pre-war economic conditions. First, we include information on pre-war turnover per worker, sampled from turnover tax statistics. This variable accounts for pre-war differences in economic conditions and development. Second, we include the share of the total workforce in a county that is employed in agriculture in 1939.

Finally, we also include a dummy for counties that are less than 75 kilometers away from the post-war inner-German border. Redding and Sturm (2008) show that cities at the inner-German border generally experienced lower population growth than other West German cities, and attribute this difference to a disproportionate loss in market access for cities at the new border. At the same time, counties at the inner-German border received higher-than-average expellee inflows, due to their proximity to the eastern territories of the German Reich (see Section 2).

TABLE 1: Expellee Inflows and Pre-war Population Growth

	Annual population growth b/w		
	1871-1939	1910-1939	1925-1939
	(1)	(2)	(3)
Inflow of expellees 1950 over population 1939	0.010	0.001	-0.003
	(0.006)	(0.003)	(0.008)
Observations	511	511	466

Notes: All regressions include our standard set of control variables, i.e., population density in 1939, the employment share in agriculture in 1939, turnover per capita in 1935, share damaged dwellings, and a dummy for counties within 75 km of the inner-German border. Data on population in 1925 is missing for counties located in the state of Rhineland-Palatinate. Robust standard errors clustered at the level of local labor markets are reported in brackets.

Expellee Inflows and Pre-war Population Growth: Before we present our main results, we show that pre-war population growth is uncorrelated with expellee inflow rates once we condition on our set of covariates. Table 1 presents the results from regressing population growth in 1871-1939, 1910-1939, and 1925-1939 on expellee inflow rates and on our standard set of covariates (Appendix A presents the corresponding conditional scatter plots). The coefficient on the expellee inflow rate is not statistically significant in any of the three regressions. This corroborates our identifying assumption that conditional on our covariates, expellee inflow rates do not correlate with a region's underlying population growth.

4 Results

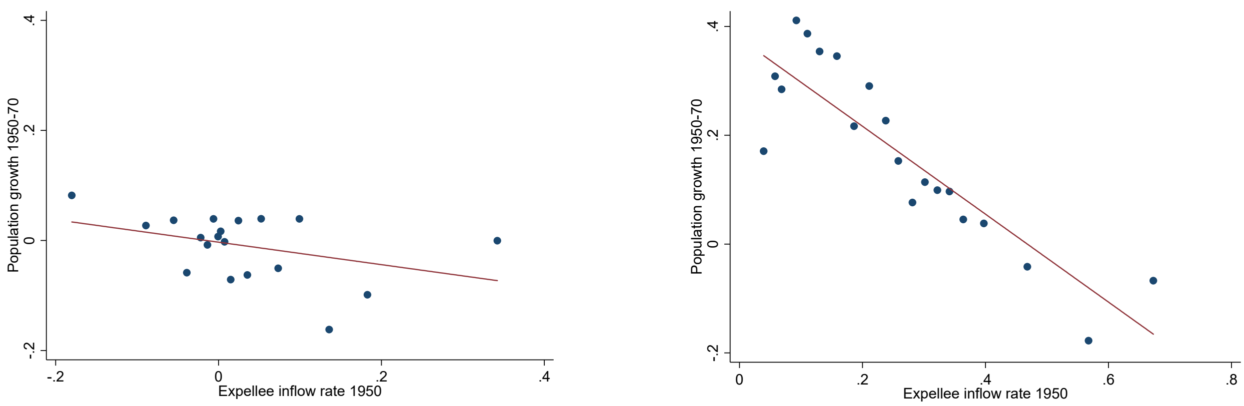
4.1 Baseline Results

Binned Scatter Plots: We begin by documenting graphically the respective importance of the two sources of variation exploited in our analysis. Figure 4 depicts unconditional binned scatter plots of population growth in 1950-70 and expellee inflow rates, grouping expellee inflow rates into 20 equal-sized bins. Figure 4a uses only variation within local labor markets, whereas Figure 4b uses only variation between local labor markets. Each scatter plot also shows the respective linear OLS regression line.

FIG. 4: Binned Scatter Plots (Unconditional)

(a) Within Local Labor Markets

(b) Between Local Labor Markets



Notes: The figures depict binned scatter plots of population growth in 1950-70 and expellee inflow rates, grouping expellee inflow rates into 20 equal-sized bins. Panel 4a relates deviations from labor-market-wide averages to each other, whereas Panel 4b considers labor-market-wide averages themselves.

Figure 4a shows a weakly negative relationship between the expellee inflow rate and post-war population growth. The binned scatter points are quite dispersed around the regression line, which suggests that its slope is only imprecisely estimated. The estimated OLS slope coefficient is -0.204 with a standard error of 0.834. The unconditional regression thus suggests that expellee-induced population growth had a persistent effect on population patterns within local labor markets, as subsequent population growth did not reverse the initial shock. This is consistent with our hypothesis that differential inflows into counties located in the same labor market will not lead to relocations between these counties.

This does not imply, however, that there has been no adjustment *between* labor market regions. In fact, Figure 4b shows that local labor markets that exhibited faster (slower) population growth in 1939-1950 grew, on average, less (more) strongly in 1950-1970. The estimated slope coefficient is -0.808 with a standard error of 0.090. This strong and statistically significant negative association is suggestive of significant population adjustments that almost completely reversed the initial population shock (a coefficient of -1 would indicate complete reversion).

Taken together, Figures 4a and 4b illustrate our main point. The persistence of population shocks might be very different, depending on whether one considers variation within or between local labor markets. In our setting, the within variation points towards a high persistence of population shocks, and thus suggests that locational fundamentals are not an important determinant of population patterns within local labor markets. The between variation, in contrast, suggests that across local labor markets, population shocks are largely reversed, which is in line with the locational fundamentals hypothesis.

Regression Results: For reasons discussed in Section 2, expellee-induced population growth in 1939-50 is unlikely to be completely orthogonal to underlying population growth potential in 1950-70. We therefore next test whether the unconditional correlation is still evident in a multivariate regression framework.

Table 2 reports our main regression results. The table reports conditional OLS (columns (1)-(3)) and IV estimates (columns (4)-(6)). For each set of estimates, we first present results that are based on the overall variation in expellee inflows (columns (1) and (4)), and then results that are based only on the variation of expellee inflows within local labor markets (columns (2) and (5)) and between local labor markets (columns (3) and (6)).

In the first specification, we regress population growth between 1950 and 1970 on our key explanatory variable, the expellee inflow rate, and our set of covariates. As shown in column (1) of Table 2, the estimated coefficient on the expellee inflow rate is -0.311 with a standard error of 0.140. A one percentage point increase in a county's expellee inflow rate thus reduced subsequent population growth in 1950-70 by 0.311 percentage points. The result–based on the overall variation for West Germany at county level–suggests that there was some reversion to the pre-shock population distribution. Overall, therefore, counties subjected to a larger positive (negative) population shock in 1939-1950 tended to show lower (higher) average population growth in subsequent decades. However, the magnitude of reversion was limited, at least until 1970 and for West Germany as a whole.

In specifications (2) and (3), we decompose the total variation of the population shock into two components, a within local labor market component and a between local labor market component. Specification (2) considers the deviation of variables from the labor-market-wide mean. Exploiting only variation between counties within the same local labor market provides evidence on the persistence of population shocks that differentially hit counties located in the same labor market. As shown in column (2), the estimated coefficient on our population shock measure turns statistically insignificant in our within regression (and is now, with 0.131, even positive). Thus, within local labor markets, the population shock appears to have been persistent, showing no sign of reversion.

In specification (3), we aggregate our county-level data to the local labor market level and then re-run our full-fledged model at this higher level of regional aggregation. This way, we exploit only variation between local labor markets. The point estimate of -0.671 indicates that between local labor markets, the initial population shock was, to a large degree, reversed in 1950-70. For any percentage point increase in the expellee inflow rate in 1950, subsequent population growth was reduced by 0.671 percentage points. Comparing the results of Specification (1) and (3) also highlights the importance of the unit of observation: Moving from counties to local labor markets more than doubles the absolute magnitude of the coefficient on

TABLE 2: Main Regression Results

	OLS			IV		
	Overall (1)	Within (2)	Between (3)	Overall (4)	Within (5)	Between (6)
Expellee inflow rate 1950	-0.311**	0.131	-0.671***	-0.498***	-0.060	-0.830***
	(0.140)	(0.124)	(0.202)	(0.130)	(0.123)	(0.115)
Pop. density 1939	-0.021***	-0.012***	-0.012***	-0.021***	-0.011***	-0.012**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)
Share agriculture 1939	-0.682***	-0.660***	-0.339***	-0.659***	-0.633***	-0.303**
	(0.098)	(0.096)	(0.105)	(0.095)	(0.091)	(0.140)
Turnover p.c. 1935	-0.003	-0.099***	0.053**	0.003	-0.104***	0.055*
	(0.021)	(0.019)	(0.024)	(0.020)	(0.020)	(0.029)
Share of damaged dwellings	0.208**	0.415***	-0.103	0.139	0.365***	-0.155
	(0.088)	(0.103)	(0.087)	(0.086)	(0.100)	(0.099)
0/1 Inner-German border	-0.129***	-0.038	-0.095***	-0.108***	-0.036	-0.074**
	(0.037)	(0.033)	(0.032)	(0.037)	(0.033)	(0.035)
R-squared	0.314	0.260	0.441	0.307	0.255	0.434
Observations	511	511	157	511	511	157
F-Statistic, excl. instrument				2023.4	716.4	1714.1
First-stage coefficient				0.924***	0.941***	0.946***
				(0.029)	(0.035)	(0.023)

Notes: The dependent variable is the change in population between 1950-70 over the population in 1939. Regression models (1) and (4) use the overall variation in the data, whereas models (2) and (5) uses only the variation within local labor markets, and models (3) and (6) the variation between local labor markets (see Section 3 for details). The IV regressions in columns (4) to (6) use the expellee inflow rate in 1946 as an instrument for the expellee inflow rate in 1950. Robust standard errors are in brackets. Standard errors are clustered at the level of local labor markets in models (1), (2), (4), and (5). *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

the expellee inflow rate.

We next estimate IV regressions to alleviate concerns that some expellees might have endogenously moved by 1950 after moving restrictions were phased out in 1949. The IV regressions isolate the variation in inflow rates that is due only to the initial placement of expellees. Their results are shown in columns (4)-(6) of Table 2.

The first-stage results suggests that we do not have a weak instrument problem. The IV results generally confirm our OLS results although the IV estimates are more negative than the OLS estimates. First, when using the overall-variation (column (1) vs. (4)), the estimated coefficient is now -0.498, considerably smaller than the OLS estimate of -0.311. Second, the within estimator now turns negative to -0.060 (column (5)). However, the estimated coefficient is not statistically significant at any conventional level. The expellee-induced population shock did not induce lower population growth in 1950-70, implying a persistent effect on the spatial distribution of population *within* local labor markets. Finally, the negative point estimate of the between specification also decreases slightly from -0.671 in specification (3) to -0.830 in specification (6). The estimate implies that a 1 percentage point increase in population growth between 1939 and 1950 reduces population growth between 1950 and 1970 by 0.830 percentage points. The population shock hence had very little effect on the spatial distribution of population *between* local labor markets 25 years after the war.

4.2 Robustness Checks

We conduct several tests to assess the robustness of our IV results. Table 3 provides the results of these robustness checks, reproducing our main results–from columns (5) and (6) of Table 2–in Panel A.

TABLE 3: Robustness Checks - IV Results on Expellee Inflow Effect

	Within (1)	Between (2)
A. Baseline regression	-0.060	-0.830***
	(0.123)	(0.115)
B. *Alternative control variables for destruction and urbanization*		
... using rubble per capita	-0.148	-0.777***
	(0.101)	(0.107)
... using loss in housing space (categorial)	-0.127	-0.779***
	(0.101)	(0.109)
... using population share in cities with at least 10,000 inhabitants in 1939	-0.095	-0.845***
	(0.121)	(0.117)
C. *Pre-war population trends*		
... adding population growth 1871-1939	-0.064	-0.837***
	(0.125)	(0.115)
... adding population growth 1910-1939	-0.046	-0.827***
	(0.124)	(0.115)
... adding population growth 1925-1939	-0.030	-0.939***
	(0.134)	(0.124)
D. *Weighted regressions*		
... weighted with 1939 population	-0.067	-0.798***
	(0.114)	(0.127)

Notes: The table reports IV estimates of the effect of the expellee inflow rate in 1950 on population growth in 1950-1970. Each cells reports estimates from a separate regression. The dependent variable is the change in population between 1950-70 over the population in 1939. Regression model (1) uses the variation within local labor markets, and model (2) uses the variation between local labor markets (see Section 3 for details). Each regression in Panel A., C., and D. includes our standard set of control variables, i.e., population density in 1939, the employment share in agriculture in 1939, turnover per capita in 1935, the share of damaged dwellings, and a dummy for counties within 75 km of the inner-German border. Regressions in Panel B. include our standard set of control variables but replace the standard covariates for wartime destruction and urbanization by alternative covariates. Regressions in Panel C. add different measures of pre-war population growth to the set of control variables. Panel D. estimates weighted regressions, using the 1939 population as weight. Robust standard errors clustered at the level of local labor markets are in brackets. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

First, we use alternative measures of wartime destruction and urbanization, our two key control variables. In our baseline analysis, we use the share of damaged dwellings as a measure of war destruction, and pre-war population density as a measure of urbanization. As a robustness check, we instead use rubble in 1945 per inhabitant in 1939 and a categorical variable that ranges from 1 ”no destruction” to 4 ”heavy destruction” as alternative measures for war destruction, and the share of population in bigger cities as an alternative measure for urbanization (see Section 3 for details). Panel B. of Table 3 shows that our results remain robust to the use of these alternative measures of war destruction and urbanization.

Second, we add different measures of pre-war population growth to our set of controls. Pre-treatment trends in population dynamics, if correlated with expellee inflows in 1950, may confound our estimates of the effect of expellees on post-war population dynamics. As shown in Panel C. of Table 3, however, our findings also prove robust to the addition of such controls.

Finally, we estimate weighted regressions, using population in 1939 as weights. Again, our results prove robust (see Panel D. of Table 3).

4.3 Net Migration Rate 1950-70

So far, we have considered the effect of the expellee inflow on post-war population growth. Our findings show that the migration-induced population shock had a persistent effect on the distribution of population within local labor markets, but only a temporary effect on the distribution of population between such labor markets. We had hypothesized ex ante that such a pattern is to be expected because migration flows of sufficient magnitude to significantly attenuate the population shock should only occur between but not within local labor markets. Until now, however, we have not produced any direct evidence that post-displacement migration flows are indeed consistent with our interpretation. In this subsection, we provide such direct evidence to further substantiate our hypothesis and corroborate our reading of the evidence.

TABLE 4: Expellee Inflows and Net Migration Rates 1950-70

	OLS			IV		
	Overall (1)	Within (2)	Between (3)	Overall (4)	Within (5)	Between (6)
Expellee inflow rate 1950	-0.340***	0.012	-0.663***	-0.513***	-0.166*	-0.788***
	(0.128)	(0.101)	(0.159)	(0.119)	(0.095)	(0.094)
R-squared	0.424	0.425	0.574	0.418	0.420	0.568
Observations	511	511	157	511	511	157

Notes: The dependent variable is net migration between 1950 and 1970 over the population in 1939. Each regression includes our standard set of control variables, i.e., population density in 1939, the employment share in agriculture in 1939, turnover per capita in 1935, the share of damaged dwellings, and a dummy for counties within 75 km of the inner-German border. Regression models (1) and (4) use the overall variation in the data, whereas models (2) and (5) use only the variation within local labor markets, and models (3) and (6) the variation between local labor markets (see Section 3 for details). The IV regressions in columns (4) to (6) use the expellee inflow rate in 1946 as an instrument for the expellee inflow rate in 1950. Robust standard errors are in brackets. Standard errors are clustered at the level of local labor markets in models (1), (2), (4), and (5). *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Specifically, we regress the net migration rate in 1950-70, defined as net migration in 1950-70 over population in 1939, on the expellee inflow rate in 1950 and our standard set of controls (see Table 4). As before, we run both OLS and IV regressions, exploiting either the overall, within, or between variation in expellee inflow rates. The estimated coefficients of the expellee inflow rate in 1950 have the same sign and are close in magnitude to the corresponding coefficients in our baseline regressions reported in Table 2. This suggests that post-displacement migration flows do indeed explain a very large share of the overall effect that expellee inflows had on post-war population growth in 1950-70, both overall and between local labor markets.[13] The estimated coefficient of the expellee inflow rate in 1950 in the IV within regression (column (5)), while marginally significant at the 10% level, is but a fifth in magnitude of that of the corresponding coefficient estimate in the IV between regression (column (6)). Post-displacement migration flows hence depend much more heavily on variation in expellee inflow rates between than within local labor markets.

These results have implications also for the literature that studies the link between immigrant inflows and population outflows. In particular, we showed that expellee inflows and net population flows are much more correlated between than within local labor markets. Previous work suggests that expellees were particularly mobile and thus responsible for a disproportionate share of population movements (Braun and Kvasnicka, 2014; Braun and Weber, 2016). Consistent with these earlier findings, Figure II in the Appendix shows that the distribution of expellee population shares at county level was much less dispersed in 1961 than in 1950 (the standard deviation decreased from 0.093 in 1950 to 0.063 in 1961). Expellees were more equally distributed in 1961 than in 1950, as they moved in disproportionate numbers from regions with high expellee inflows to regions with low expellee inflows.

4.4 Alternative Units of Observation

We conclude by highlighting once more–but in an alternative and more direct way of exposition that also considers an additional and larger regional unit than the local labor market–the importance of the unit of observation for the estimated effect of the expellee inflow on subsequent population growth. We have already shown that moving from counties to local labor markets as the unit of observation considerably increases the absolute magnitude of the coefficient on the expellee inflow rate in both OLS and IV regressions. Panels A. and B. of Table 5 reproduce these earlier results from Table 2 (columns (1), (3), (4), and (6)).[14] Panel C. adds a third level of regional aggregation, and estimates our standard regression at the level of *Raumordnungsregionen*, of which there are 36 in post-war West Germany. *Raumordnungsregionen* are also based on a functional definition, but cover a larger set of counties than local labor market regions.

Table 5 shows that at each level of aggregation, the expellee inflow rate exerts a negative impact on population growth in 1950-70. Most importantly, however, the absolute magnitude of this effect increases considerably with the level of aggregation. It is lowest for counties (Panel A.), i.e., the smallest unit considered, and highest for *Raumordnungsregionen* (Panel C.), the largest aggregation level. Local labor markets (Panel B.) fall in between these two, both in terms of aggregation level and in the absolute size of the estimated effect. The higher the level of aggregation, therefore, the less persistent proves the initial population shock. In fact, the IV coefficient estimate of -0.960 for *Raumordnungsregionen* suggests that at this largest aggregation level considered, the initial population shock was completely reversed by 1970.

[13] The net migration rate is one component of total population growth. The latter is made up of the sum of net migration and net natural changes of population. Since we normalize both population growth and net migration by population in 1939, coefficients in Tables 2 and 4 are directly comparable.

[14] Note that the between variation at the county level equals the total variation at the level of local labor markets.

TABLE 5: Alternative Units of Observation

	OLS	IV
	(1)	(2)
A. Counties (N=511)	-0.311**	-0.498***
	(0.140)	(0.130)
B. Local Labor Markets (N=157)	-0.671***	-0.830***
	(0.202)	(0.155)
C. Raumordnungsregionen (N=36)	-0.832***	-0.960***
	(0.188)	(0.187)

Notes: The table reports OLS and IV estimates of the effect of the expellee inflow rate in 1950 on population growth in 1950-1970. Each cells reports estimates from a separate regression. The dependent variable is the change in population between 1950-70 over the population in 1939. Control variables are population density in 1939, the employment share in agriculture in 1939, turnover per capita in 1935, the share of damaged dwellings, and a dummy for counties within 75 km of the inner-German border. Panel A. considers the 511 counties in West Germany, Panel B. the 157 local labor markets, and Panel C. the 36 *Raumordnungsregionen* in West Germany. Robust standard errors are in brackets. Standard errors in Panel A. are clustered at the level of local labor markets. *,**, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

5 Conclusion

What explains the spatial distribution of economic activity across space? The economic literature provides two main explanations. A first group of theories stresses the importance of locational fundamentals, or first-nature characteristics. These theories predict that fundamentals uniquely determine the distribution of economic activity. A second group of theories stresses the importance of second-nature characteristics and economies of scale. These theories predict that temporary shocks can have permanent effects on the spatial distribution of economic activity by shifting the economy from equilibrium to another. These two explanations are not necessarily mutually exclusive, and previous studies have found empirical support for both explanations.

In this paper, we have explored the importance of local labor markets for the persistence of a major population shock, the inflow of eight million expellees to different parts of West Germany after World War II. Our results show that the estimated regional persistence of this shock depends crucially on the type of regional unit considered and the type of variation in expellee inflows exploited. The migration-induced population shock proved persistent only within but not between local labor markets.

Consistent with our conjectures, this systematic difference indicates that locational fundamentals determine the spatial distribution of population across larger, inter-connected regional units, but not within these larger units. This insight can help to better understand the disparate findings in the literature on the persistence of population shocks. It also suggests that the choice of the regional unit should be carefully motivated when drawing conclusions from the persistence of population shocks about the determinants of the spatial distribution of economic activity.

References

Bauer, T., Braun, S., and Kvasnicka, M. (2013). The economic integration of forced migrants: Evidence for post-war Germany. *The Economic Journal*, 123(571):998–1024.

Borjas, G. J. (2006). Native internal migration and the labor market impact of immigration. *Journal of Human Resources*, 41(2):221–258.

Boustan, L. P., Fishback, P. V., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the Great Depression. *Journal of Labor Economics*, 28(4):719–746.

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. and Kvasnicka, M. (2014). Immigration and structural change: Evidence from post-war Germany. *Journal of International Economics*, 93(2):253–269.

Braun, S. and Omar Mahmoud, T. (2014). The employment effects of immigration: Evidence from the mass arrival of German expellees in postwar Germany. *The Journal of Economic History*, 74:69–108.

Braun, S. and Weber, H. (2016). How do regional labor markets adjust to immigration? A dynamic analysis for post-war Germany. Kiel Working Papers 2025, Kiel Institute for the World Economy (IfW).

Braun, S. T. and Dwenger, N. (2017). The local environment shapes refugee integration: Evidence from post-war Germany. Discussion Paper Series, Department of Economics 201711, Department of Economics, University of St. Andrews.

Burchardi, K. B. and Hassan, T. A. (2013). The economic impact of social ties: Evidence from German reunification. *The Quarterly Journal of Economics*, 128(3):1219 – 1271.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

Card, D. and DiNardo, J. (2000). Do immigrant inflows lead to native outflows? *American Economic Review*, 90(2):360–367.

Connor, I. (2007). *Refugees and expellees in post-war Germany*. Manchester University Press.

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289.

Davis, D. R. and Weinstein, D. E. (2008). A search for multiple equilibria in urban industrial structure. *Journal of Regional Science*, 48(1):29–65.

Deutscher Städtetag (1949). Statistisches Jahrbuch deutscher Gemeinden. Volume 37.

Douglas, R. M. (2012). *Orderly and humane: The expulsion of the Germans after the Second World War*. New Haven: Yale University Press.

Eder, C. and Halla, M. (2016). The long-lasting shadow of the Allied occupation of Austria on its spatial equilibrium. *IZA Discussion Papers 10095*.

Head, K. and Mayer, T. (2003). The empirics of agglomeration and trade. Working Papers 2003-15, CEPII Research Center.

Henderson, J. V. (1974). The sizes and types of cities. *American Economic Review*, 64(4):640–56.

IfW (1974). Infrastruktur, räumliche Verdichtung und sektorale Wirtschaftsstruktur als Bestimmungsgründe des regionalen Entwicklungspotentials in den Arbeitsmarktregionen der Bundesrepublik Deutschland. Kiel.

Institut für Raumforschung (1955). *Die Kreismappe des Instituts für Raumforschung: Statistische Übersichten*, volume various volumes. Bonn: Institut für Raumforschung.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–99.

Miguel, E. and Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, 96(1):1 – 15.

MPIDR [Max Planck Institute for Demographic Research] and CGG [Chair for Geodesy and Geoinformatics, University of Rostock] (2011). MPIDR Population History GIS Collection (partly based on Hubatsch and Klein 1975 ff. and Bundesamt für Kartographie und Geodäsie 2011). Rostock.

Müller, W. and Simon, H. (1959). *Die Vertriebenen in Westdeutschland. Ihre Eingliederung und ihr Einfluss auf Gesellschaft, Wirtschaft, Politik und Geistesleben*, chapter 'Aufnahme und Unterbringung', pages 300–346. Ferdinand Hirt.

Peters, M. (2017). Refugees and local agglomeration: Evidence from Germany's post-war population expulsions. Mimeo.

Redding, S. J. (2010). The empirics of new economic geography. *Journal of Regional Science*, 50(1):297–311.

Redding, S. J. and Sturm, D. M. (2008). The costs of remoteness: Evidence from German division and reunification. *American Economic Review*, 98(5):1766–1797.

Sarvimäki, M. (2011). Agglomeration in the periphery. *SERC Discussion Papers 80*.

Schulze, R. (2011). *The Disentanglement of Populations. Migration, Expulsion and Displacement in Postwar Europe, 1944-49*, chapter Forced Migration of German Populations during and after the Second World War: History and Memory, pages 27–47. Basingstoke: Palgrave Macmillan.

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.

Statistisches Amt des Vereinigten Wirtschaftsgebietes (1950). Die Flüchtlinge in Deutschland, Ergebnisse der Sonderauszählungen aus der Volks- und Berufszählung vom 29. Oktober 1946.

Statistisches Bundesamt (1952). Die Bevölkerung der Bundesrepublik Deutschland nach der Zählung vom 13.9.1950. *Statistik der Bundesrepublik Deutschland*, 35.

Statistisches Bundesamt (1956). Gebäude- und Wohnungszählung in der Bundesrepublik Deutschland vom 13. September 1950. *Statistik der Bundesrepublik Deutschland*, 38.

Statistisches Bundesamt (1974). *Volkszählung vom 27. Mai 1970*. Verlag W. Kohlhammer.

Statistisches Reichsamt (1939). Umsatzsteuerstatistik 1935. *Statistik des Deutschen Reichs*, 511.

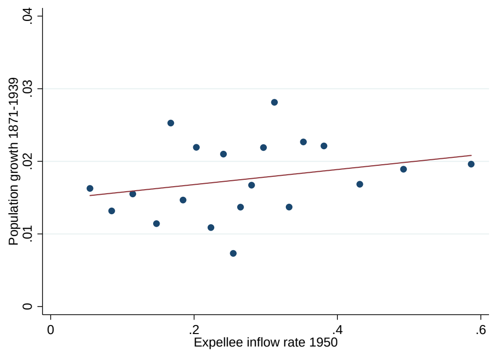
Statistisches Reichsamt (1940). *Statistisches Jahrbuch für das Deutsche Reich 1939/40*. Berlin: Statistisches Reichsamt.

Statistisches Reichsamt (1943). Berufszählung. Die Berufstätigkeit der Bevölkerung in den Reichsteilen. Volks-, Berufs- und Betriebszählung vom 17. Mai 1939. *Statistik des Deutschen Reichs*, 557.

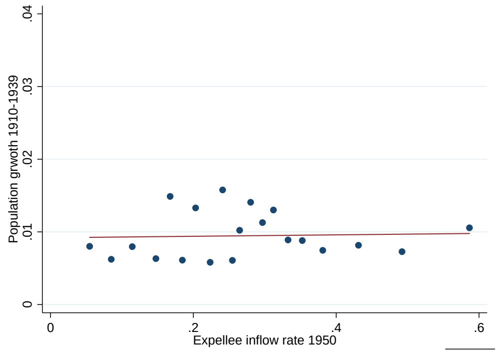
A Binned Scatter Plots – Expellee Inflows and Pre-war Population Growth

FIG. I: Binned Scatter Plots (Conditional)

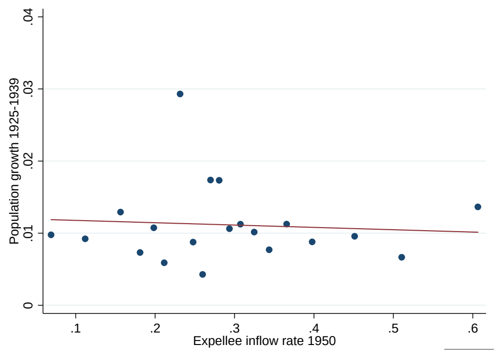
(a) Population growth 1871-1939



(b) Population growth 1910-1939



(c) Population growth 1925-1939



Notes: The figures in Panel (a), (b) and (c) depict binned scatter plots of residualized population growth in 1871-1939, 1910-1939, and 1925-1939 and residualized expellee inflow rates in 1950, grouping expellee inflow rates into 20 equal-sized bins. Covariates include our standard set of control variables, i.e., population density in 1939, the employment share in agriculture in 1939, turnover per capita in 1935, the share of damaged dwellings, and a dummy for counties within 75 km of the inner-German border. See Table 1 in the main text for the corresponding regression results.

B Kernel Density Estimates of Expellee Share in 1950 and 1961

FIG. II: Kernel Density Estimates

Notes: The figure shows Kernel density estimates of the expellee population share at county level on 17 September 1950 (solid line) and 6 June 1961 (dashed line).

C Merging of Counties

The administrative borders of some West German counties changed between 1939 and 1970. In order to make county borders comparable over time, we follow the procedure outlined in Braun and Dwenger (2017) (Appendix A) for changes between 1939 and 1950. We replicate their description in the following and extend the list of counties merged to also account for border changes between 1950 and 1970.

We first merge counties which, at any time between 1939 and 1970 formed one county. The counties of Hildesheim and Marienburg, for instance, were separate entities in 1939, but were merged to join the new county of Hildesheim-Marienburg in 1946. Consequently, the 1946 and 1950 censuses only contain data on Hildesheim-Marienburg. We thus merge Hildesheim and Marienburg already in the 1939 census. We proceed analogously for the counties of Bremerhaven and Wesermünde; city and rural districts of Bremen; Rhein-Wupper Kreis and Leverkusen; Kreis der Eder, Kreis des Eisenberges and Kreis der Twiste; city and rural districts of Konstanz; Coburg and Rodach bei Coburg; city and rural districts of Dinkelsbühl; city and rural districts of Donauwörth; city and rural districts of Göttingen; Gifhorn and Wolfsburg; Kempen-Krefeld and Viersen; city and rural districts of Herford; city and rural districts of Lüdenscheid; city and rural districts of Siegen.

In addition, there were some smaller border changes, in which municipalities were moved from one county to another. To deal with these border changes, we first compare the 1939 population of each county in its 1950 borders to the 1939 population of the same county in its 1939 borders. Since the majority of administrative borders remained unchanged between 1939 and 1950, the 1939 population figure is usually the same regardless of whether we use 1939 or 1950 borders. Moreover, we do not take any action if the difference between the two population figures is less than 5%. If the difference is larger than 5%, we merge the counties that exchanged municipalities. This applies to the counties of Osterholz, Verden and Bremen; Bergstraße, city and rural districts of Worms; Goslar, Wolfenbüttel and Salzgitter; Mainz, Groß-Gerau and Wiesbaden; Böblingen, Eßlingen and Stuttgart; city and rural districts of Osnabrück; city and rural districts of München; city and rural districts of Kulmbach; Lörrach and Neustadt; Norden and Emden; Braunschweig and Peine; city and rural districts of Erlangen; Sinsheim and Heilbronn; ; city and rural districts of Schwabach; Grevenbroich and Kempen-Krefeld; Bonn and Rhein-Siegkreis; Bielefeld, Paderborn and Wiedenbrück; Detnold and Höxter; Hamm and Unna; Meschede and Olpe; Beckum and Soest; city and rural districts of Ingolstadt.

Finally, we drop counties that have lost or gained more than 5% of its 1939 population to regions outside West Germany, in particular to counties in the Soviet Occupation Zone. These counties include Blankenburg (Rest); Helmstedt; Birkenfeld; Zweibrücken; Saarburg; Trier; Mellrichstadt; Osterode; rural and city districts of Lüneburg.

D Data sources

TABLE B1: Data Sources

Variable	Description and data source
Dependent variables	
Population growth 1950-70	Population change 1950-70 over population in 1939, based on Statistisches Bundesamt (1974).
Migration rate 1950-70	Net migration 1950-70 over population in 1939, based on Statistisches Bundesamt (1974).
Main explanatory and instrumental variable	
Expellee inflow rate 1950	Expellees in 1950 over the population in 1939, based on Statistisches Bundesamt (1952).
Expellee inflow rate 1946	Expellees in 1946 over the population in 1939, based on Statistisches Amt des Vereinigten Wirtschaftsgebietes (1950).
Control variables	
Share of damaged dwellings	Share of dwellings built before 1945 that were damaged in the war, based on Statistisches Bundesamt (1956).
Rubble per capita	Untreated rubble at the end of the war over the population in 1939, based on Deutscher Städtetag (1949).
Loss in housing space	Classifies the loss in housing space in four categories, ranging from 'no losses' (1) to 'very substantial losses' (4). This indicator is taken from Institut für Raumforschung (1955).
Pop. density 1939	Population in 1939 per square kilometer, based on Statistisches Bundesamt (1974).
Population share in cities with at least 10,000 inhabitants in 1939	The 1939 share of population living in cities with at least 10,000 inhabitants, based on Statistisches Reichsamt (1940).
Share agriculture 1939	The share of the workforce in agriculture in 1939, based on Statistisches Reichsamt (1943).
Turnover p.c. 1935	Turnover in 1935, taken from Statistisches Reichsamt (1939), over the workforce in 1939, taken from Statistisches Reichsamt (1943).
0/1 Inner-German border	Dummy for whether a county is located within 75 kilometers of the inner-German border.
Population growth 1871-1939 (1910-39, 1925-39)	Population change 1871-1939 (1910-39, 1925-39) over the population in 1871 (1910, 1925), based on various publications of the statistical agencies of the federal states.