NBER WORKING PAPER SERIES

# MARKET SIZE AND SPATIAL GROWTH - EVIDENCE FROM GERMANY’S POST-WAR POPULATION EXPULSIONS

Michael Peters

Working Paper 29329
http://www.nber.org/papers/w29329

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2021

I hereby declare that I have no relevant or material financial interests that relate to the research described in the paper "Market Size and Spatial Growth - Evidence from Germany’s Post-War Population Expulsions" The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Market Size and Spatial Growth - Evidence from Germany’s Post-War Population Expulsions
Michael Peters
NBER Working Paper No. 29329
October 2021
JEL No. O11,O4,R11

**ABSTRACT**

Virtually all theories of economic growth predict a positive relationship between population size and productivity. In this paper I study a particular historical episode to provide direct evidence for the empirical relevance of such scale effects. In the aftermath of the Second World War about 8m ethnic Germans were expelled from their domiciles in Eastern Europe and transferred to West Germany. This inflow increased the German population by almost 20%. Using variation across counties I show that the settlement of refugees had a large and persistent effect on the size of the local population, manufacturing employment and income per capita. I show that these findings are quantitatively consistent with an idea-based model of spatial growth if population mobility is subject to frictions and productivity spillovers occur locally. The model implies that the refugee settlement increased aggregate income per capita by about 12% after 25 years and that the historical settlement rule triggered persistent industrialization of rural areas.

Michael Peters
Department of Economics
Yale University
28 Hillhouse Avenue
New Haven, CT 06511
and NBER
m.peters@yale.edu

A data appendix is available at http://www.nber.org/data-appendix/w29329

# 1 Introduction

Can increases in the size of the population raise productivity? There are ample theoretical reasons to believe that the answer to this question ought to be yes. Most theories of growth predict a positive relationship between innovation incentives and population size, standard models of international trade imply that larger countries benefit from variety gains and many models of development and economic geography incorporate exogenous agglomeration forces, presumably as a reduced form for such considerations. This paper exploits a particular historical episode to provide direct evidence for the quantitative importance of such scale effects.

The historical setting I am studying concerns the forced population expulsions in post-war Germany. At the end of the Second World War, the Governments of the US, the UK and Russia expelled millions of ethnic Germans from their domiciles in Eastern Europe and transferred them to West Germany and the Soviet Occupied Zone. The ensuing expulsion was implemented between 1945 and 1948 and represents one of the largest forced population movements in world history. By 1950, about 8m people had been transferred to West Germany. Given the population at the time, this amounted to an increase in the population of more than 20%.

In order to use this historical setting to estimate the relationship between population size and productivity, I proceed in three steps. First, I provide direct evidence on the link between the settlement of refugees and subsequent income growth and industrialization. To do so, I exploit the fact that counties in West Germany differed vastly in their exposure to the inflow of refugees and that the specifics of the historical allocation rule allow me to address the obvious endogeneity concern that the incoming refugees might have settled in locations with favorable growth prospects. Second, motivated by the historical context, I build a model of spatial growth, where individuals are mobile across space (subject to frictions) and local productivity evolves endogenously. Third, I use the cross-sectional estimates from step one to estimate the structural parameters of my theory and to quantify the productivity effects of population inflows both at the regional and aggregate level.

To estimate the cross-sectional relationship between refugee inflow and local economic development, I constructed a novel panel dataset for more than 500 West German counties since the 1930s from original historical sources. Two features of the refugee settlement allow me to use it as a shifter of local labor supply. First, the refugees were not free to settle in the location of their choice but the population transports were organized by the Military Governments of the US and the UK, the governing bodies of West Germany at the time. Second, the dominant consideration to allocate the inpouring refugees to particular regions was the availability of housing rather than future economic prospects. With millions of refugees being transferred, the population of West Germany in 1950 exceeded its pre-war level by about 20%. At the same time, the Allied bombing campaign had reduced the housing stock by almost 25% on average and in many cities by more than 75%. As a consequence, refugees were assigned to rural, low population density localities where housing was relatively abundant. These aspects of the historical setting allow me to tease out the exogenous component of the initial refugee allocation both by directly controlling for the determinants of the allocation rule and by using an instrumental variable strategy, which exploits the distance to the pre-war population centers in Eastern Europe.

My results imply a positive relationship between population size and local productivity. First, I show that the initial allocation of refugees was very persistent. Even decades after the initial settlement, counties that received more refugees were still substantially larger and the share of refugees was still higher. Second, I establish a robust positive relationship between the allocation of refugees and manufacturing employment growth in the 1950s and 60s. Third, I document that the inflow of refugees raised local productivity and that such gains accrued slowly over time: while the effect of refugee inflows on income per capita in 1950 is statistically indistinguishable from zero, it is positive and large in later decades.

To rationalize these findings, I then propose a model of spatial growth. I combine a standard model of economic geography with a canonical idea-based growth model. The growth-part of the theory delivers an explicit model of regional productivity, which is determined endogenously and responds positively to the size of the local workforce. The geography-part of the theory generates an endogenous law-of-motion for the spatial distribution of population. If spatial mobility is subject to frictions, both local productivity and regional population size are slow-moving state variables that evolve jointly in equilibrium.

The model highlights an important distinction between the short-run and the long-run elasticity of productivity with respect to population size. The short-run elasticity describes the relationship between productivity and the local population holding current productivity constant. This elasticity depends on the elasticity of substitution across varieties and is isomorphic to agglomeration externalities commonly used in quantitative models of economic geography. By contrast, the long-run elasticity describes the relationship between local productivity and the local population along a spatial balanced growth path. The crucial parameter for the size of the long-run elasticity is the - what I call - inter-temporal knowledge elasticity, which determines how quickly the costs of creating new ideas decline in the existing stock of ideas. If this elasticity is positive, the long-run elasticity exceeds the short-run elasticity as the dynamic accumulation of local productivity acts as an amplifying force. Moreover, if this elasticity is sufficiently large and mobility subject to frictions, population shocks can have persistent effects.

I structurally estimate the model using the empirical variation from the natural experiment. The main moments of interest are the relationship between refugee inflows and income per capita and population size at different time horizons, the spatial persistence of the refugee population, and the response of local manufacturing employment. My empirical estimates imply that moving frictions were substantial and that the dynamic amplification of the initial shock was powerful. The long-run scale elasticity is almost 4 times as large as the short-run scale elasticity. Even though the estimated knowledge elasticity is small enough so that I can comfortably reject the case of non-stationarity, my estimates imply that the refugee settlement had persistent effects: the economy converges to a unique stationary equilibrium, which, however, is determined by initial allocation of refugees.

Finally I use the model to quantify the aggregate and spatial consequences of the refugee settlement. The combination of decreasing returns to scale in the agricultural sector and increasing returns to scale in the manufacturing sector imply that the effect is a priori ambiguous. It is also not identified from the cross-sectional estimates because of general equilibrium interactions. I find that the inflow of refugees reduced income per capita by about 3% in the short-run but increased it by about 12% after 25 years. Moreover, the model also implies that the policy of the Military Government to settle refugees in less developed, agriculturally specialized locations triggered a persistent increase in industrialization in such rural areas.

**Related Literature** The paper is related to a large literature in economic growth, which highlights the importance of market size effects (see, for example, the survey articles by Jones (2005) or Akcigit (2017)). Of particular relevance is the semi-endogenous growth model by Jones (1995), which highlights the importance of the inter-temporal knowledge elasticity to distinguish models of endogenous and semi-endogenous growth in the time series. My empirical results based on cross-sectional data are consistent with models of semi-endogenous growth where changes in population size affect the level of productivity but not the long-run growth rate. Recent papers that focus on nexus between population and productivity growth are Jones (2019) and Peters and Walsh (2020).

The paper also contributes to a recent literature on dynamic models of trade and economic geography. Desmet et al. (2018), Desmet and Rossi-Hansberg (2014), Nagy (2017) and Walsh (2019) also present models where local productivity is endogenously determined and responds to changes in local population size. This dynamic interaction between spatial mobility and local productivity, in particular the potential for shocks to have persistent effects, is also studied in Allen and Donaldson (2020), albeit in a more reduced form way. With respect to these studies, the main contribution of my paper is the explicit link to a natural experiment that generates large local changes in labor supply. A dynamic model of trade and migration is also analyzed in Caliendo et al. (2019), who however assume that regional productivity is exogenous.

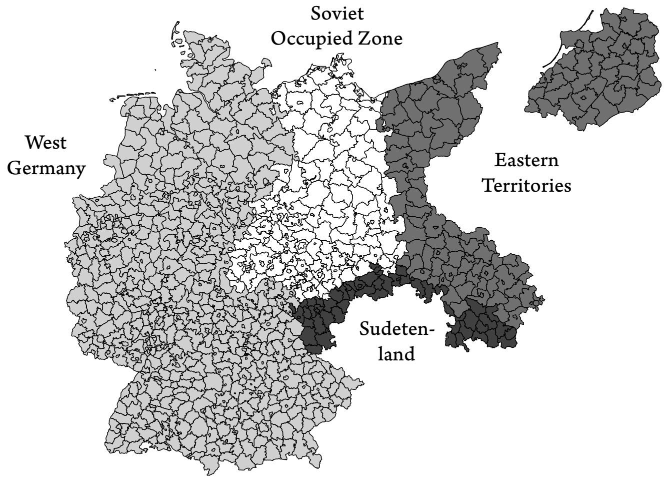
There is also a close connection to the large economic geography literature that often relies on exogenous agglomeration economies - see for example Ahlfeldt et al. (2015), Ramondo et al. (2016), Faber and Gaubert (2019) or the recent survey by Redding and Rossi-Hansberg (2017). These reduced-form specifications imply that scale elasticities are stable and time-invariant. My theoretical and empirical results highlight that such scale elasticities differ substantially in the short- and long-run. This finding is reminiscent of the literature on directed technological change, which also stresses the difference between short- and long-run elasticities (Acemoglu, 2002, 2007).

The paper also speaks to the literature on the long-run effects of immigration. The majority of contributions are concerned with the short-run impact of immigrants within local labor markets (see e.g. Card (1990), Burstein et al. (2017), Dustmann et al. (2017) or Peri (2016) for a survey). Exceptions are Sequeira et al. (2019), Burchardi et al. (2019), Bazzi et al. (2016), Bazzi et al. (2019) or Hornung (2014), which however are mostly empirical in nature and do not attempt a structural analysis.

Finally, various papers use the German context as a source of historical experiments.[1] Burchardi and Hassan (2013) use the fall of the Berlin Wall to measure the importance of social ties, Ahlfeldt et al. (2015) exploit the partition of Berlin to estimate agglomeration forces within city-blocks in Berlin and Redding and Sturm (2008) use the division of Germany as a shift in market access. The post-war population expulsions, which are the focus of this paper, have also been analyzed in Braun and Mahmoud (2014) and Braun and Kvasnicka (2014). These contributions, however, do not focus on the effect on local productivity.

The remainder of the paper is structured as follows. In the next section I describe the historical setting and the initial settlement of refugees in West Germany. Section 3 contains the main empirical analysis. In Section 4 I develop the theoretical model, which I estimate in Section 5. Section 6 concludes. An Appendix contains derivations of the main theoretical results, a variety of robustness checks and additional empirical results.

[1]See Fuchs-Schündeln and Hassan (2015) and Nakamura and Steinsson (2018) for recent surveys on the use of well-identified experiments to identify macroeconomic models.




Notes: The figure shows the German Reich in the boundaries of 1939. The light grey shaded part in the west is the area of to-be West Germany. The medium-grey shaded parts in the east are the Eastern Territories of the German Reich. The dark shaded area in the south-east is the Sudetenland, which used to be part of Czechoslovakia and was annexed by Germany in 1938. The white shaded part in the middle is the area of the Soviet Occupied Zone, i.e. the to-be GDR. The intra-regional spatial units are counties.

Figure 1: The German Reich in 1939

# 2 The Historical Setting

### The Presence of Ethnic Germans in Eastern Europe before 1939

The presence of ethnic Germans in Eastern Europe dates back to the Middle Ages.[2] The geography of the German Reich in 1939 is shown in Figure 1. In the west, shown in a light shade, is the area which is going to become West Germany in 1949. In the far east, shown in medium dark, are the Eastern Territories that encompassed the regions of East Prussia and Silesia and are part of todays' Poland and Russia. In the south-east, shown in dark, is the so-called Sudetenland. This part in the north of Czechoslovakia has a long tradition of German settlements and was annexed by the Nazi Government in 1938. Finally, the light shaded area in the middle will become the Soviet Occupied Zone (in 1945) and then turn into the German Democratic Republic (in 1949).

As shown in Table 1, on the eve of the Second World War about 17m Germans inhabited regions to the east of what is Germany today. Roughly 13.5m people lived in the Eastern Territories and the Sudetenland. In addition, there were sizable German minorities in other countries of Eastern Europe like Poland, Hungary and Romania.

[2]For recent historical treatments of this episode I refer to Douglas (2012) or Kossert (2008).

| East German Territories | Czechoslovakia | Hungary | Romania | Poland | Others | Total |
|---|---|---|---|---|---|---|
| 9.6m | 3.5m | 0.6m | 0.8m | 1m | 1.4m | 16.9m |

Notes: The table shows the ethnic German population in different regions in Eastern Europe in 1939. The category "Others" comprises Danzig, the Baltic States and Yugoslavia. Source: Federal Statistical Office (1953, p. 3)

Table 1: The German Population in Eastern Europe in 1939

In terms of their economic structure, West Germany and the areas in the East differed substantially, because the East had a comparative advantage in agriculture. In 1939, West Germany had an agricultural employment share of 27% and more than 40% of the local population worked in manufacturing. In the Eastern Territories and the Sudetenland, the agricultural sector was still the dominant source of employment and comprised more than 37% of the workforce.

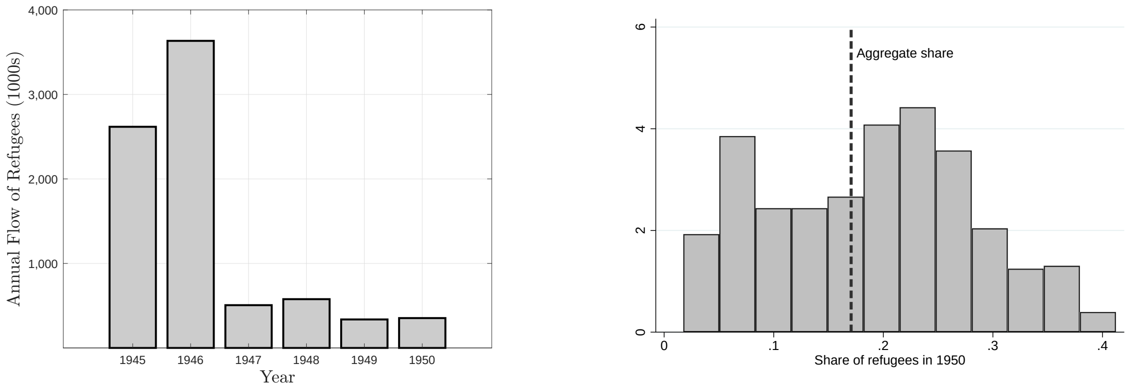
### The Expulsions and the Potsdam Conference in 1945

The Second World War brought an abrupt end to the presence of ethnic Germans in Eastern Europe because the entire German population either fled or was expelled in the aftermath of the war. This population transfer, where roughly 12 million ethnic Germans were forced to leave their domiciles, is one of the largest in world history.

The expulsion can be broadly divided into three phases. The first wave of refugees arrived in West Germany during the last months of the war when Soviet forces made their appearance at the eastern German border. After the German defeat in May 1945, the so-called wild expulsions started. These took place in the spring and summer of 1945 mainly in Poland and Czechoslovakia, where both the army and privately organized militias systematically expelled the remaining German population. It is only at the Potsdam Conference in the Summer of 1945, that the Military Governments of the US, UK and Russia tried to put an end to these unorganized expulsions and legalized them ex-post.[3] In the official protocol of the conference they noted that *"the Three Governments, having considered the question in all its aspects, recognize that the transfer to Germany of German populations, or elements thereof, remaining in Poland, Czechoslovakia and Hungary, will have to be undertaken. They agree that any transfers that take place should be effected in an orderly and humane manner."* Within the following two years, the majority of the ethnic German population was transferred from Eastern Europe to West Germany and the Soviet Occupied Zone.[4]

[3] The Potsdam Conference took place from 17 July to 2 August 1945. In addition to the expulsion of the German population, the governments of Russia, the US and the UK also decided (among other things) on the redrawing of Germany's eastern border, the trials of the German war criminals, the division of Germany and Austria into different occupation zones and the payments of war reparations.

[4] Becker et al. (2020) study the impact of the population transfer in Poland. They focus on the polish




Notes: The left panel shows the flow of refugees arriving in West Germany in thousands. The right panel shows the distribution of the share of refugees across counties in West Germany.

Figure 2: The Settlement of Refugees in West Germany

| Population | Population Losses 1939-50 | | | | Population Gains 1939-50 | | | Population |
|---|---|---|---|---|---|---|---|---|
| 1939 | Military Losses | Civilian Losses | Non-military Deaths | Others | Refugees | Inflows from SOZ | Births | 1950 |
| 39.3m | 2m | 0.4m | 5.2m | 0.5m | 7.9m | 1.5m | 7m | 47.6m |

Notes: The table reports aggregate population trends in West Germany between 1939 and 1950. "Inflows from SOZ" are individuals who fled the Soviet Occupied Zone. Source: Edding (1951, p. 2)

Table 2: The Population of West Germany: 1939 - 1950

In the left panel of Figure 2 I depict the flow of refugees that arrived in West Germany. By the end of 1946, almost 6m refugees had arrived in West Germany. Between 1947 and 1950 another half million refugees arrived per year. By 1950, the inflow of refugees had increased the population in West Germany by about 8m individuals.[5] To put this population inflow into perspective, Table 2 reports a decomposition of the population dynamics in West Germany between 1939 and 1950. From the initial population of about 40m in 1939, West Germany suffered military and civilian losses of about 2.5m during the Second World War. At the same time, the country saw the arrival of 8m refugees and an additional 1.5m people fleeing the Soviet Occupied Zone. Hence, despite the casualties during the war, the population of West Germany increased by 20%, between 1939 and 1950.

In terms of their demographic characteristics the refugee and native population were very similar. The first two panels of Table 3 show that refugees and natives had the same share of males and that their age distribution was almost identical. In the third panel, I show that

population that was re-settled in the areas from which the Germans were expelled.

[5] There are additional refugees coming into West Germany after 1950. These flows are not only much smaller in magnitude, but most of them moved to West Germany after an initial spell in the Soviet Occupied Zone after their expulsion from the Eastern Territories. As I will measure the initial allocation of refugees across Wester German counties in 1950, these continuing flows are not the focus of this paper.

the distribution of educational attainment was also very similar.

| | Male | Age distribution | | | Educational attainment | | | |
|---|---|---|---|---|---|---|---|---|
| | share | <15 | 20-65 | 65+ | Elem. school | High school | Vocational school | College |
| Natives | 46.5 | 20.4 | 68.5 | 11.1 | 66.8 | 26.3 | 4.9 | 1.9 |
| Refugees | 46.9 | 21.9 | 68.5 | 9.7 | 67.4 | 25.6 | 4.9 | 2.1 |

Notes: The first panels report the share of males and the age distribution in 1958. The last panel reports the distribution of educational attainment of the cohort born before 1920 as observed in the Census in 1970. These individuals were are at least 25 years old in in 1945 and hence completed their educational attainment prior to the expulsion. Source: Besser (2007).

Table 3: Characteristics of Refugees and Natives

### The Initial Settlement in West Germany

Upon their arrival in West Germany, the refugees were not free settle where they wanted to but their assignment was organized and implemented by the Military Governments of the US and the UK. They received the inflowing refugee treks, which arrived from Eastern Europe either by train or by foot, and allocated them across counties in West Germany. Moreover, in the immediate post-war period until the late 1940s, labor mobility was severely restricted and the Occupying Forces deployed armed forces at the state boundaries to prevent internal migration. William H. Draper, Director of the Economic Division of the Office of the Military Government of the US (OMGUS), observed that "Germany has been virtually cut into four Zones of Occupation - with the Zone borders not merely military lines, but almost air-tight economic boundaries" (Office of the Military Government for Germany, 1945, p. 10).

One consequence of these policies was that the settlement of refugees was strikingly unbalanced. According to the German historian Gerhard Reichling "there is no aspect where the Federal Republic of Germany shows a similar degree of heterogeneity as in the absorption and distribution of expellees" (Reichling, 1958, p. 17). This heterogeneity is depicted in the right panel of Figure 2, which shows the histogram of the local share of refugees across counties. In the aggregate, refugees amounted to roughly 18% of the population. However, this statistic hides substantial spatial heterogeneity: some counties received hardly any refugees and other counties received so many that their population almost doubled.

To appreciate this unequal spatial distribution, remember that an orderly settlement was an almost impossible task in war-torn Germany. A particular concern was the availability of housing as the rising population came hand in hand with a sharply diminished housing stock, which was heavily destroyed during the Allied Bombing Campaign.[6] Werner Nellner, one of the leading post-war economic historians, described the situation as follows: "In the midst

[6]About 23% of the aggregate housing stock was damaged during the Allied bombing campaign. Moreover, there is considerable heterogeneity as a large share of counties saw more than 60% of their housing stock damaged during the war (see Section A-2.3 in the Appendix).

of the chaotic post-war circumstances arrived the refugee transports. The entirely confusing political and economic situation paired with the abruptness of this pouring-in simply did not allow a sensible distribution of the expellees into areas where they could find work. The ultimate goal was to find shelter for those displaced persons" (Nellner, 1959, p. 73). This uncoordinated settlement was already considered an enormous problem at the time. As early as in 1946, P.M. Raup, Acting Chief of the Food and Agricultural Division of the Office of the Military Government of the US (OMGUS) complained that "both the planning and the execution of the support measures for German expellees was conducted entirely under welfare perspectives. The people in charge at the Military Government are social service officials. Similarly on the side of the German civil government, the department in charge is the social service agency. Entire communities are moved so that the population of some counties is increased by 25-30% and the agency in charge was founded to support the elderly, disabled people and the poor. ... The whole problem has not been handled as one of settlements of entire communities but as an emergency problem supporting the poor." (Grosser and Schraut, 2001, p. 85).

These descriptions of the refugee settlement are strikingly visible in the data. In Table 4 I report the results of a set of bivariate regressions of the share of refugees in 1950 on different pre-war county characteristics and state fixed effects and report the coefficients on the respective characteristics. In column 1, I show that the share of refugees is strongly negatively correlated with the population-weighted distance to the expulsion region (the "expulsion distance" $ED_c$), which I calculate as

$$ED_c = \ln\left(\sum_{r \in ER} d_{c,r} \times pop_r^{1939}\right), \tag{1}$$

where $d_{c,j}$ is the geographical distance between county $j$ and $r$, $ER$ denotes the set of expulsion regions, i.e. the areas, which the german population had to leave and $pop_r^{1939}$ is the size of their population in 1939. Hence, counties that were closer to the population centers of ethnic Germans in the pre-war period, experienced larger refugee inflows. This is exactly what one would expect if the Military Governments experienced an institutional overload in distributing the refugees, which kept pouring in at the eastern border.

In columns 2 and 3 I focus on the availability of housing. The share of refugees was much larger in regions with a low population density in the pre-war period and in counties, that experienced a larger destruction of their housing stock during the war. Hence, refugees were settled in rural and thus less developed locations. This is seen in the remaining columns of Table 4. A county's share of refugees is negatively correlated with the share of manufacturing employment (both in 1933 and 1939) and positively correlated with its agricultural employ-

| | Expulsion Distance | ln pop dens 1939 | War time Dest. | Manufac. share 1939 | Manufac. share 1933 | Ag. share 1933 | Rural share 1933 | GDP pc 1935 |
|---|---|---|---|---|---|---|---|---|
| $\beta$ | -0.159*** | -0.023*** | -0.190*** | -0.132*** | -0.109*** | 0.087*** | 0.080*** | -0.017*** |
| | (0.026) | (0.002) | (0.011) | (0.022) | (0.020) | (0.011) | (0.008) | (0.005) |
| $N$ | 536 | 536 | 536 | 535 | 523 | 523 | 536 | 523 |
| $R^2$ | 0.662 | 0.724 | 0.752 | 0.656 | 0.662 | 0.691 | 0.705 | 0.651 |

Note: Standard errors are clustered at the level of 37 larger administrative units (Regierungsbezirke). *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. Each column reports the coefficient $\beta$ of a regression $\mu_r = \delta_s + \beta x_r + u_r$ where $\mu_r$ is the share of refugees in 1950, $x_r$ are the different regional characteristics in the respective columns and $\delta_s$ is a set of state fixed effects. The wartime destruction in column 3 is measured as the share of the housing stock that was destroyed during the war.

Table 4: Spatial Correlates of Refugee Inflows

ment share. Moreover, counties with a larger share of refugees are more likely to be rural (as measured by the share of the population living in small cities) and have lower GDP per capita in 1935. My empirical strategy will take these systematic correlations into account.[7]

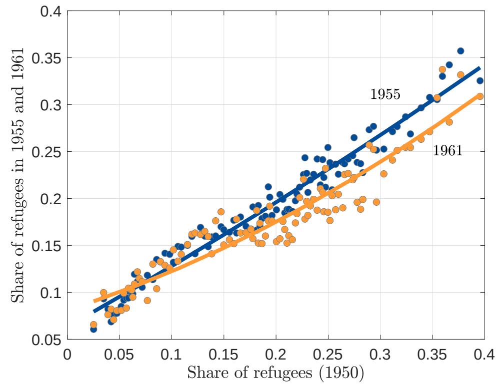
The economic consequences of this initial settlement naturally depend crucially on its persistence. If the refugees left their originally assigned areas relatively quickly, the impact would likely be small. This, however, was not the case. Figure 3, which shows the correlation between the share of refugees in 1950 and 1955 and 1961, highlights that counties with a large initial share of refugees still featured a large share in the subsequent decade. Hence the initial assignment was highly persistent.[8]

# 3 Refugees, Industrialization and Local Growth

Did these persistent shocks to the local population have important economic implications for the local economy? In this section I estimate the effects of refugee inflows on population growth, changes in sectoral specialization and growth in income per capita. These cross-sectional estimates form the backbone of my structural analysis because I estimate the structural parameters of my theory with indirect inference to fit these regression results.

[7]Even though the size of the refugee settlement was systematically correlated with local characteristics, there was very little spatial sorting of particular refugees into particular localities. If refugees had been spatially sorted by the government authorities, the composition of the settled refugees would vary systematically with the pre-war industrial make-up and one would expect refugees from the manufacturing-intensive Sudetenland to be sent to locations with a higher pre-war manufacturing share. This is not the case. As I show in Section A-2.3 in the Appendix, neither the manufacturing share nor GDP per capita predicts the composition of the refugee population. Importantly, this is not the case for individuals fleeing the Soviet Occupied Zone, who were not part of the organized refugee treks but were free to settle. These migrants do locate systematically in richer and more manufacturing intensive locations.

[8]For a subset of counties I also observe the share of refugees in 1946. This share is also strongly correlated with the share in 1950. A bivariate regression yields a coefficient of 0.91 with an $R^2$ of 0.952.




Notes: The figure shows the correlation between the share of refugees in 1950 and 1955 (blue) and 1961 (orange) as binned scatter plots for 100 percentiles of the refugee share in 1950.

Figure 3: The Regional Persistence of Refugee Inflows

## 3.1 Data

My analysis relies on a variety of novel historical datasets, many of which were digitized for this project. Using this data I constructed a spatially harmonized panel dataset for more than 500 counties in West Germany spanning the time-period from 1933 to the end of the 20th century.[9] The basis of my dataset is comprised of the population censuses for the years 1933, 1939, 1950 and 1961, which are published individually for each of the nine states. For each of these years, the publications report a variety of outcomes at the county-level like the level of population, sectoral employment shares, occupational employment shares, sex ratios and various other characteristics

I augmented this dataset with six additional pieces of information. The first concerns the regional allocation of refugees, which I digitized from a special statistical publication published in 1955 (Statistisches Bundesamt, 1955b). Secondly, in the 60s, 70s, 80s and 90s, the statistical offices from the respective German states constructed measures of GDP at the county-level. These results were published and could be digitized (Statistische Landesämter, 1968, 1976, 1992). I was not able to find county-level GDP measures for the pre-war and immediate post-war period. As a substitute I digitized county-level information from tax records, which report value added taxes for each county in 1935 (Statistisches Reichsamt, 1938) and 1950 (Statistisches Bundesamt, 1955a). I take these measures as being proportional to local GDP. However, for the structural estimation of my model, I also present results that do not rely on this information.

Fourth, I digitized the county-level results for four waves of the manufacturing census

[9] See Section A-2.1 in the Appendix for the detailed references and Section A-2.2 for details on the construction of time-invariant boundaries.

in 1933, 1939, 1950 and 1956 (Statistisches Bundesamt, 1957). They report the number of plants at the county-level and hence allow me to directly measure plant entry, which is an important theoretical mechanism of my model. Fifth, I provide new measures of the extent of war time destruction and regional housing supply at the county-level, which I digitized from the historical housing census conducted in 1950 (Statistisches Bundesamt, 1950).[10] Finally, I digitized the historical migration census from 1955, which reports inflows and outflows for each of the 500 counties (Statistisches Bundesamt, 1955c). This information is useful to estimate the structure of mobility costs in the quantitative model. To corroborate my baseline results at the county-level, I also digitized data for all 6000 cities and villages for the state of Bavaria (Bayerisches Statistisches Landesamt, 1944, 1952, 1963a) for the years 1939, 1950 and 1961. Like for the county data, I observe population growth, the share of refugees and sectoral employment at this more granular level of aggregation.

I complement my analysis with two micro datasets. The first is the Mikrozensus Zusatzerhebung 1971 (MZU 71), a special appendix to the census conducted in 1971 to measure social mobility. It includes identifiers on individuals' refugee status and retrospective information about individuals' employment characteristics in 1939, 1950, 1960 and 1971. The MZU 71 has roughly 200,000 observations, 40,000 of which are refugees. The second is the Einkommens- und Verbrauchsstichprobe 1962/63 (EVS 62), which is a micro dataset conducted in 1962 to measure household expenditure and hence similar to the Consumer Expenditure Survey in the US. It has about 32,000 observations and allows me to measure relative earnings of refugees and natives.

## 3.2 The Economic Effects of Refugee Inflows

To estimate the effects of the refugee settlement on the local economy, I focus on six outcomes: population growth, changes in the sectoral employment shares (for manufacturing, agriculture and services), growth in income per capita, and growth in the number of industrial plants, both in the short run and the long-run. I consider a specification of the form

$$z_{rt} - z_{r,\text{pre-war}} = \delta_s + \beta \mu_{r1950} + \alpha z_{r,\text{pre-war}} + \phi \ln \ell_{r1939} + \varphi \text{wd}_r + x_r' \zeta + u_r, \tag{2}$$

where $z_{rt}$ and $z_{r,\text{pre-war}}$ denote the respective outcome of interest at time $t$ and in the pre-war period and $\mu_{r1950}$ is the share of refugees in 1950. Furthermore, I control for a set of state fixed

[10]This data is different from the one used in Brakman et al. (2004) and Burchardi and Hassan (2013). These papers focus on the extent of wartime destruction in selected cities. The housing census contains information on war damages for each county covering the entire landmass of Germany. Because refugees were predominantly allocated to rural areas outside of cities, it is important to measure the extent of war-time destruction at the county level.

effects ($\delta_s$), population density in 1939 ($\ell_{r1939}$) and the extent of wartime destruction ($\text{wd}_r$), which are the important determinants of the housing supply (and hence refugee flows), and a set of additional spatial controls ($x_r$). I estimate this specification both via OLS and with an instrumental variable strategy (see Table 6 below). For brevity I only report the coefficient of interest $\beta$. In Section A-2.4.1 in the Appendix I also report the full results for all covariates for completeness.

Consider first the OLS results reported in Table 5. The six different panels refer to the six different outcomes of interest. The first four columns capture the short-run effect in 1950. The last four columns highlight the long-run effect in 1961. The different specifications include a varying extent of regional controls. Columns 1 and 5 only control for state fixed effects ($\delta_s$) and hence capture the unconditional correlation with refugee inflows. In columns 2 and 6 I control for initial population density, wartime destruction and the distance to the inner german border. These variables are important determinants of the refugee allocation and could be directly correlated with regional growth. I also control for the initial level of the dependent variable, $z_{r,\text{pre-war}}$, to allow for mean reversion. In the third and seventh column I include the pre-war levels of all the six dependent variables on the right hand side. Finally, in the last columns, I control for a host of additional pre-war characteristics at the district level, such as the average urbanization rate, population density and the manufacturing share in 1933 (in addition to 1939) and the regional GDP share of agriculture and manufacturing in 1935. Standard errors are clustered at the level of 37 *Regierungsbezirke,* the next larger administrative unit.

Table 5 paints a cohesive picture of the regional impact of refugee inflows. First, given the size and the persistence of such inflows shown in Figure 3, one would expect the initial allocation of refugees to be an important determinant local population growth. Panel A shows that this is the case: the semi-elasticity of 1.3 implies that an increase in the share of refugees by 10 percentage points increases the local population by 13%. Note that the short run elasticity is much higher if the extent of war-time destruction is not controlled for (see column 1). This reflects the negative correlation between refugee inflows and war-time destruction. Interestingly, the long-run elasticity in columns 5 to 8 is statistically identical across specifications and does not depend on whether the extent of war-time destruction is controlled for. This is consistent with the results of Davis and Weinstein (2002) and Brakman et al. (2004), who show that war-time destruction has a transitory effect on population size.

The following three panels document the stark sectoral reallocation in response to refugee inflows. The manufacturing employment share increases, the agricultural employment share decreases and the share of service employment is not affected. Moreover, this reallocation effect is not merely transitory but the manufacturing employment share is still systematically

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Panel A:* Population growth: $\ln L_{rt} - \ln L_{r1939}$ | | | | | | | |
| | 1939-1950 | | | | 1939-1961 | | | |
| Share of refugees in 1950 | 1.999*** | 1.359*** | 1.377*** | 1.428*** | 0.968*** | 1.029*** | 1.086*** | 1.219*** |
| | (0.106) | (0.112) | (0.107) | (0.097) | (0.139) | (0.211) | (0.182) | (0.157) |
| $N$ | 536 | 523 | 519 | 472 | 488 | 475 | 472 | 472 |
| $R^2$ | 0.610 | 0.683 | 0.698 | 0.732 | 0.173 | 0.175 | 0.283 | 0.338 |
| | *Panel B:* Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$ | | | | | | | |
| | 1939-1950 | | | | 1939-1961 | | | |
| Share of refugees in 1950 | 0.203*** | 0.317*** | 0.322*** | 0.353*** | 0.451*** | 0.241*** | 0.244*** | 0.255*** |
| | (0.064) | (0.074) | (0.075) | (0.054) | (0.053) | (0.086) | (0.087) | (0.073) |
| $N$ | 535 | 535 | 519 | 472 | 535 | 535 | 519 | 472 |
| $R^2$ | 0.301 | 0.393 | 0.423 | 0.539 | 0.230 | 0.351 | 0.357 | 0.424 |
| | *Panel C:* Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$ | | | | | | | |
| | 1933-1950 | | | | 1933-1961 | | | |
| Share of refugees in 1950 | -0.454*** | -0.186** | -0.228*** | -0.423*** | -0.716*** | -0.097 | -0.151** | -0.326*** |
| | (0.099) | (0.072) | (0.062) | (0.052) | (0.133) | (0.078) | (0.060) | (0.057) |
| $N$ | 523 | 523 | 519 | 472 | 523 | 523 | 519 | 472 |
| $R^2$ | 0.091 | 0.701 | 0.776 | 0.842 | 0.122 | 0.761 | 0.817 | 0.858 |
| | *Panel D:* Service employment: $\pi^S_{rt} - \pi^S_{r1933}$ | | | | | | | |
| | 1933-1950 | | | | 1933-1961 | | | |
| Share of refugees in 1950 | -0.089 | 0.014 | -0.059 | 0.051 | -0.098 | 0.017 | -0.054 | 0.057 |
| | (0.057) | (0.058) | (0.055) | (0.061) | (0.061) | (0.071) | (0.068) | (0.074) |
| $N$ | 523 | 523 | 519 | 472 | 523 | 523 | 519 | 472 |
| $R^2$ | 0.211 | 0.359 | 0.441 | 0.602 | 0.053 | 0.172 | 0.276 | 0.448 |
| | *Panel E:* GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$ | | | | | | | |
| | 1935-1950 | | | | 1935-1961 | | | |
| Share of refugees in 1950 | -1.219*** | -0.083 | -0.017 | -0.017 | 1.159*** | 0.502** | 0.658*** | 0.746*** |
| | (0.296) | (0.382) | (0.369) | (0.323) | (0.419) | (0.227) | (0.210) | (0.199) |
| $N$ | 523 | 523 | 519 | 472 | 519 | 519 | 515 | 468 |
| $R^2$ | 0.110 | 0.511 | 0.540 | 0.582 | 0.101 | 0.889 | 0.905 | 0.903 |
| | *Panel F:* Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$ | | | | | | | |
| | 1933-1950 | | | | 1933-1956 | | | |
| Share of refugees in 1950 | -0.450 | 0.726* | 0.653** | 0.817*** | -0.819 | 0.697 | 0.830* | 1.169*** |
| | (0.383) | (0.410) | (0.283) | (0.247) | (0.744) | (0.756) | (0.456) | (0.353) |
| $N$ | 520 | 520 | 519 | 472 | 520 | 520 | 519 | 472 |
| $R^2$ | 0.045 | 0.393 | 0.664 | 0.680 | 0.140 | 0.372 | 0.617 | 0.626 |
| State FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Pop. density (1939) | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Wartime destr. | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Geography | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Levels of dep. variable | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Pre-war controls | | | ✓ | ✓ | | | ✓ | ✓ |
| Addtl. pre-war controls | | | | ✓ | | | | ✓ |

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke.* *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The dependent variables are population growth (Panel A), changes in sectoral employment shares (Panels B - D), income per capita growth (Panel E) and the growth in the number of industrial plants (Panel F). The various specifications control for the share of destroyed housing stock ("Wartime destr."), the distance to the inner german border and a fixed effect for whether a county is a border county ("Geography"), the respective dependent variable in levels in the pre-war period ("Levels of dep. variable"), all six dependent variable in levels in the pre-war period in Panels A-F ("Pre-war controls") and the population share in cities with less than 2000 inhabitants in 1939, population density in 1933, the manufacturing share in 1933 and the GDP share in manufacturing and agriculture in 1935 ("Addtl. pre-war controls").

Table 5: The Effects of Refugee Inflows on the Local Economy

higher in the 1960s. This pattern of reallocation is exactly what one would expect if services are non-traded and agricultural production is subject to decreasing returns. In addition, as I will show below, refugees not only increased the size of the local population, but their labor supply was biased towards the manufacturing sector. Quantitatively, an increase in the share of refugees by 10 percentage points increases the manufacturing employment share by around 2.5 percentage points.

In panel E, I estimate the effect of the refugee settlement on income per capita growth. Columns 2 - 4 show that income per capita growth between 1935 and 1950 is essentially unrelated to the inflow of refugees. Columns 5 to 8 show that the relationship between refugee inflows and long-run income per capita growth is positive. According to these estimates, an increase in the share of refugees by 10% increases income per capita by roughly 5-6% after 15 years. Hence, the long-run effect is positive, suggesting a form of dynamic agglomeration.[11]

It is important to reiterate an important caveat for the interpretation of the results in Panel E. Because data on GDP per capita on the county level do not exist in 1950, I have to rely on value added taxes per capita. Hence, my measure of GDP capita differs between the long-run and the short-run specification. In Section 5.2 below I explicitly address this discrepancy by relying on additional data on GDP pc in the late 1950s, 1970s and 1980s.

In the structural model presented in Section 4, the positive effect on long-run GDP pc is rationalized through dynamic variety gains in the spirit of Romer (1990). Interestingly this mechanism also appears explicitly in the historical sources. In 1949, M. Bold, the Deputy Director of the US Military Government in Bavaria for example noted that "since refugees and bombed-out Bavarians now living in rural areas cannot move nearer to industrial jobs, such jobs must go to them. In fact many world famous industries wanting to reestablish in Bavaria have already sought locations in non-industrial areas near idle workers" (Office of the Military Government for Germany, 1949, p. 26). Panel F provides direct evidence for this mechanism by documenting that refugee inflows are correlated with an increase in the entry of manufacturing plants. Interestingly, and similar to the results for GDP pc in Panel E, the long-run elasticity is larger then the short-run elasticity. However, these differences are too small to detect statistically.

The results in Table 5 hinge on the assumption of parallel trends, i.e. local economic development would have been similar, conditional on the determinants of the refugee settlement.

[11]Note that the unconditional relationship between refugee inflows and income growth reported in columns 1 and 5 are different. This reflects the fact that income growth is systematically related to pre-war population density and that local income shows mean reversion. The main difference between columns 1 and 2 (columns 5 and 6) is the inclusion of $\ln \ell_{r1939}$ and $\ln y_{r1935}$. The coefficient on the share of refugees in a regression that controls for log population density in 1939 and log income per capita in 1935 is given by $-0.065$ (with a standard error of 0.298) for 1950 and 0.343 (with a standard error of 0.2) in 1961.

The stability of the coefficients across the different specifications is therefore reassuring. In Section A-2.4.2 in the Appendix I provide additional evidence for the plausibility of this assumption. First, I show that (conditional on population density in 1939) the share of refugees in 1950 is uncorrelated with sectoral employment shares in 1933 and 1939 and with population growth and growth in the number of industrial plants between 1933 and 1939. Moreover, the correlation with the change in the manufacturing employment share between 1933 and 1939 is, if anything, negative. Hence, there is no indication that counties with higher refugee inflows were on a more promising trajectory in the pre-war period. I also address the concern that pre-war population density might have had non-linear effects on future population growth and industrialization (see, e.g., Desmet and Rossi-Hansberg (2009)). The results in Table 5 are almost unchanged even when I control non-parametrically for pre-war population density and pre-war urbanization with detailed fixed effects.

**Instrumental Variable Estimates** As complementary evidence that these results reflect the causal effect of changes in the local refugee population I now present an instrumental variable strategy that exploits the systematic geographic variation between the share of refugees and the distance to the expulsion regions. More specifically, I estimate the same specification as reported in Table 5 but use the expulsion distance within states defined in (1) to instrument for the share of refugees in 1950. The results are contained in Table 6 whose structure exactly parallels the one of Table 5. For each of the six outcomes, I report the coefficient and standard error on the instrumented share of refugees in 1950 and the F-statistic. Again, I cluster standard errors at the level of the 37 Regierungsbezirke.

The results are very similar to the corresponding OLS estimates, both qualitatively and quantitatively. The semi-elasticity of population growth is slightly larger but not statistically different from the OLS estimates given the size of the standard errors. The effects on sectoral employment shares are also comparable: the manufacturing share increases, the agricultural share declines and the service-share is not significantly affected. As in the OLS, the IV strategy also finds a noisy and statistically insignificant effect on short-run income growth. The long-run effect is positive and the point estimate is - in the specifications with controls - similar to the OLS results.[12] Finally, the relationship between refugee inflows and plant entry is also positive, with the long-run elasticity generally exceeding the short-run elasticity.

The main concern with this identification strategy is that the distance to the expulsion regions is - by construction - correlated with the distance to the new inner german border.

[12]The reason why the unconditional correlation in column 5 differs between the OLS and the IV is that the IV specification only exploits the variation in the refugee share that is explained by the distance to the expulsion regions. Because counties with low initial population density grow faster on average, this form of regional convergence is captured in the OLS but less so in the IV.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *Panel A:* Population growth: $\ln L_{rt} - \ln L_{r1939}$ | | | | | | | | |
| | 1939-1950 | | | | 1939-1961 | | | |
| Share of refugees in 1950 | 1.897*** | 1.459*** | 1.563*** | 1.614*** | 1.018*** | 1.227*** | 1.450*** | 1.501*** |
| | (0.191) | (0.159) | (0.189) | (0.158) | (0.207) | (0.253) | (0.255) | (0.234) |
| *N* | 526 | 526 | 509 | 463 | 526 | 526 | 509 | 463 |
| F-Stat | 56.026 | 17.632 | 19.575 | 18.114 | 97.733 | 20.721 | 24.233 | 21.488 |
| *Panel B:* Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$ | | | | | | | | |
| | 1939-1950 | | | | 1939-1961 | | | |
| Share of refugees in 1950 | 0.124 | 0.271** | 0.297** | 0.406*** | 0.279*** | 0.199 | 0.222* | 0.333*** |
| | (0.114) | (0.118) | (0.122) | (0.064) | (0.082) | (0.135) | (0.134) | (0.098) |
| *N* | 535 | 535 | 519 | 472 | 535 | 535 | 519 | 472 |
| F-Stat | 97.785 | 28.434 | 23.443 | 21.888 | 97.785 | 28.434 | 23.443 | 21.888 |
| *Panel C:* Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$ | | | | | | | | |
| | 1933-1950 | | | | 1933-1961 | | | |
| Share of refugees in 1950 | -0.337*** | -0.441** | -0.573*** | -0.607*** | -0.261 | -0.294 | -0.449** | -0.472** |
| | (0.121) | (0.188) | (0.169) | (0.183) | (0.197) | (0.193) | (0.176) | (0.185) |
| *N* | 523 | 523 | 519 | 472 | 523 | 523 | 519 | 472 |
| F-Stat | 92.790 | 27.365 | 23.443 | 21.888 | 92.790 | 27.365 | 23.443 | 21.888 |
| *Panel D:* Service employment: $\pi^S_{rt} - \pi^S_{r1933}$ | | | | | | | | |
| | 1933-1950 | | | | 1933-1961 | | | |
| Share of refugees in 1950 | 0.146 | 0.307 | 0.228 | 0.188 | -0.007 | 0.271 | 0.202 | 0.143 |
| | (0.097) | (0.219) | (0.203) | (0.183) | (0.089) | (0.227) | (0.220) | (0.198) |
| *N* | 523 | 523 | 519 | 472 | 523 | 523 | 519 | 472 |
| F-Stat | 92.790 | 25.060 | 23.443 | 21.888 | 92.790 | 25.060 | 23.443 | 21.888 |
| *Panel E:* GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$ | | | | | | | | |
| | 1935-1950 | | | | 1935-1961 | | | |
| Share of refugees in 1950 | -0.400 | 0.221 | 0.341 | -0.003 | -0.671 | 0.208 | 0.517* | 0.471** |
| | (0.457) | (0.743) | (0.716) | (0.617) | (0.615) | (0.370) | (0.271) | (0.238) |
| *N* | 523 | 523 | 519 | 472 | 519 | 519 | 515 | 468 |
| F-Stat | 92.790 | 22.366 | 23.443 | 21.888 | 82.777 | 23.307 | 24.381 | 22.087 |
| *Panel F:* Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$ | | | | | | | | |
| | 1933-1950 | | | | 1933-1956 | | | |
| Share of refugees in 1950 | -0.290 | 1.675** | 1.553** | 1.851*** | 1.449 | 1.583 | 2.097** | 2.567*** |
| | (0.653) | (0.678) | (0.622) | (0.516) | (1.116) | (1.270) | (0.900) | (0.741) |
| *N* | 520 | 520 | 519 | 472 | 520 | 520 | 519 | 472 |
| F-Stat | 93.760 | 23.611 | 23.443 | 21.888 | 93.760 | 23.611 | 23.443 | 21.888 |
| State FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Pop. density (1939) | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Wartime destr. | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Geography | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Levels of dep. variable | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Pre-war controls | | | ✓ | ✓ | | | ✓ | ✓ |
| Addtl. pre-war controls | | | | ✓ | | | | ✓ |

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke.* *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The dependent variables are population growth (Panel A), changes in sectoral employment shares (Panels B - D), income per capita growth (Panel E) and the growth in the number of industrial plants (Panel F). The various specifications control for the share of destroyed housing stock ("Wartime destr."), the distance to the inner german border and a fixed effect for whether a county is a border county ("Geography"), the respective dependent variable in levels in the pre-war period ("Levels of dep. variable"), all six dependent variable in levels in the pre-war period in Panels A-F ("Pre-war controls") and the population share in cities with less than 2000 inhabitants in 1939, population density in 1933, the manufacturing share in 1933 and the GDP share in manufacturing and agriculture in 1935 ("Addtl. pre-war controls"). The share of refugees is instrumented with the population-weighted distance to the expulsion regions (see (1)) interacted with state fixed effects.

Table 6: The Effects of Refugee Inflows on the Local Economy: IV Estimates

Hence, if regions closer to the border are directly affected by the German division through political uncertainty or - as argued by Redding and Sturm (2008) - through a larger loss in market access, the identification assumption would be violated. This concern is unlikely to affect the conclusions drawn from Table 6 for three reasons. First, note that I include in all specifications a fixed effect for whether or not a particular county is a border county and I also control for the geographical distance to the inner german border. Second, both of these arguments would imply a negative correlation between the instrument and regional income growth or the growth of the manufacturing sector, that produces tradable goods. Hence, such concerns would induce a negative bias, pushing against the main findings reported in Table 6. Third, in Section A-2.4.1 of the Appendix I also offer an additional instrumental variable strategy, which is less subject to these concerns but also less precisely estimated.[13]

**Robustness** In Section A-2.4.1 in the Appendix I report a battery of robustness checks for the results reported in Tables 5 and 6. In particular, (i) I control for spatial variation in labor supply (as proxied by the aggregate employment share and the share of males) and local demand for reconstruction (as proxied by the share of the housing stock built after 1945), (ii) I report the results when counties are weighted by their population size to ease the concern that many small counties drive most of the variation (iii) I use the refugee share as of 1946 (instead of 1950) as the dependent variable, (v) I show that the results are not driven by particular cities or states by controlling for a full set of city times state fixed effects and (vi) I replicate the results using robust instead of clustered standard errors.

Throughout these specifications I find that most results are essentially identical to the baseline results. In terms of the OLS estimates reported in Table 5, the main difference is that the long-run relationship between GDP per capita and population growth with the refugee share in 1946 is not statistically significant. This is not entirely unsurprising given that a large number of refugees arrived only in 1946 and the following years (see Figure 2). Similarly, the IV estimates are largely robust to these concerns. There are three instances where the results are qualitatively different. First, like for the OLS, focusing on the refugee share in 1946 lowers the precision of the estimates and and renders the long-run impact on population growth and income growth insignificant. Second, if Bavaria, which is the largest state that accounts for almost 200 counties, is dropped from the analysis, the IV estimates for long-run income growth and plant entry cease to be significant. Third, if I only focus on the reduced form, the results are sometimes not statistically different from zero.

[13]This strategy exploits the fact that the inflowing refugees were often housed within the apartments of natives whenever housing was particularly scarce. Because doing so was easier if native homes were multi-room houses, the interaction between the expulsion distance and the supply of multi-room houses should predict the allocation of refugees.

| | 1939-1950 | | | 1961 | 1939-1961 | | |
|---|---|---|---|---|---|---|---|
| | Pop. Growth | Change ... share Manuf. | Agric. | Refugee share | Pop. Growth | Change ... share Manuf. | Agric. |
| Share of refugees (1950) | 1.131*** | 0.225*** | -0.322*** | 0.317*** | 0.574*** | 0.085*** | -0.126*** |
| | (0.070) | (0.020) | (0.039) | (0.022) | (0.063) | (0.023) | (0.041) |
| County FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Pre-war controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $N$ | 6035 | 6018 | 6035 | 5965 | 6018 | 6018 | 6021 |
| $R^2$ | 0.412 | 0.508 | 0.120 | 0.302 | 0.384 | 0.173 | 0.120 |

Note: Standard errors are clustered at the county level. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. All specifications control for county fixed effects, population density in 1939 and 1933 and the manufacturing employment share in 1939.

Table 7: The Effects of Refugee Inflows on the Local Economy: Variation Within Counties

**Within-County Evidence** To further corroborate these results, I also collected data for all local communities ("villages") for the state of Bavaria. This village-level data contains information on the local population, sectoral employment shares and the presence of refugees for more than 6000 villages. By combining the historical village data for the years 1939, 1950 and 1961, I can perform the same analysis as reported in Table 5 using only variation within counties.[14] The results, shown in Table 7, confirm the results of Table 5 and show that (across villages within counties) refugees are an important source of population growth and that they shift the village-level employment share from agriculture to manufacturing. Moreover, in 1950, the estimated elasticities based on the within county variation are almost the same as the ones based on the cross-county variation in Table 5.[15]

## 3.3 Refugees as a Source of Manufacturing Labor

One important reason for the stark expansion of the local manufacturing sector was that the incoming refugees often ended up as manufacturing workers. In Table 8 I report the distribution of refugees' sectoral employment shares relative to natives *within* counties, i.e. $\pi_{rs}^{Ref}/\pi_{rs}^{Nat}$ where $\pi_{rs}^{Ref}$ ($\pi_{rs}^{Nat}$) is the employment share of refugees (natives) in sector $s$ in county $r$.[16] A value of unity indicates that refugees and natives have the same sectoral employment

[14]Expectedly, these villages are tiny: in 1950, the median village has a population of around 550. The allocation of refugees across villages within counties, however, is still very dispersed (see Section A-2.3 in the Appendix). This high degree of variation in the initial allocation of refugees at very different levels of aggregation is consistent with the historical narrative of the non-organized "pouring-in" of refugees.

[15]In 1961 the effects, while still large and positive, are appreciably smaller. As I show in Section A-2.3 in the Appendix, this is a consequence of spatial mobility: within counties, refugees leave the most rural locations and move into near-by towns that offer more opportunities for industrial jobs. This type of "short-distance" mobility is not visible in the cross county variation.

[16]I can only report these statistics for the state of Bavaria, which is the only state which published sectoral employment in each county separately for refugees and natives.

|  | Mean | Distribution of $\pi_{rst}^{Ref}/\pi_{rst}^{Nat}$ Quantiles | | | | |
|---|---|---|---|---|---|---|
|  |  | 10% | 25% | 50% | 75% | 90% |
| Manufacturing | 1.36 | 0.93 | 1.10 | 1.31 | 1.54 | 1.81 |
| Agriculture | 0.37 | 0.18 | 0.21 | 0.29 | 0.40 | 0.76 |

Note: The table reports the distribution of refugees' relative sectoral employment shares $\pi_{rst}^{Ref}/\pi_{rst}^{Nat}$ across counties for the state of Bavaria.

Table 8: The Manufacturing Bias of Refugees' Labor Supply

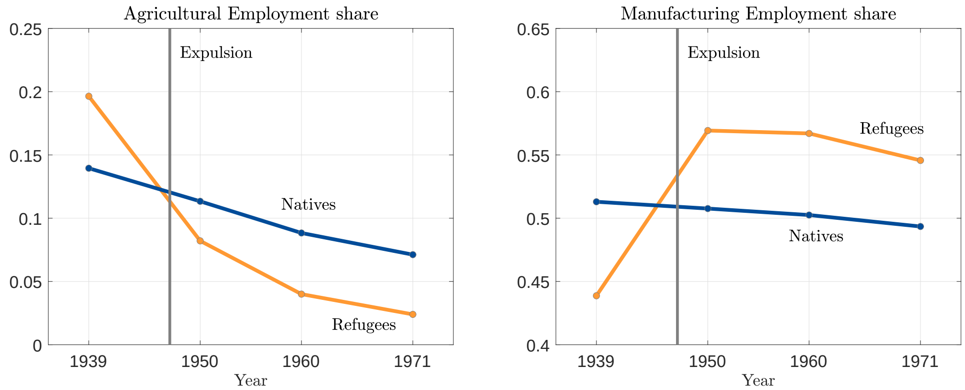
shares.

The table shows a clear pattern of comparative advantage: within local labor markets refugees are on average 36% more likely to work in manufacturing. In contrast, the average agricultural employment share among refugees is only 37% as large as the one of natives. The remaining columns show that these patterns hold throughout the entire distribution of counties. In less than 20% of counties are refugees less likely to work in manufacturing and there is no instance for refugees to be more likely to work in the agricultural sector.

This sectoral sorting is also apparent in the longitudinal microdata of the MZU 71. This unique supplement to the 1971 population census asked every respondent where he or she lived in 1939 and in which occupation and sector he or she worked in 1939, 1950, 1960 and 1971. By analyzing the time-series of these retrospective questions, I can measure snapshots of the employment life-cycle for both refugees and natives for a 40 year window. In Figure 4 I depict the sectoral life-cycle profile for the cohort of individuals born between 1915 and 1919. This cohort is 20-25 years old in 1939 and in their late twenties or early thirties at the time of the expulsion around 1947. In 1971, this cohort is 50-55 years old, i.e. still in the labor force. The two panels show the agricultural employment share (left panel) and the manufacturing employment share (right panel). The vertical line indicates the time of the expulsion. The process of reallocation is vividly apparent. Among refugees, 20% of the twenty year olds in 1939 used to work in the agricultural sector.[17] After the expulsion and their resettlement to West Germany, only 8% still did so. In contrast, the share of manufacturing employment, within the same cohort of individuals, increases from 44% to 57% after the settlement. The pattern for natives is strikingly different as the time period of the expulsion is hardly noticeable.[18]

[17] Note that this number is substantially smaller than the average agricultural employment share in 1939, which is closer to 50%. This is consistent with Porzio and Santangelo (2019) and Hobijn et al. (2018) who show that a large share of the structural transformation is accounted for by changes in employment shares across cohorts.

[18] The secular decline in agricultural and manufacturing employment for both natives and refugees in the post-war period reflects the process of structural change towards the service sector. In Section A-2.3 in the Appendix I analyze this data in more detail. The patterns shown in Figure 4 are not specific to this particular cohort. Interestingly, the patterns are different for young refugees that entered the labor market in Western




Note: The figure shows the agricultural employment share (left panel) and the manufacturing share (right panel) for the cohort of workers born between 1915 and 1919 by refugee status.

Figure 4: The Life-Cycle of the 1915-1919 Cohort

# 4 Theory: A Model of Spatial Growth

The settlement of refugees had three important consequences at the local level: (i) it had a large and persistent effect on the size of the local population, (ii) it was associated with industrialization at the local level and (iii) it led to increases in per-capita income, particularly in the long-run. In this section I develop a theory that can rationalize this evidence, both qualitatively and quantitatively.

## 4.1 Environment

I consider an economy with $R$ regions (counties in the data). Individuals face a consumption choice, i.e. how to allocate their expenditure across different goods, a sectoral labor supply choice, i.e. which sector to work in, and a migration choice, i.e. in which region to live and work. For tractability I assume that consumers are myopic and take optimal actions to maximize per-period utility. They derive utility from consuming both agricultural and manufacturing goods according to a Cobb-Douglas utility function $u\left(c_A, c_M\right) = c_A^{\alpha} c_M^{1-\alpha}$. Both goods $s = A, M$ are in turn CES aggregates from a set of differentiated, regional varieties that are tradable across space (subject to an iceberg trade cost $\tau_{rj}$) and aggregated according to $Y_{st} = \left(\sum_{r=1}^{R} Y_{rst}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$. Letting $P_{rst}$ denote the price of sector $s$ goods from $r$ in $r$, consumers in region $j$ pay $\tau_{rj} P_{rst}$ for region $r$ goods. The price index of sector $s$ goods in region $j$ is therefore given by $\overline{P}_{jst} = \left(\sum_r \left(\tau_{rj} P_{rst}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$.

Germany. This suggests an important role for social mobility across generations, a finding I also corroborate using self-reported data on social status.

**Production** The agricultural good is produced using labor and land according to

$$Y_{rAt} = \mathcal{Q}_{rt} \mathcal{T}_r^{1-\gamma} H_{rAt}^{\gamma},$$

where $\mathcal{T}_r$ denotes agricultural land in region $r$, $H_{Art}$ denotes the total amount of labor employed for agricultural production and $\mathcal{Q}_{rt}$ is productivity in region $r$ at time $t$, that evolves exogenously. Agricultural land $\mathcal{T}_r$ is in fixed supply, so that agricultural production is subject to decreasing returns to scale. The returns to land accrue to a set of immobile land-owners that only consume.

The manufacturing good is subject to variety gains as in Romer (1990) and is produced according to

$$Y_{Mrt} = \mathcal{Q}_{rt} \left( \int_0^{N_{rt}} x_{it}^{(\rho-1)/\rho} di \right)^{\rho/(\rho-1)},$$

where $N_{rt}$ denotes the endogenous measure of varieties, $x_{it}$ denotes the quantity of input $i$ used and $\rho > 1$ is the elasticity of substitution across inputs. Such inputs are produced using only manufacturing labor, i.e. $x_{it} = h_{it}$.

The regional productivity term $\mathcal{Q}_{rt}$ evolves according to the persistent process

$$\ln \mathcal{Q}_{rt} = (1 - \varrho) \ln \mathcal{Q}_r + \varrho \ln \mathcal{Q}_{rt-1} + \varpi u_{rt},$$

where $\mathcal{Q}_r$ is a fixed, region-specific level of innate productivity, $\varrho$ governs the regional autocorrelation and $u_{rt}$ is a productivity shock, which is distributed iid with a unit variance. Hence, the county fixed effect $\mathcal{Q}_r$ determines the long-run value of exogenous productivity in region $r$ and $\varpi$ governs the variance of regional productivity shocks.

**Entry** The measure of input varieties $N_{rt}$ is determined endogenously and provides the link between local productivity and labor supply. At the end of each period (after production has taken place) an exogenous fraction $\delta$ of firms exits. Firm entry takes place in the beginning of the period. The labor requirement to start a new firm in region $r$ at time $t$, $h_{rt}^E$, is given by

$$h_{rt}^E = f_E N_{rt-1}^{-\lambda}, \tag{3}$$

where $\lambda \leq 1$. The parameter $\lambda$ governs the extent of dynamic spillovers as in Jones (1995) and, as I show below, is a crucial parameter to determine the long-distribution of economic activity across space and whether population shocks have persistent effects. Because $\lambda$ determines how the existing state of knowledge $N_{rt-1}$ affects the costs of creating new knowledge, I refer to it as the *inter-temporal knowledge elasticity.* The parameter $f_E$ determines the size of entry costs.

The entry technology in (3) nests three important benchmark models as special cases and the structural estimation allows me to distinguish between them. If $\lambda = 0$ and $\delta = 1$, the model is the static model of Krugman (1980): firms only live for a single period and the cost of entry do not depend on the number of varieties already available. The case of $\lambda = 1$ is the specification of Romer (1990), where the costs of creating new varieties are inversely proportional to the level of knowledge $N_{r,t-1}$. This specification of the model leads to fully endogenous growth. The intermediate case of $0 < \lambda < 1$ is the semi-endogenous growth model of Jones (1995), where growth in the long-run is fully determined by population growth. As I discuss in detail below, these different parameterizations have strikingly different implications for the dynamic effects of refugee inflows on local income per capita.

**Sectoral Labor Supply** I model the sectoral supply of human capital using the usual Roy-type machinery. Individuals are characterized by a two-dimensional efficiency vector $z_{it} = (z_{iAt}, z_{iMt})$, where $z_{ijt}$ denotes the number of efficiency units individual $i$ can supply to sector $j$ and is drawn from a Fréchet distribution, i.e. $F_j(z) = e^{-\phi_j z^{-\theta}}$. To meaningfully talk about the composition of the local workforce, I allow for persistent differences in average human capital that is parametrized by $\phi_j$. I assume there exist two latent types, "industrial workers" ($I$) and "farmers" ($F$), that have a comparative advantage in the respective sectors, $\phi_M^I/\phi_A^I > \phi_M^F/\phi_A^F$. The share of individuals of type $\nu \in \{F, I\}$ working in sector $j$ in region $r$ is then given by

$$\pi_{rjt}^{\nu} = \phi_j^{\nu}\left(\frac{w_{rjt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta}, \quad \text{where} \quad \overline{w}_{rt}^{\nu} = \left(\phi_A^{\nu} w_{rAt}^{\theta} + \phi_M^{\nu} w_{rMt}^{\theta}\right)^{1/\theta}.$$

By allowing for latent types, the model provides a distinct role for the composition of the local population to determine aggregate labor supply. In particular, denote the number of natives and refugees of type $\nu$ at time $t$ in region $r$ by $L_{rt}^{N\nu}$ and $L_{rt}^{R\nu}$, the population share of refugees by $\mu_{rt}$ and the share of workers of type $\nu$ among refugees and natives in region $r$ by $\omega_{rt}^{R\nu}$ and $\omega_{rt}^{N\nu}$. The manufacturing share among refugees relative to natives, $\pi_{rMt}^R - \pi_{rMt}^N$, is then given by

$$\pi_{rMt}^R - \pi_{rMt}^N = \left(\pi_{rMt}^I - \pi_{rMt}^F\right) \times \left(\omega_{rt}^{IR} - \omega_{rt}^{IN}\right). \tag{4}$$

This expression highlights that the empirical pattern shown in Table 8 that refugees were more likely to work in manufacturing within labor markets suggests that the share of industrial workers among refugees exceeds the ones of natives.[19]

For my baseline analysis I assume that that the *aggregate* share of industrial workers is given by $\chi$ for both groups. This implies that the *spatial* distribution of industrial types differs across groups if native workers are endogenously sorted but refugees are randomly assigned. In particular, within rural areas the average native who chose to remain in an agriculturally specialized labor market might have had a comparative advantage in the agricultural sector relative to a randomly selected refugee. Intuitively, the share of engineers within an arriving refugee trek might have been higher than in the rural native population. Refugee inflows thus increase both the size of the local population and change the composition of the workforce. This has specific implications for the differential impact of refugee inflows across space and I show below that the model-implied sorting is consistent with the empirical relationship implied by (4).

**Spatial Mobility** Individuals are mobile across space, but subject to a friction. Individuals know their type before making their moving decisions, but they do not observe their particular skill realization $z_{it}$. The utility for individual $i$ who currently lives in region $j$ and moves to region $r$ at time $t$ is thus given by $\mathcal{U}^i_{jrt} = \mathcal{A}_{rt}\overline{u}^i_{rt}\eta_{jr}\xi^i_{rt}$. Here, $\mathcal{A}_{rt}$ denotes an amenity in region $r$, $\overline{u}^i_{rt}$ is the expected utility individual $i$ achieves in region $r$, $\eta_{jr}$ parametrizes the moving costs from $j$ to $r$ and $\xi^i_{rt}$ is a regional taste shock which is independent both across individuals and across locations for any given individual and Fréchet distributed with shape parameter $\varepsilon > 1$. The share of people of type $\nu$ moving from region $j$ to region $r$, $m^\nu_{jrt}$, is thus given by

$$m^\nu_{jrt} = \frac{\left(\mathcal{A}_{rt}\eta_{jr}\overline{u}^\nu_{rt}\right)^\varepsilon}{\sum_d \left(\mathcal{A}_{dt}\eta_{jd}\overline{u}^\nu_{dt}\right)^\varepsilon}, \tag{5}$$

where expected utility of individual $i$ of type $\nu$ in region $r$ is given by $\overline{u}^\nu_{rt} \propto \overline{w}^\nu_{rt}/\left(\overline{P}^\alpha_{rAt}\overline{P}^{1-\alpha}_{rMt}\right)$. Note that (5) encapsulates the economics of spatial sorting: because industrial types put a higher weight on manufacturing wages, they move towards locations with a comparative advantage in manufacturing.

Motivated by the fact that the spatial allocation of refugees was highly persistent, I allow for (in addition to the moving costs encapsulated in $\eta_{jr}$) a second mobility friction a la Calvo: at each point in time individuals have the option to move with probability $\psi > 0$ (see also Bilal (2019)). The combination of $\psi < 1$ and $\eta_{jr} \neq \eta_{kr}$ for $j \neq k$ parsimoniously captures the intensive and extensive margin of costly migration. The "Calvo shock" $\psi$ mostly governs the persistence of the initial population distribution: the lower $\psi$, the longer it takes for the

[19] Note that $\pi^I_{rMt} - \pi^F_{rMt} = \left(\frac{\phi^I_M/\phi^I_A}{\phi^F_M/\phi^F_A} - 1\right)\pi^I_{rAt}\pi^F_{rMt} > 0$ because of the comparative advantage of industrial workers in the manufacturing sector.

initially assigned refugees to percolate spatially. The bilateral migration frictions $\eta_{jr}$ govern the spatial proximity of moving flows conditional on moving. In my quantitative application I assume that $\eta_{jr} \propto d_{jr}^{-\kappa}$, where $d_{jr}$ is the geographic distance between $j$ and $r$ and $\kappa$ is a parameter, which I estimate. Similarly, as in Allen and Donaldson (2020), I assume that local amenities are a power function of the local population, $\mathcal{A}_{rt} = \mathcal{A}_r L_{rt}^{-\beta}$. The parameter $\beta$ captures congestion forces such as the scarcity of local housing or rivalries in the usage of public goods and $\mathcal{A}_r$ is the time-invariant component of location amenities.

## 4.2 Equilibrium

The timing of events is as follows. At the beginning of period $t$, the set of state variables in region $r$ is given by its exogenous productivity $\mathcal{Q}_{rt-1}$, the number of existing varieties $N_{rt-1}$ and the local population of industrialists and farmers $\mathscr{L}_{rt-1} = \left(L_{rt-1}^F, L_{rt-1}^I\right)$.[20] Then the exogenous productivity shock $\mathcal{Q}_{rt}$ is realized, individuals make their mobility decision, and new firms decide whether or not to enter. These choices determine the future set of state variables $(\mathcal{Q}_{rt}, N_{rt}, \mathscr{L}_{rt})$. Production, consumption and factor prices are then determined as the outcomes of a static trade equilibrium.

**Static Equilibrium** To solve for the static equilibrium allocations, consider first the manufacturing sector.[21] Because the market for intermediate inputs is monopolistically competitive, firms charge a constant markup and receive a share $1/\rho$ of firm revenue as profits. Production workers thus receive a share $(\rho - 1)/\rho$ of revenue as labor payments. This implies that profits of firm $i$ in region $r$ are given by

$$\pi_{ir} = \frac{1}{\rho}\frac{P_{rM}Y_{rM}}{N_r} = \frac{1}{\rho - 1}\frac{w_{rMt}H_{rPt}}{N_r},$$

where $H_{rPt}$ is the aggregate production labor input in region $r$ at time $t$.

The mass of varieties $N_{rt}$ is determined by free entry. As for workers, I assume that entering firms act myopically, only considering static profits as part of their entry decision.[22]

[20] Note that, because refugees and natives are identical conditional on their type $\nu$, the relevant state variable is only the local distribution of types $L_{rt-1}^{\nu} = L_{rt-1}^{R\nu} + L_{rt-1}^{N\nu}$.

[21] See Section A-1.1 in the Appendix for details.

[22] As for the owners of land, I assume that firm profits accrue to a set of spatially immobile entrepreneurs.

Free entry therefore requires that[23]

$$\pi_r = w_{rMt} h_{rt}^E = w_{rMt} f_E N_{rt-1}^{-\lambda}. \tag{6}$$

Using the expressions for profits $\pi_{ir}$ this yields a simple expression for the evolution $N_{rt}$:

$$N_{rt} = \frac{1}{f_E}\frac{1}{\rho-1} \times \underbrace{H_{rPt}}_{\text{Market size}} \times \underbrace{(N_{rt-1})^{\lambda}}_{\text{Dynamic agglomeration}}. \tag{7}$$

Equation (7) is the key equation of the model as it highlights the two determinants of variety creation and hence productivity growth at the local level. The first term is the usual scale effect: a larger workforce $H_{rPt}$ triggers the entry of new varieties because it goes hand in hand with larger profits. Note that $H_{rPt}$ emerges as a sufficient statistic that summarizes all equilibrium effect on sectoral wages and aggregate demand, which are determined as part of the trade and spatial equilibrium. The second term captures the dynamic agglomeration force. As long as $\lambda > 0$, the equilibrium features persistence whereby the existing number of varieties positively predicts the future number of varieties.

Armed with equation (7) I can also solve for the endogenous aggregate production function of the manufacturing sector, which is given by

$$Y_{rMt} = \varsigma_1 \mathcal{Q}_{rt} N_{rt}^{\frac{1}{\rho-1}} H_{rPt} = \varsigma_2 \mathcal{Q}_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{1+\vartheta} \quad \text{where} \quad \vartheta = \frac{1}{\rho-1}, \tag{8}$$

and $\varsigma_1$ and $\varsigma_2$ are inconsequential constants. The first equality of equation (8) shows the usual variety gains: a larger mass of varieties $N_{rt}$ increases productivity. Given $N_{rt}$, the manufacturing sector has constant returns to scale. The second equality exploits the fact $N_{rt}$ is itself increasing in the size of the workforce. This implies that the manufacturing sector has increasing returns holding a location's pre-determined state variables $(\mathcal{Q}_{rt}, N_{rt-1})$ fixed. I thus refer to $\vartheta$ as the *short-run scale elasticity.* Note also that - holding $(\mathcal{Q}_{rt}, N_{rt-1})$ fixed - the expression in (8) is isomorphic to a setting with exogenous agglomeration forces common in many models of economic geography (Redding and Rossi-Hansberg, 2017).

Given the aggregate sectoral production functions, the static equilibrium can be fully characterized given the vector of pre-determined productivities $(\mathcal{Q}_{rt}, N_{rt-1})$ and the population distribution $\mathscr{L}_{rt}$.

**Definition 1.** *Given $\{\mathcal{Q}_{rt}, N_{rt-1}, \mathscr{L}_{rt}\}_r$, a static equilibrium is a set of wages and land rents*

[23] While the free entry condition always holds with equality in the steady-state, it might be slack during the transitional dynamics. To avoid a taxonomic presentation of the results, I focus on the case where (6) holds with equality in the main text. In the quantitative application I of course allow for the general case where (6) might be slack.

$\{w_{rAt}, w_{rMt}, R_{rt}\}_r$, *intermediate varieties, input prices and quantities* $\left\{N_{rt}, [p_{irt}, x_{irt}]_{i=0}^{N_{rt}}\right\}_r$, *sectoral employment allocations* $\{H_{rAt}, H_{rPt}, H_{rEt}\}_r$ *and quantities of tradable goods* $\{Y_{rAt}, Y_{rMt}\}_{rt}$, *such that (i) firms and consumers behave optimally and (ii) labor and good markets clear.*

Because this static trade equilibrium is characterized by the typical labor and goods market clearing condition, I relegate the formal derivation to the Appendix (see Section A-1.2).

**Dynamic Equilibrium** This static equilibrium that determines the distribution of factor prices $\{w_{rAt}, w_{rMt}\}_r$ depends on the population distribution $\mathscr{L}_{rt}$. Moreover, the evolution of regional varieties $\{N_{rt}\}$ also has to be consistent with firms' entry decisions. The dynamic equilibrium of this economy is thus defined in the following way:

**Definition 2.** *Given a path of exogenous productivity* $\{\mathcal{Q}_{rt}\}_{rt}$ *and an initial condition* $\{\mathscr{L}_{r0}, N_{t0}\}_r$, *a dynamic equilibrium is a path of local populations* $\{\mathscr{L}_{rt}\}_{rt}$ *and local varieties* $\{N_{rt}\}_{rt}$, *such that (i)* $\{\mathscr{L}_{rt}\}_{rt}$ *is consistent with individuals' optimal mobility decisions, (ii)* $\{N_{rt}\}_{rt}$ *is consistent with free entry and (iii) the resulting allocations represent a static equilibrium at each point in time.*

The laws of motion of the two endogenous state variables $\{\mathscr{L}_{rt}, N_{rt}\}_{rt}$ are given by

$$N_{rt} = \frac{1}{\rho - 1}\frac{1}{f_E} \times H_{rPt}\left(\{\mathcal{Q}_{rt}, N_{rt-1}, \mathscr{L}_{rt}\}\right) \times N_{rt-1}^{\lambda} \tag{9}$$

$$L_{rt}^{\nu k} = (1-\psi)\, L_{rt-1}^{\nu k} + \psi \sum_{j=1}^{R} L_{jt-1}^{\nu k} m_{jrt}^{\nu}\left(\{\mathcal{Q}_{rt}, N_{rt-1}, \mathscr{L}_{rt}\}\right), \tag{10}$$

where the notation highlights that $H_{rPt}$ and $m_{jrt}^{\nu}$ are determined as part of the static equilibrium because they are functions of $\{\mathcal{Q}_{rt}, N_{rt-1}, \mathscr{L}_{rt}\}$ via equilibrium wages and prices.

Equations (9) and (10) are the key dynamic equations of my theory because they describe the joint evolution of local productivity $N_{rt}$ and the local population $\mathscr{L}_{rt}$. Local productivity $N_{rt}$ depends on the size of the local population $\mathscr{L}_{rt}$ through the size of the manufacturing workforce: a larger population raises $H_{rPt}$ and thus triggers variety creation. Similarly, $\mathscr{L}_{rt}$ depends on the mass of local varieties through equilibrium factor prices and agents' migration choices.

## 4.3 Population Inflows and Persistent Local Productivity Dynamics

The joint dynamics of local varieties and the local population depend crucially on the extent of spatial mobility and the knowledge elasticity $\lambda$. Note that (9) implies that the equilibrium process for local productivity $N_{rt}$ is akin to an AR(1) process:

$$\ln N_{rt} = \alpha_0 + \lambda \ln N_{rt-1} + \ln H_{rPt},$$

where $\alpha_0 = \ln\left(\frac{1}{\rho-1}\frac{1}{f_E}\right)$. Hence, the knowledge elasticity $\lambda$ emerges as the key parameter governing the persistence of changes in market size. For any $\tau \geq t_0$, the level of productivity $N_{r\tau}$ is thus given by

$$\ln N_{r\tau} = \Lambda(t, t_0) + \lambda^{\tau-(t_0-1)} \ln N_{rt_0-1} + \sum_{j=t_0}^{\tau} \lambda^{\tau-j} \ln H_{rPj}, \tag{11}$$

where $\Lambda(t, t_0) = \alpha_0 \sum_{j=t_0}^{\tau} \lambda^{j-t_0}$. Local productivity depends on the entire *history* of the manufacturing workforce $\{H_{rPj}\}_{j=t_0}^{t_0+\tau}$, discounted by $\lambda$. Intuitively: local productivity encapsulates the entire history of local scale, because past market size led to plant entry, which made the creation of future varieties easier.

Expression (11) highlights that local labor supply shocks can have transitory effects, long-lasting effects or lead to persistence, where the long-run outcomes depend on the history of past shocks. With free mobility, i.e. $\psi = 1$ and $\eta_{jk} = 1$, the distribution of people across space ceases to be a state variable and a population shock to an individual region lasts only for a single period. If in addition there are no dynamic spillovers, i.e. $\lambda = 0$, the model is a static model with agglomeration forces as in Allen and Arkolakis (2014) or Ahlfeldt et al. (2015). This parameterizations is at odds with my empirical findings that the population shock was persistent and that the effect on income per capita was increasing over time.

With frictions to spatial mobility and $0 < \lambda < 1$, the initial allocation of refugees has long-lasting effects. If $\lambda > 0$, even a one-time increase in $H_{rPt}$ affects local productivity in all future periods. As long as $\lambda < 1$, the productivity response is long-lasting, but subsides eventually. For the limiting case of $\lambda = 1$, the productivity process is a random walk, shocks have permanent effects and the cross-sectional productivity distribution is not stationary. Furthermore, with frictions to spatial mobility, a population shock in $t$ induces an increase in $H_{rPt}$ for future periods and hence complements the long-lasting the productivity response. However, as long as the shock does not increase $H_{rPt}$ permanently, the productivity response is also not permanent.[24]

[24] As a specific example, suppose there is a positive shock to $H_{rP}$ at $t_0$, which subsides at rate $p \leq 1$, that is $d \ln H_{rPd+t_0} = d \ln H_{rPt_0} \times p^d$. As I show in Section A-1.3 in the Appendix, the elasticity of local varieties with respect to the initial shock is given by $d \ln N_{rd+t_0 /} d \ln H_{rPt_0} = \Psi_d(p, \lambda) = \frac{\lambda^{d+1}-p^{d+1}}{\lambda-p}$. If the shock is transitory, $\Psi_d(0, \lambda) = \lambda^d \to 0$, i.e. the productivity response is long-lasting but declining over time. If the shock was permanent, $\Psi_d(1, \lambda) = \frac{1-\lambda^{d+1}}{1-\lambda} \to \frac{1}{1-\lambda}$, i.e. the effect is increasing over time. If $0 < p < 1$, the productivity response subsides in the long-run, but the impulse response $\Psi_d(p, \lambda)$ is hump-shaped if $\lambda + p > 1$.

Finally, the model also admits the possibility of full persistence (or, in the terminology of Allen and Donaldson (2020), path-dependence), where the initial allocation determines the allocation in the long-run. According to (11), a temporary increase in local labor supply triggers the creation of local varieties. This in turn raises local wages and can prevent individuals from leaving, leading to a permanent increase in $H_{rPt}$. This feedback-loop is more likely to occur if agglomeration forces are large (i.e. $\lambda$ are large and $\rho$ is small), spatial dispersion forces are limited (i.e. $\varepsilon$ and $\sigma$ are large and $\beta$ is small) and mobility is subject to frictions (i.e. $\psi$ is small).

## 4.4 Balanced Growth and the Long-Run Scale Elasticity

To study the long-run implications, consider the behavior of the economy along a non-stochastic spatial balanced growth path (SBGP), which I define as an allocation where the population distribution is stationary and regional wages grow at a common rate. Along a SBGP innate productivity $\mathcal{Q}_{rt}$ is constant and equal to its long-run level $\mathcal{Q}_r$. With a stationary population, goods market clearing implies that regional varieties grow at a common rate:

$$g_N = \frac{N_{rt}}{N_{rt-1}} = \frac{1}{\rho-1}\frac{1}{f_E} H_{rPt} N_{rt-1}^{\lambda-1}. \tag{12}$$

Equation (12) has obvious similarities to the growth equation analyzed in Jones (1995). For $g_N$ to be constant across space, the local mass of varieties along a SBGP is given by

$$N_{rt} = \left(\frac{1}{g_N^{\lambda}}\frac{1}{\rho-1}\frac{1}{f_E}\right)^{\frac{1}{1-\lambda}} H_{rPt}^{\frac{1}{1-\lambda}}. \tag{13}$$

and thus tied to local employment in the manufacturing sector. Hence, if $\lambda < 1$ this is a model of semi-endogenous growth as in Jones (1995), where in the absence of population growth, variety and income growth is bound to be zero in the long-run and the economy converges to a steady-state. The case of $\lambda = 1$ is qualitatively different. As is apparent from (12), generically there does not exists a SBGP as this would require the amount of human capital to be equalized across space. The linear relationship between growth and the level of population is of course exactly the case of "strong scale effects", which is at the heart of most models of endogenous growth. Equation (12) can therefore be read as the spatial analog of the distinction between endogenous and semi-endogenous models of growth: the spatial distribution of economic activity is stationary in the latter but not stationary in the former.

Equation (13) highlights the importance of local scale effects: regions where $H_{rPt}$ is large have high productivity. Crucially, the long-run relationship between productivity and the manufacturing workforce is fundamentally different from the short-run relationship. Combining the balanced growth relationship (13) with the equilibrium production function (8) implies local labor productivity in manufacturing to be

$$Y_{rMt}/H_{rPt} \propto \mathcal{Q}_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{\vartheta} \propto \mathcal{Q}_r H_{rPt}^{\frac{\vartheta}{1-\lambda}}.$$

Thus, whereas the short-run elasticity $\vartheta$ describes the relationship between local productivity and local scale, holding $N_{rt-1}$ constant, the long-run elasticity takes the endogeneity of $N_{rt-1}$ into account and is given by $\vartheta/(1-\lambda)$. As long as $\lambda > 0$, the long-run scale elasticity exceeds the short-run elasticity because the dynamic accumulation of ideas amplifies the static differences in regional scale.

# 5 Structural Estimation and Quantitative Analysis

I now estimate the structural parameters of my theory by fitting the empirical results of Section 3. This exercise has two main purposes. First, I show that the theory can quantitatively rationalize the empirical results presented in Section 3. Second, the model allows me quantify the effect of the refugee-settlement on aggregate income and to study how the government policy of sending refugees to the countryside ignited persistent rural industrialization.

## 5.1 Estimation and Identification Strategy

The model is fully parametrized by 17 structural parameters and a tupel of fundamentals $\{\mathcal{Q}_r, \mathcal{A}_r, \mathcal{T}_r\}_r$ per region. I calibrate 5 parameters externally and estimate the remaining 12 parameters within the context of this paper

$$\Omega = \left\{ \underbrace{\rho, \lambda}_{\text{Growth}}, \underbrace{\varepsilon, \psi, \kappa, \beta}_{\text{Spatial Mobility}}, \underbrace{\chi, \phi_M^I, \phi_A^I}_{\text{Skill distribution}}, \underbrace{\varpi}_{\text{Process of}\mathcal{Q}_{rt}}, \underbrace{\alpha}_{\text{Agricult. spending}}, \underbrace{\xi}_{\text{Trade costs}} \right\}.$$

My empirical strategy to identify $\Omega$ and $\{\mathcal{Q}_r, \mathcal{A}_r, \mathcal{T}_r\}_r$, which I describe in more detail in Section A-2.5 in the Appendix relies on two steps:

First, given the parameters $\Omega$, I identify the fundamentals $\{\mathcal{Q}_r, \mathcal{A}_r, \mathcal{T}_r\}_r$ by calibrating the model to the cross-regional data on GDP per capita, sectoral employment shares and population size in 1933, which I assume to correspond to a steady-state of the system.[25]

[25]Formally, given a set of structural parameters, there is a one-to-one mapping between $\{\mathcal{Q}_r, \mathcal{A}_r, \mathcal{T}_r\}_r$ and the three moments for each region. The steady-state also determines the endogenous distribution of $N_{r1933}$ and the extent of spatial sorting, i.e. $\omega_{r1933}^I$. In principle, one could identify the fundamentals without the steady-state assumption. This would, however, require at least two periods where the above mentioned data was observed. I only have access to the data on GDP per capita at the county level for a single period prior

Then I replicate the historical experiment of the refugee settlement in my model. To do so, I simulate the dynamic evolution of the economy starting in 1933 and "shock" the economy with the inflow of refugees in the post-war period. Because the majority of refugees arrived around the year 1947, I assume that all refugees arrived in 1947 and I allocate them according to the empirically observed share of refugees in 1950.[26] Hence, the model - by construction - replicates the correlation between the share of refugees and population density, GDP per capita and sectoral employment shares in 1933 and therefore captures the systematic correlation with local fundamentals. In terms of fundamentals, refugee-rich localities tend to have low permanent productivity $\mathcal{Q}_r$ and a comparative advantage in the productivity of agricultural goods (i.e. high $\mathcal{T}_r/\mathcal{Q}_r$).

**Estimation Moments** I estimate the parameters $\Omega$ through a combination of calibration and indirect inference. In total I target 16 moments. Eleven of these moments directly exploit the regional variation in the allocation of refugees induced by the historical experiment. First, I target the regression coefficients between the share of refugees in 1950 and population growth, income per capita growth and the growth of manufacturing employment in 1950 and 1961 reported in columns 2 and 6 of Table 5 and run the same specification in the model-generated data. By relying on a specification that controls for population density and economic outcomes in the pre-war period, the regressions implicitly control for the variation in fundamentals. And because the allocation of refugees is uncorrelated with the regional productivity shock $u_{rt}$, this specification is - in the context of my model - consistent with the identification assumptions underlying my OLS strategy.[27] Second, I target the correlation between the refugee share in 1950 and 1955 and 1961 depicted in Figure 3.

I augment this indirect inference strategy with three additional regression moments that directly speak to the short-run dynamics of local population growth and the long-run response of local productivity and population size. First I target a regression between the share of refugees in 1950 and subsequent population growth between 1950 and 1955 after controlling for state fixed effects, pre-war population density and war-time destruction. I find a precisely estimated coefficient of -0.342 indicating that local congestion plays an important role. Second, in addition to the effects in 1950 and 1961, I also target the relationship between refugee inflows

to the war.

[26]Even though the model-implied refugee share in 1950 is therefore not exactly equal to the one in data, the difference is very small because the estimated mobility hazard $\psi$ is small.

[27]By focusing on the OLS estimates, I can directly use the observed share of refugees and hence ensure that the model matches the cross-sectional distribution of refugees and its correlation with other county characteristics. If had opted to use the IV estimate as a moment for identification, I would have had to model the endogeneity and the first stage explicitly. Given that empirically the OLS and the IV estimates are broadly similar, I chose to target the OLS results.

and income per capita and population size in 1980. I use the same regression specification as reported in columns 2 and 6 in Table 5. I find a coefficient of 0.201 (with a standard error of 0.198) for income growth and 1.041 (with a standard error of 0.521) for population growth. Hence, refugee-rich counties in 1950 are still larger and richer in 1980. Economically, these patterns points towards a parametrization where the initial population shock was very persistent. In Table A-21 in the Appendix I summarize all the regression moments above in a unified table.

I utilize five additional moments to identify the model. First, I use two moments on relative factor payments. I discipline the average earnings premium in manufacturing relative to agriculture, i.e. the "agricultural productivity gap". Because my data does not contain direct information on local earnings by sector, I target a value of 1.5, which is in line with the results reported in Gollin et al. (2014). I also exploit differences in earnings between refugees and natives at the micro-level by estimating a Mincer-type regression. Empirically, refugees earn about 7.5% less than natives. Second, to estimate the size of spatial trade and migration frictions, I first use a historical migration survey in 1955 that reports - for each county - the share of out-migrants that remains in their state. Empirically, the 2/3 of migration flows occur within the same state and I target this number to discipline the extent to which migration costs are increasing in distance. I recover the elasticity of trade costs with respect to distance $\zeta$ from the gravity relationship of within-country trade flows. Because I do not have access to historical trade flow data from Germany at the county-level, I target the moment reported in Monte et al. (2018), who use data on shipments within the US and estimate a distance elasticity of -1.29.[28] Finally, to estimate the dispersion of productivity shocks $\varpi$, I run the county-level panel regression $\ln y_{rt} = \delta_r + \beta \ln y_{rt-1} + v_{rt}$ both in the model and the data and target the dispersion of the estimated residuals, i.e. $sd\left(\hat{v}_{rt}\right)$.

**Mapping to Parameters** Even though I estimate all moments jointly, they map directly to the main parameters of the interest. The two scale parameters $\rho$ and $\lambda$ are mostly identified by the response of income per capita and manufacturing employment at different horizons. The Calvo-type mobility friction $\psi$, the labor supply elasticity $\varepsilon$ and the strength of local congestion $\beta$ are important determinants of the extent of spatial mobility and are informed by the spatial auto-correlation of refugee shares and the correlation between refugee inflows and population growth. The data on earnings across sectors and between refugees and natives aid in identifying the human capital parameters. Holding $\phi_A^I$ fixed, $\phi_M^I$ increases relative human capital of industrialists and hence the measured agricultural gap. And the extent of sorting,

[28]This elasticity is consistent with the findings reported in Wolf (2009), who analyzes data on trade flows across 21 regions in Germany in the pre-war period. He estimates a distance elasticity of around -1.4.

which is influenced by the share of industrial workers $\chi$, affects relative earnings because refugees are, on average, located in rural locations that feature lower factor prices.

**Estimation** I minimize the distance between these empirically observed moments and the moments in the model using Sobol grids. To account for the sampling variation induced by the stochastic productivity process, I replicate this entire analysis 50 times and calculate the average of all moments and regression coefficients. The five parameters I set externally are the trade elasticity $\sigma$, the labor share in the agricultural sector $\gamma$, the dispersion of skills $\theta$, the correlation of the productivity process $\varrho$ and the exogenous exit rate $\delta$. I assume that $\sigma = 5$, $\gamma = 0.5$, $\theta = 1.5$, $\varrho = 0.9$ and $\delta = 0.1$. The fixed cost of entry $f_E$ can be normalized by an appropriate choice of units for $N_{rt}$.

**The German Division and the Loss of Market Access** I augment my estimation strategy by one additional important historical feature. As highlighted in my empirical analysis and stressed in Redding and Sturm (2008), the spatial allocation of refugees is correlated with a second "spatial shock": the division of Germany also represented a loss of market access for counties closer to the inner german border (and hence with a higher share of refugees). To capture this correlation in my quantitative analysis, I allow for trade between Western and East Germany in the pre-war period and then model the German division (and the resulting loss in market access) as a prohibitive increase in both trade and mobility costs.[29] Because trade costs prior to the war are a function of distance, counties that are closer to the border are more affected by this shock. To implement this shock, I model East Germany as an "$R + 1$"th region in the pre-war period and estimate its economic size by targeting the regression coefficient on distance to the inner German border and local income growth between 1939 and 1961 corresponding to my main specification in column 6 of Table 5. Intuitively, I discipline the amount of trade there must have been between East and West to force the model to replicate the positive cross-sectional correlation between distance and income growth once trade is prohibited.

## 5.2 Estimation Results and Model Fit

In Table 9 I report the estimated structural parameter and the fit of the model. The model is able to replicate the targeted moments well. In particular, it matches the persistent positive correlation between refugee inflows and population growth (rows 1 - 3) and manufacturing

[29] My assumption of East and West Germany being relatively integrated is consistent with Wolf (2009, p. 876) who finds that "the nearly impregnable border between East and West that existed between about 1946 and 1989 was therefore hardly predictable in 1939."

| Structural Parameters | | | Moments | Data | Model |
|---|---|---|---|---|---|
| | *Scale Elasticities* | | *Experimental Moments* | | |
| $\lambda$ | Inter-temporal elasticity | 0.71 | Pop growth 39-50 (Table 5) | 1.36 | 1.19 |
| $\rho$ | Elasticity of substitution | 5.02 | Pop growth 39-61 (Table 5) | 1.029 | 0.934 |
| | *Human Capital* | | Pop growth 39-80 (Table A-21) | 1.06 | 0.914 |
| $\phi_M^I$ | HC of industrialists in manuf. | 13.61 | Manuf. growth 39-50 (Table 5) | 0.317 | 0.272 |
| $\phi_A^I$ | HC of industrialists in agric. | 0.84 | Manuf. growth 39-61 (Table 5) | 0.241 | 0.299 |
| $\chi$ | Share of industrial workers | 0.58 | Income growth 39-50 (Table 5) | -0.083 | -0.003 |
| | *Spatial Mobility* | | Income growth 39-61(Table 5) | 0.502 | 0.358 |
| $\varepsilon$ | Spatial labor supply elasticity | 2.12 | Income growth 39-80 (Figure A-21) | 0.201 | 0.388 |
| $\psi$ | Frequency of mobility shocks | 0.07 | Refugee share 1955 (Figure 3) | 0.735 | 0.763 |
| $\beta$ | Congestion elasticity of amenities | -0.16 | Refugee share 1961 (Figure 3) | 0.586 | 0.556 |
| $\kappa$ | Dist. elasticity of mov. costs | -1.09 | Pop growth 50-55 (Figure A-21) | -0.342 | -0.183 |
| | *Other* | | Distance and income growth (Tab A-21) | 0.06 | 0.012 |
| $\varpi$ | Disp. of prod. shocks | 0.05 | *Additional Moments* | | |
| $y_{33}^{East}$ | Rel income in East Germany | 2.4 | Agricul. prod. gap | 1.5 | 1.516 |
| $\alpha$ | Spending share on agricult. | 0.24 | Earnings diff. of refugees | -0.075 | -0.0729 |
| $\xi$ | Dist. elasticity of trade costs | 0.32 | Share of outflows within states | 0.67 | 0.611 |
| | | | Distance elasticity of trade | -1.29 | -1.29 |
| | | | Std dev of resid. of regional $y$ growth | 0.041 | 0.037 |

Note: The table reports the structural parameters and the targeted moments in both the data and the model. The exogenously set parameters are $\sigma = 5$, $\gamma = 0.5$, $\theta = 2$, $\varrho = 0.9$ and $\delta = 0.1$.

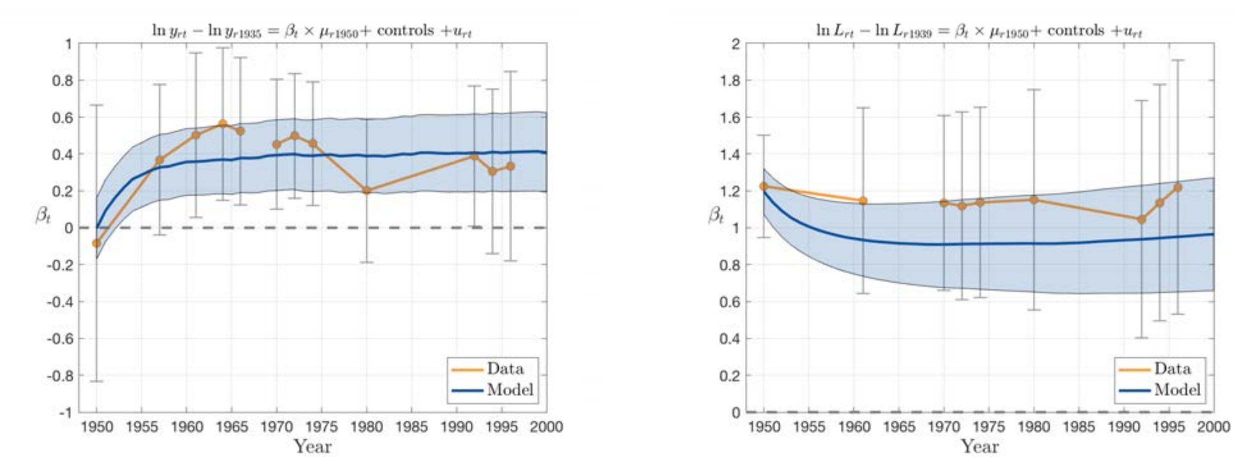
Table 9: Structural parameters & Model Fit

employment (rows 3 and 4) and the fact that the short-run effect on GDP per capita is small (row 5) but the long-run effect is robustly positive (rows 6 and 7). It also matches the spatial persistence of refugee flows (rows 8 and 9) and the correlation between refugee inflows and subsequent population outlows (row 10). Finally, the model also features a positive correlation between income growth and the distance to East Germany due to the loss of market access.

As a graphical description of the fit of the model, consider Figure 5, where I report the regression coefficient of income growth (left panel) and population growth (right panel) on the share of refugee share (and the usual controls as in specifications 2 and 6 in Table 5) at different time horizons. The blue line stems from the calibrated model, the orange line depicts the data. For both the model and the data I also plot the respective 95% confidence intervals. Figure 5 highlights that the model captures the main features of the persistent effects on income growth and population growth.

In terms of the implied structural parameters, I estimate the inter-temporal knowledge elasticity $\lambda$ to be 0.71 and the elasticity of substitution $\rho$ to be 5. Hence, the short-run scale elasticity in the non-agricultural sector is equal to $\vartheta = \frac{1}{\rho-1} = 0.25$ and the long-run scale elasticity is $\frac{1}{1-\lambda} \approx 3.5$ times as large.[30] As highlighted above, the estimate of $\lambda$ is

[30] Note that $\vartheta \approx 0.25$ is not directly comparable to typical estimates of regional agglomeration as it applies




Note: The figures report the coefficient $\beta$ of the regression $y_{rt} = \beta_t \mu_{r1950} + x'_{rt}\gamma + u_{rt}$for different time horizons and for income growth (left panel) and population growth (right panel) as dependent variables. The vector $x_{rt}$controls for state fixed effects, population density in 1939, war-time destruction, log income per capital (population) in 1939 and the distance to the inner german border (see columns 2 or 6 of Table 5). For both the model and the data I also report 95% confidence intervals.

Figure 5: Model Fit: The Dynamic Effects on Income and Population Size

tightly linked to the importance of scale effects at the aggregate level (Jones, 1999). My finding of $\lambda < 1$ implies that growth is semi-endogenous so that population shocks increase the level of productivity but not the long-run growth rate.[31] This is consistent with existing empirical work, that has mostly relied on time-series data and also points towards models of semi-endogenous growth.[32]

To match the persistence of the refugee settlement, the model implies a moving hazard of $\psi = 0.07$. The remaining parameters are also in line with the existing findings in the literature. The migration elasticity $\varepsilon \approx 2.12$ is consistent with Allen and Donaldson (2020) and Monte et al. (2018), who report estimates between 2 and 4 and the distance elasticity of migration cost $\kappa = -1.1$ is in the ballpark of the findings of Allen and Donaldson (2020) and Bryan and Morten (2019), whose estimates are between 0.7 and 2.

**Persistence and Path Dependence** As highlighted by Allen and Donaldson (2020), population shocks can have persistent effects if they lead to higher productivity. Environments particularly prone to such effects feature strong degrees of agglomeration and small spatial

only to the manufacturing sector. Given the decreasing returns in agriculture, the "overall" short-run scale elasticity at the local level is below 0.25.

[31] It is, however, important to point out that this cross-sectional evidence is not necessarily conclusive. If, for example, ideas were to diffuse across space, the cross-sectional evidence could underestimate the aggregate scale elasticity. Alternatively, the cross-sectional elasticity could be an overestimate, if for example local population shocks were to lead to a spatial reallocation of firms rather than new firm creation. Both of these channels are mute in my theoretical framework.

[32] Economic growth seems unrelated to the number of researchers or the size of the population (Jones, 1995) and research productivity seems to be declining over time (Bloom et al., 2020).

dispersion forces. In the context of my model, persistence requires $\lambda$ and $\vartheta$ to be large, $\varepsilon$ and $\sigma$ to be large and $\beta$ and $\psi$ to be small.

As I show in detail in Section A-2.6 in the Appendix, the estimated parameters reported in Table 9 put the model in a range, where such persistence occurs. More specifically, for any history of shocks, the model converges to a steady state but this steady-state depends on initial conditions. The main data moment that pushes towards a parametrization with persistent effects is the large long-run elasticity between refugee inflows and population size. This feature of the model implies that the historical policy of settling refugees in rural locations might have affected the long-run path of industrialization in rural labor markets, a topic I will come back to in Section 5.5.

**The Assumption of Myopic Agents** My estimation methodology requires me to solve the model's transitional dynamics for different histories of shocks and to then run the empirical regressions of Section 3. A key feature of my theory that facilitates the computational implementation is that the difference equations that describe the evolution of the endogenous state vector $\{N_{rt}, \mathscr{L}_{rt}\}$ are backward looking - see (9) and (10). This property is a direct consequence of my assumption that individuals behave myopically. In Section A-3.1 in the Appendix I characterize my model with forward-looking agents and solve for the analogues to (9) and (10).[33] In that case, the equilibrium mobility and entry decisions depend on the entire distribution of future wages. This makes estimating the model while still preserving the rich spatial heterogeneity to connect the theory to my empirical analysis in Section 3 computationally challenging.

As I discuss in more detail in Section A-3.1 in the Appendix, however, for the questions of interest of this paper, my estimation based on short-lived agents might still lead to informative results. Because the static equilibrium is not affected by the myopia assumption, a model with forward-looking agents would have the same implications for local wages and employment shares if it were to match the same path of state variables $\{N_{rt}, \mathscr{L}_{rt}\}$. Many of my targeted moments, in particular the estimated response of population growth and income growth shown in Figure 5, are of course tightly linked to precisely these state variables. Hence, many implications might look quite similar if a model with forward-looking agents was calibrated to the same moments. Of course, the implied structural parameters would be different and a fully forward-looking model would respond differently to other shocks (e.g. the announcement of a refugee inflow in the future) and have different welfare consequences.

[33] In Section A-3.1 in the Appendix I also provide a more detailed comparison with the contributions by Desmet et al. (2018) and Caliendo et al. (2019), both of which allow for forward-looking agents.

**Robustness of Quantitative Results** In Section A-2.8 in the Appendix I discuss two important extensions for the robustness of my results. First, as highlighted above, in 1950 I had to rely on data for value added taxes because data on local GDP per capita only starts in 1957. This naturally raises the concern that the discrepancy between the short- and the long-run effects are in part driven by the differences in income growth measures.[34] I therefore re-estimated the model without relying on income growth in 1950 as an estimation moment. Secondly, I also extend and re-estimated the model by allowing for innate human capital differences between refugees and natives. If, for example, refugees' skills might only have been partly transferable across space, such differences could account for their lower earnings and might change the mapping between refugee inflows and local income growth because of a deterioration of the local human capital stock. The re-estimated model shows that, quantitatively, both of these concerns are not very important. The estimated parameters are very similar and so is the match with the targeted moments.

## 5.3 Non-targeted Moments: Evidence on Spatial Sorting

To validate the model along an important non-targeted dimension note that equation (4) makes specific predictions about the spatial heterogeneity of the impacts of refugee inflows. Because the local population of refugees was not selected on their skills, the type composition did not vary across space, i.e. $\omega_{rt}^{RI} = \chi$. By contrast, the theory implies that the native population was spatially sorted whereby industrial types locate in regions that have a comparative advantage in the production of manufacturing goods. Hence, $\omega_{rt}^{NI}$ is positively correlated with the local manufacturing share in the pre-war period. This implies that the refugee bias, i.e. the relative employment share of refugees $\pi_{rMt}^{R} - \pi_{rMt}^{N}$, is particularly large in rural location. As a consequence, the effect of refugee inflows on local manufacturing employment should be particularly large in rural locations where $\omega_{rt}^{NI}$ is low.

In Table 10 I document these predictions both in the model and in the data. In the first four columns I report regression run in the model. The first column shows the sorting of natives: there is a strong correlation between the pre-war manufacturing share $\pi_{r1939}^{M}$ and the share of industrialists. Column two implements (4) and regresses the refugee bias $\pi_{rMt}^{R} - \pi_{rMt}^{N}$ on the pre-war manufacturing share. The refugee bias is particularly high in rural areas. Finally, the last two columns focus on the spatial heterogeneity of the impact of refugee inflows on local industrialization. For comparison, column three reports the basic cross-sectional relationship and in column four I allow the effect of the share of refugees $\mu_{r1950}$ to vary with the pre-

[34] As discussed in more detail in the Appendix, for a single year (1970) I was able to find both data on value added taxes and local GDP. I indeed find that an analysis based on the value added data yields smaller and less precisely estimated results.

| | Model | | | | Data | | |
|---|---|---|---|---|---|---|---|
| | Ind. share $\omega^I_{r1939}$ | Refugee bias | Manufac. growth $\pi^M_{r1950} - \pi^M_{r1939}$ | | Refugee bias | Manufac. growth $\pi^M_{r1950} - \pi^M_{r1939}$ | |
| $\pi^M_{r1933}$ | 0.654*** | -0.249*** | -0.056*** | 0.072*** | -0.062** | -0.129*** | -0.005 |
| | (0.009) | (0.010) | (0.008) | (0.013) | (0.027) | (0.044) | (0.059) |
| $\mu_{r1950}$ | | | 0.324*** | 0.600*** | | 0.292*** | 0.482*** |
| | | | (0.031) | (0.033) | | (0.081) | (0.110) |
| $\mu_{r1950} \times \pi^M_{r1939}$ | | | | -0.595*** | | | -0.692** |
| | | | | (0.052) | | | (0.325) |
| Controls | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 500 | 500 | 500 | 500 | 499 | 499 | 499 |
| $R^2$ | 0.986 | 0.934 | 0.565 | 0.653 | 0.413 | 0.386 | 0.403 |

Notes: Standard errors are clustered at the level of 37 *Regierungsbezirke*. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. All specifications control for state fixed effects, population density in 1939, the share of the destroyed housing stock, the distance to the inner german border and a fixed effect for whether a county is a border county.

Table 10: Spatial Sorting and Rural Industrialization

war manufacturing share. The model implies that the effect is *weaker* the higher the initial manufacturing share because refugees' manufacturing bias is smaller in these locations.

The three remaining columns run the same specifications in the data. Of course, column one does not have an empirical counterpart, because the type composition of the local workforce is unobserved. However, the remaining patterns between the local refugee bias and the heterogenous impact of the refugee settlement are qualitatively and quantitatively very similar although none of these aspects was targeted in the estimation.[35][36]

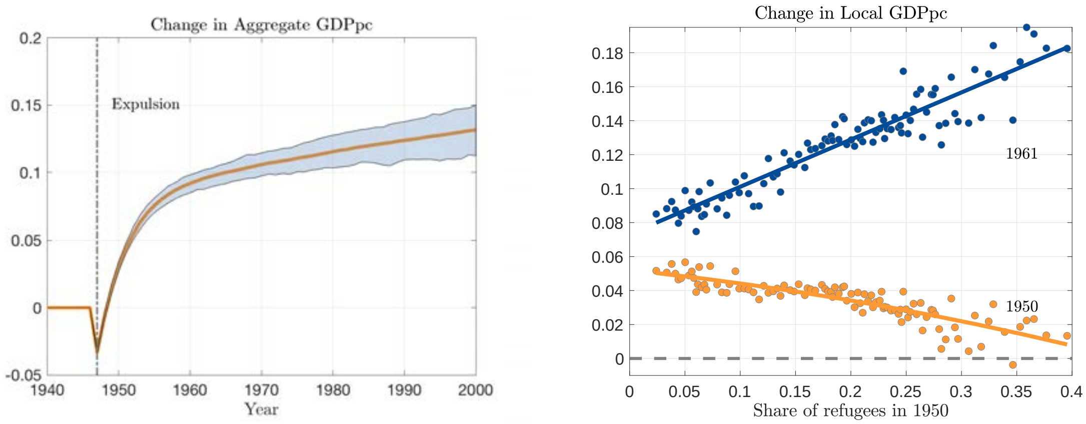
## 5.4 The Aggregate and Local Effects of the Refugee-Settlements

How large was the aggregate impact of the refugee settlement on economic activity in West Germany? This object is not identified from the cross-sectional regression due to the usual "missing-intercept" problem (see for example Adao et al. (2021), Chodorow-Reich (2019) and Wolf (2019)). However, it can be computed in the calibrated model by comparing the equilibrium with refugee inflows with a counterfactual West Germany where the refugees did not arrive.

In the left panel of Figure 6 I plot the time path of the percentage change in aggregate income per capita due to the refugee settlement. More precisely, for a given sequence of regional productivity shocks, I compute the effect of the refugee settlement on aggregate

[35] Recall that I only observe the refugee bias for the state of Bavaria, hence the smaller number of observations. However, for these 167 counties, the bias in the model and the data is highly correlated. A simple cross-sectional regression between data and model yields a coefficient of 0.38 with a standard error of 0.045.

[36] In Section A-2.10 in the Appendix, I provide additional evidence for this pattern of spatial sorting using the expenditure micro data from 1962.




Note: The left panel shows aggregate GDP per capita for the model with refugee inflows relative to a counterfactual economy without the refugee inflow. The orange line shows one particular sample path of the productivity process $\mathcal{Q}_{rt}$. The shaded area displays a 95% confidence interval from the bootstrap distribution. The right panel shows the spatial impacts in 1950 (orange) and 1961 (blue) as binned scatter plots for 100 percentiles of the refugee share in 1950. I calculate the spatial impact as $\left(y_{rt}^{with} - y_{rt}^{no}\right)/y_{rt}^{no}$, where $y_{rt}^{with}$ and $y_{rt}^{no}$ denote income per capita in the equilibrium with and without the refugee settlement.

Figure 6: The Aggregate and Spatial Impact of the Refugee Settlement

income. Redoing this experiment for a different sequence of local productivity shocks allows me to estimate the distribution of this aggregate impact and I plot both the average effect in orange and a 95% confidence interval in light blue.

The graph shows that the influx of refugees initially reduced GDP per capita by about 3%. This is mostly due to the fact that agricultural production is subject to decreasing returns. Due to the endogenous nature of technological progress this initial drop is short-lived and the population increase causes income per capita to increase. Given the estimated parameters, the average effect rises to about 8% in 1961 and reaches 12% by 2000.[37]

Note that the confidence interval around the aggregate GDP effect reflects two sources of uncertainty. First, the presence of productivity shocks implies that the aggregate impact of refugee inflows is a random variable. Intuitively, the aggregate impact of refugee inflows is larger along a sample path where locations with lots of refugees happen to receive positive productivity shocks. Second, as discussed above, my model features persistence, where the initial allocation potentially affects long-run outcomes. Because the confidence interval is computed from the distribution of outcomes of solving the model repeatedly with different histories of shocks, they capture both these sources of uncertainty. Quantitatively, they can

[37]To put these numbers into perspective, if there was only a single region, the elasticity of long-run income per capita with respect to population size would be given by $d\ln y/d\ln L = (1-\alpha)\,\vartheta/(1-\lambda) - \alpha\gamma$. Hence, the aggregate scale elasticity is an $\alpha-$weighted average between the long-run scale elasticity in manufacturing $\vartheta/(1-\lambda)$ and the returns to scale in agriculture $-\gamma$. Using the estimated parameters in Table 9, this expressions suggests an elasticity of 0.53 in the long-run. The inflow of refugees, which increased the aggregate population by around 18%, should have increased income per capita by about 10% in the long-run.

change the aggregate GDP impact by about one percentage point at the 50 year horizon.[38]

The cross-sectional estimates provide a misleading answer for the aggregate impact of the refugee settlement. Not only is the cross-sectional estimate between refugee inflows and GDP per capita in 1950 negative (even though the aggregate effect is positive) but the long-run estimates are also downward biased. The point estimate of 0.2 in 1980 for example suggests that a 18% increase in the share of refugees increases GDP per capita by 3.6%, even though the true aggregate impact is around 10%.

The reason is of course that non-treated regions also benefited from the refugee inflow in general equilibrium. This is shown in the right panel of Figure 6, where I depict the correlation between the counterfactual percentage change in income per capita and the share of refugees, both in 1950 (orange dots) and 1961 (blue dots).[39] In 1950 there is a negative correlation, in line with the negative cross-sectional estimate. However, the entire locus is shifted upwards due to general equilibrium effects that are differenced out in the empirical cross-sectional estimates. If we fast-forward by a decade and look at the impact on income per capita in 1961, we see a different picture. First, the relationship is now strongly positive, reflecting the slow accumulation of local productivity. Empirically, this slope reflects the positive cross-sectional relationship between refugee inflows and income per capita in the long-run. Second, the entire locus is further shifted upwards because regions that were initially non-treated benefit both from refugees' migration response and through trade linkages.

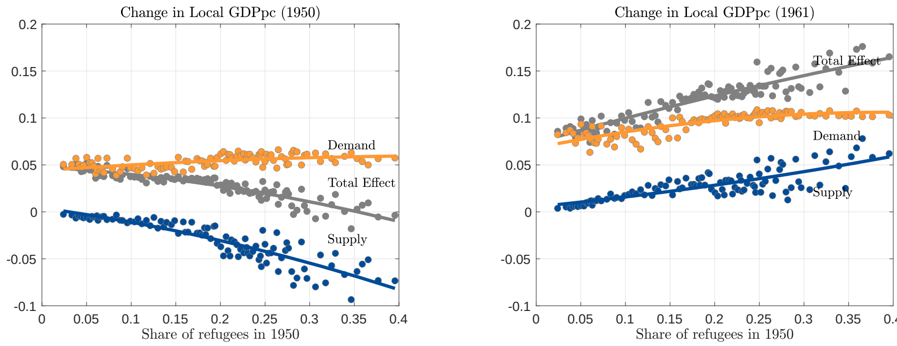
One way to rationalize these patterns is as the combination of supply and demand forces. In order to decompose the importance of these two effects, I define the supply effect for region $r$ as the counterfactual change in income per capita if *only* region $r$ experienced an inflow of refugees. Conversely, I define the demand effect for region $r$ as the counterfactual change if every region *but* region $r$ experienced an inflow of refugees. In the first scenario, demand is - almost - unaffected if region $r$ is small. In the second scenario, region $r$ directly benefits from "foreign" demand and only experiences changes in labor supply dynamically once the inflowing refugees start relocating within Germany.

In Figure 7 I depict the results of conducting these experiments for each of the 500 regions in my sample. The demand effect is depicted in orange, the supply effect is depicted in blue. For comparison I also depict the total effect shown in Figure 6.[40] In the short-run, the supply

[38] In Section A-2.6 the Appendix, I also compare the model with a parametrization that does not feature persistence and show that the aggregate implications are quantitatively similar.

[39] I calculate the equilibrium path for a given realization of exogenous productivity shocks with and without the refugee inflow and calculate the percentage difference between income capita for region $r$, i.e. $\frac{y_{rt}^{with}-y_{rt}^{no}}{y_{rt}^{no}}$, where $y_{rt}^{with}$ ($y_{rt}^{no}$) denotes income per capita in the equilibrium with (without) the refugee settlement.

[40] Due to non-linearities, the sum of the supply and the demand effect is not numerically equivalent to the total effect. In practice, however, they are almost indistinguishable.




Note: The figure shows the change in income per capita in 1950 (left row) and 1961 (right row) as binned scatter plots for 100 percentiles of the refugee share in 1950. In each case it displays the total effect, the supply effect (i.e. if refugees had only arrived in the particular region) and the demand effect (i.e. if refugees had only arrived in all other regions).

Figure 7: The Spatial Impact of Refugee Inflows: Demand vs Supply

effect is negative and explains most of the cross-sectional variation. Expectedly, the supply effect is zero for a county that did not receive any refugees. By contrast, the demand effect is positive, only weakly correlated with the refugee-share and thus plays the role of the "missing intercept".[41] The right panel shows that the supply effect also explains regional differences in income *growth* between 1950 and 1961.[42]

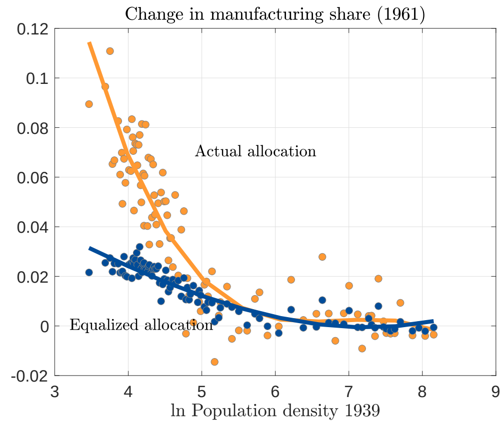
## 5.5 Persistence of the Historical Allocation and Rural Industrialization

The persistent effects of the refugee settlement raise the intriguing question whether the government policy of settling refugee in rural labor markets might have changed the path of local industrialization in West Germany. To study the quantitative importance of this form of path dependency of the government policy, I compare the equilibrium with a counterfactual allocation rule where the initial share of refugees was equalized in 1950 but the size of the aggregate inflow is held constant. In Figure 8 I report - for both scenarios - the change in the local manufacturing share in 1961 relative to an allocation without the refugee settlement as a function of pre-war population density.

Figure 8 vividly shows how the specific historical allocation rule affected the process of

[41] The main reason why the demand effect is weakly positively correlated with the refugee share is that empirically the allocation of refugees is spatially correlated. In the presence of trade costs, this implies that refugee-rich counties experienced a slightly larger demand shock.

[42] See also Section A-2.7 in the Appendix where I provide more details for this exercise. There I also show that the effect on manufacturing employment is fully captured by the supply force. The demand effect is negative because other regions increase their supply of manufacturing products but is quantitatively very small. Hence, even though the manufacturing sector is the sole source of endogenous productivity gains, an increase in the local employment share does not necessarily go hand-in-hand with an increase in income per capita if local technology accumulates slowly.




Note: The figure shows the change in the local manufacturing employment share, relative to an allocation without refugee inflows, for the historical allocation (orange) and a counterfactual allocation policy that equalizes the initial share of refugees across counties.

Figure 8: Persistent Effects of the Refugee Settlement: Rural Industrialization

industrialization in rural counties in post-war Germany. Under the actually implemented allocation rule, low-density, rural communities were the dominant receivers of the inflowing refugee population and industrialized as a consequence. Quantitatively, the model implies that these inflows raised the local manufacturing share by around 5-7% in traditional rural communities. This specific form of rural industrialization would not have happened with a more equitable refugee allocation in 1950. If the US and the UK Government had been able to equalize the share of refugees in 1950 across counties, rural areas would have only experienced a 2% increase in their manufacturing share. While rural counties would still have industrialized faster even in the presence of an equalized allocation, the "rural bias" of the refugee settlement would have been far less pronounced. The specific rural nature of the historical allocation rule thus acted as a form of place-based policy that triggered local industrialization and might have played an important role in the emergence of the German manufacturing base that even today is often found in the countryside outside the large cities.[43]

# 6 Conclusion

The positive relationship between population size and productivity is at the heart of virtually all theories of economic growth. In this paper I analyzed a particular historical setting to provide direct evidence for the empirical relevance of such scale effects. I focused on the expulsion of the ethnic German population in the aftermath of the Second World War that

[43] In Section A-2.9 in the Appendix I analyze these two allocation rules in more detail. There I show that rural counties, in line with their faster industrialization, also experience faster income and population growth and that these effects are still visible 50 years after the initial settlement.

was implemented by the Military Governments of the US, the UK and Russia as part of the post-war restructuring of continental Europe. Between 1945 and 1948 almost 8m people were transferred to West Germany. At the time, this amounted to an increase in the population by about 20%.

Because regions in West Germany differed substantially in the extent to which they were exposed to the refugee settlement, I use the cross-sectional variation in refugee inflows to estimate the relationship between changes in population size, income per capita and industrialization in both the short- and long-run. I find that the refugee settlement led to persistent increases in the local population, the manufacturing share and income per capita. I then propose a parsimonious idea-based model of spatial growth and estimate its parameters by using the cross-sectional regression results of the natural experiment as identified moments. The model can rationalize the empirical findings both qualitatively and quantitatively and delivers a persistent effect of the refugee settlement if spatial mobility is subject to frictions and dynamic productivity spill-overs occur at the local level and are sufficiently potent. At the aggregate level, the settlement of refugees increased income per capita by about 12% after 25 years. Moreover, the government policy of settling refugees predominantly in the countryside had long-run effects and markedly increased rural industrialization.

A natural question is of course whether these results are quantitatively portable to predict the consequences of immigration episodes today. While I expect the basic mechanism to apply more generally, there are at least three aspects of this study that seem particularly context-specific. First and foremost, the German economy just emerged from the Second World War and firm creation might have been particularly mobile across space. Second, the refugees were allocated to rural areas and not to urban centers. This is in stark contrast to most episodes of voluntary migration both in the modern era and in the past. Finally, the 1950s and 1960s were characterized by a secular rise in the manufacturing sector. To the extent that the mechanisms highlighted in this paper are less potent in services, the productivity effects of changes in population size might be smaller today.

# References


**Acemoglu, Daron**, "Directed Technical Change," *Review of Economic Studies*, 2002, *69* (4), 781–809.

_ , "Equilibrium Bias of Technology," *Econometrica*, 2007, *175*, 1371–1410.

**Adao, Rodrigo, Costas Arkolakis, and Federico Esposito**, "General Equilibrium Effects in Space: Theory and Measurement," 2021.

**Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf**, "The Economics of Density: Evidence From the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

**Akcigit, Ufuk**, "Economic growth: The past, the present, and the future," *Journal of Political Economy*, 2017, *125* (6), 1736–1747.

**Allen, Treb and Costas Arkolakis**, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

**_ and Dave Donaldson**, "The Geography of Path Dependence," 2020. Working Paper.

**Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946**, *Volks- und Berufszählung com 29. Oktober 1946 in den vier Besatzungszonen und Gross-Berlin*, Berlin-München: Duncker und Humboldt, 1949.

**Bayerisches Statistisches Landesamt**, "Bayerische Gemeinde- und Kreisstatistik," in "Beiträge zur Statistik Bayerns," Vol. 132 J. Lindauersche Universitäts-Buchhandlung München 1944.

_ , "Bayerische Gemeinde- und Kreisstatistik," in "Beiträge zur Statistik Bayerns," Vol. 177 J. Lindauersche Universitäts-Buchhandlung München 1952.

_ , "Volks- und Berufszählung am 13. September 1950 in Bayern," in "Beiträge zur Statistik Bayerns," Vol. 171 J. Lindauersche Universitäts-Buchhandlung München 1953.

_ , "Bayerische Gemeinde- und Kreisstatistik," in "Beiträge zur Statistik Bayerns," Vol. 231 J. Lindauersche Universitäts-Buchhandlung München 1963.

_ , "Bevölkerung und Erwerbstätigkeit: Ergebnisse der Volks- und Berufszählung am 6. Juni 1961," in "Beiträge zur Statistik Bayerns," Vol. 231 J. Lindauersche Universitäts-Buchhandlung München 1963.

**Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong**, "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia," *American Economic Review*, 2016, *106* (9), 2658–98.

**_ , _ , _ , and _** , "Unity in diversity? How intergroup contact can foster nation building," *American Economic Review*, 2019, *109* (11), 3978–4025.

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

**Besser, Christoph**, "Die deutschen Vertriebenen in der Bundesrepublik Deutschland von 1945 bis 1969," Technical Report, GESIS Datenarchiv, Köln. histat., Studiennummer 8217 2007.

**Bilal, Adrien**, "The Geography of Unemployment," 2019. Working Paper.

**Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb**, "Are ideas getting harder to find?," *American Economic Review*, 2020, *110* (4), 1104–44.

**Brakman, Steven, Harry Garretson, and Marc Schramm**, "The Strategic Bombing of German Cities during WWII and its Impact in City Growth," *Journal of Economic Geography*, 2004, *4* (2), 201–18.

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

**_ and Toman Omar Mahmoud**, "The employment effects of immigration: evidence from the mass arrival of German expellees in postwar Germany," *The Journal of Economic History*, 2014, *74* (01), 69–108.

**Bryan, Gharad and Melanie Morten**, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

**Burchardi, Konrad B and Tarek Alexander Hassan**, "The Economic Impact of Social Ties: Evidence from German Reunification," *The Quarterly Journal of Economics*, 2013, *128* (3), 1219–1271.

**_ , Thomas Chaney, Tarek A Hassan, Lisa Tarquinio, and Stephen J Terry**, "Immigration, Innovation, and Growth," 2019. Working Paper.

**Burstein, Ariel, Gordon Hanson, Lin Tian, and Jonathan Vogel**, "Tradability and the Labor Market Impact of Immigration: Theory and Evidence for the U.S," 2017. Working Paper.

**Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro**, "Trade and labor market dynamics: General equilibrium analysis of the china trade shock," *Econometrica*, 2019, *87* (3), 741–835.

**Card, David**, "The Impact of the Marial Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 1990, *43* (2), 245–257.

**Chodorow-Reich, Gabriel**, "Regional Data in Macroeconomics: Some Advice for Practitioners," Technical Report, National Bureau of Economic Research 2019.

**Davis, Donald R. and David E. Weinstein**, "Bones, Bombs and Break Points: The Geography of Eonomic Activity," *American Economic Review*, 2002, *92* (5), 1269–89.

**Desmet, Klaus and Esteban Rossi-Hansberg**, "Spatial growth and industry age," *Journal of Economic Theory*, 2009, *144* (6), 2477–2502.

**_ and _** , "Spatial development," *American Economic Review*, 2014, *104* (4), 1211–43.

**_ , Dávid Krisztián Nagy, and Esteban Rossi-Hansberg**, "The Geography of Delevopment," *Journal of Political Economy*, 2018, *126* (3), 903–983.

**Douglas, Ray M**, *Orderly and humane: The expulsion of the Germans after the Second World War*, Yale University Press, 2012.

**Dustmann, Christian, Uta Schönberg, and Jan Stuhler**, "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 2017, *132* (1), 435–483.

**Edding, Friedrich**, "The Refugees as a Burden a Stimulus, and a Challenge to the West German Economy," in "Publications of the Research Group for European Migration Problems," Hague, M. Nijhoff, 1951.

**Faber, Benjamin and Cecile Gaubert**, "Tourism and Economic Development: Evidence from Mexico's Coastline," *American Economic Review*, 2019, *109* (6), 2245–93.

**Fuchs-Schündeln, Nicola and Tarek Hassan**, "Natural Experiments in Macroeconomics," in "The Handbook of Macroeconomics," Elsevier, 2015.

**Gollin, Douglas, David Lagakos, and Michael E. Waugh**, "The Agricultural Productivity Gap," *The Quarterly Journal of Economics*, 2014, *129* (2), 939–993.

**Grosser, Thomas and S. Schraut**, *Fluechtlinge und Heimatvertriebene in Wuerttemberg-Baden nach dem Zweiten Weltkrieg: Dokumente und Materialien zu ihrer Aufnahme und Eingliederung.*, Mannheim, 2001.

**Hessisches Statistisches Landesamt**, "Sonderrreihe: Berufszählung 1950. Wirtschaftliche und soziale Gliederung der Bevölkerung in den hessischen Kreisen.," in "Beiträge zur Statistik Hessens" Hessisches Statistisches Landesamt Wiesbaden 1952.

_ , "Amtliches Verzeichnis der Gemeinden in Hessen: Endgültige Ergebnisse d. Volkszählung vom 6. Juni 1961," in "Beiträge zur Statistik Hessens" Hessisches Statistisches Landesamt Wiesbaden 1962.

**Hobijn, Bart, Todd Schoellman, and Alberto J. Vindas Q.**, "Structural Transformation:A Race between Demographics and Technology," 2018. Mimeo, Arizona State University.

**Hornung, Erik**, "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *The American Economic Review*, 2014, *104* (1), 84–122.

**Jones, Charles**, "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 1995, *103*, 759–784.

— , "Growth: With or Without Scale Effects?," *American Economic Review (P&P)*, 1999, *89*, 139–144.

— , "Growth and ideas," *Handbook of economic growth*, 2005, *1*, 1063–1111.

— , "The End of Economic Growth? Unintended Consequences of a Declining Population," 2019. Working Paper.

**Kossert, Andreas**, *Kalte Heimat: die Geschichte der deutschen Vertriebenen nach 1945*, Siedler Verlag, 2008.

**Krugman, Paul**, "Scale Economies, Product Differentiation, and the Pattern of Trade," *The American Economic Review*, 1980, pp. 950–959.

**Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg**, "Commuting, migration, and local employment elasticities," *American Economic Review*, 2018, *108* (12), 3855–90.

**Nagy, Dávid Krisztián**, "City Location and Economic Development," *Working Paper*, 2017.

— , "Quantitative Economic Geography Meets History: Questions, Answers and Challenges," *Regional Science and Urban Economics*, 2021, p. 103675.

**Nakamura, Emi and Jón Steinsson**, "Identification in Macroeconomics," *Journal of Economic Perspectives*, 2018, *32* (3), 59–86.

**Nellner, Werner**, "Grundlagen und Hauptergebnisse der Statistik," in Eugen Lember and Friedrich Edding, eds., *Die Vertriebenen in Westdeutschland: ihre Einglierung und ihr Einluss auf die Gesellschaft, Wirtschaft, Politik und Geistesleben*, Verlag Ferdinand Hirt, 1959.

**Niedersächsisches Amt für Landesplanung und Statistik**, "Volkszählung und Wohnungszählung am 13. September 1950," in "Gemeindestatistik für Niedersachsen" Niedersächsisches Amt für Landesplanung und Statistik Hannover 1952.

— , "Gemeindestatistik Niedersachsen 1960/61. Teil 1: Bevölkerung und Erwerbstätigkeit," in "Gemeindestatistik für Niedersachsen" Niedersächsisches Landesverwaltungsamt (Hrsg.) Hannover 1964.

**Office, Germany Federal Statistical**, *Statistical Pocket-Book on Expellees in the Federal Republic of Germany and West Berlin* 1953.

**Office of the Military Government for Germany**, "A year of Potsdam: The German economy since surrender," Technical Report, Military Governor 1945.

— , "Information Bulletin No. 169," Technical Report, Provisional Office of Administration, Reports & Statistics Branch, US Army 1949.

**Peri, Giovanni**, "Immigrants, Productivity, and Labor Markets," *The Journal of Economic Perspectives*, 2016, *30* (4), 3–29.

**Peters, Michael and Conor Walsh**, "Population Growth and Firm Dynamics," *Working Paper*, 2020.

**Porzio, Tommaso and Gabriella Santangelo**, "Does Schooling Cause Structural Transformation?," 2019. Working Paper.

**Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez**, "Trade, domestic frictions, and scale effects," *The American Economic Review*, 2016, *106* (10), 3159–3184.

**Redding, Stephen J. and Daniel M. Sturm**, "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, December 2008, *98* (5), 1766–97.

_ **and Esteban Rossi-Hansberg**, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

**Reichling, Gerhard**, *Die Heimatvertriebenen im Spiegel der Statistik*, Berlin: Duncker & Humboldt, 1958.

**Romer, Paul M.**, "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), 71–102.

**Sequeira, Sandra, Nathan Nunn, and Nancy Qian**, "Immigrants and the Making of America," *Review of Economic Studies*, 2019.

**Statistische Landesämter**, "Das Bruttoinlandsprodukt der kreisfreien Städte und Landkreise 1957-1966," Technical Report 3, Gemeinsschaftsveroeffentlichung der Statistischen Landesaemter 1968.

_ , "Das Bruttoinlandsprodukt der kreisfreien Städte und Landkreise 1970-1976," Technical Report 8, Gemeinsschaftsveröffentlichung der Statistischen Landesämter 1976.

_ , "Das Bruttoinlandsprodukt der kreisfreien Städte und Landkreise 1980-1992," Technical Report 31, Gemeinsschaftsveröffentlichung der Statistischen Landesämter 1992.

**Statistisches Bundesamt**, "Statistik der Bundesrepublik Deutschland," in "Gebäude- und Wohnungszählung in der Bundesrepublik Deutschland vom 13. 9. 1950," Vol. 38, W. Kohlhammer GmbH, 1950.

_ , "Die Umsätze der Umsatzsteuerpflichtigen und deren Besteuerung," in "Statistik der Bundesrepublik Deutschland," Vol. 112 Kohlhammer Stuttgart 1955.

_ , "Die Vertriebenen und Flü in der Bundesrepublik Deutschland 1946-1953," in "Statistik der Bundesrepublik Deutschland," Vol. 114 Kohlhammer Stuttgart 1955.

_ , "Die Wanderungen im Jahre 1955," in "Statistik der Bundesrepublik Deutschland," Vol. 171 Kohlhammer Stuttgart 1955.

_ , "The regionale Verteilung der Industriebetriebe," in "Die Industrie der Bundesrepublik Deutschland," Vol. 4, W. Kohlhammer GmbH, 1957.

**Statistisches Landesamt Nordrhein-Westfalen**, "Gemeindestatistik des Landes Nordrhein-Westfalen," in "Beiträge zur Statistik des Landes Nordrhein-Westfalen," Vol. Sonderreihe Volkszälung 1950 1952.

_ , "Gemeindestatistik des Landes Nordrhein-Westfalen," in "Beiträge zur Statistik des Landes Nordrhein-Westfalen," Vol. Sonderreihe Volkszälung 1961 1964.

**Statistisches Landesamt Rheinland-Pfalz**, "Die Wohnbevölkerung in Rheinland-Pfalz 1950," in "Statistische Bände," Vol. 17 Statistisches Landesamt Rheinland-Pfalz Bad Ems 1950.

_ , "Die Bevölkerung in Rheinland-Pfalz im Jahre 1961," in "Statistische Bände," Vol. 114 Statistisches Landesamt Rheinland-Pfalz Bad Ems 1961.

**Statistisches Reichsamt**, "Volks-, Berufs- und Betriebszählung vom 16. Juni 1933," in "Statistik des Deutschen Reichs," Berlin, Verl. für Sozialpolitik, Wirtschaft u. Statistik P. Schmidt, 1936.

— , "Umsatzsteuerstatistik 1935," in "Statistik des Deutschen Reichs," Berlin, Verl. für Sozialpolitik, Wirtschaft u. Statistik P. Schmidt, 1938.

— , "Volks-, Berufs- und Betriebszählung vom 17. Mai 1939," in "Statistik des Deutschen Reichs," Berlin, Verl. für Sozialpolitik, Wirtschaft u. Statistik P. Schmidt, 1939.

**Walsh, Conor**, "Firm Creation and Local Growth," 2019. Working Paper.

**Wolf, Christian**, "The Missing Intercept: A Demand Equivalence Approach," 2019. Working Paper.

**Wolf, Nikolaus**, "Was Germany ever united? Evidence from intra-and international trade, 1885-1933," *The Journal of Economic History*, 2009, pp. 846–881.