Market Size and Spatial Growth - Evidence from Germany's Post-War Population Expulsions

Michael Peters*

August 2019

Abstract

Can increases in the size of the local workforce raise productivity and spur economic development? This paper uses a particular historical episode to study this question empirically. After the Second World War, between 1945 and 1948, about 8m Ethnic Germans were expelled from their domiciles in Middle and Eastern Europe and transferred to Western Germany. At the time, this inflow amounted to almost 20% of the Western German population. Using variation across counties I show that refugee inflows are positively correlated with income per capita and manufacturing employment. Importantly, the long-run effects are much larger than the short-run effects. I show that these findings are quantitatively consistent with a parsimonious model of spatial growth. The model makes tight predictions on the spatial distribution of economic activity in the long-run and shows that the productivity effects of a shock to the local population can build up slowly if labor mobility is limited. The theory also shows that the spatial distribution of economic activity is stationary if and only if growth is semi-endogeneous. When calibrated to the empirical estimates, the model implies that the inflow of refugees increased aggregate income per capita by 7% after 15 years.

*eMail: m.peters@yale.edu. I am very grateful to Julieta Caunedo, Penny Goldberg, Tarek Hassan, Chad Jones, Pete Klenow, Virgiliu Midrigan, Giuseppe Moscarini, Steve Redding and Alwyn Young for their comments. I particularly thank Steve Redding and Ben Faber for their insightful discussions at the NBER SI EFEG and ITI meetings. I also thank seminar participants at Chicago, Columbia, Cornell, Harvard, MIT, NYU, Princeton, UCSD, Stanford and UCLA for their comments and the staff at the GESIS Institute in Mannheim, especially Bernhard Schimpl-Neimanns, for their hospitality and help in accessing the historical microdata of the 1971 census.

1 Introduction

Does local productivity respond to changes in factor supplies? There are ample theoretical reasons to believe that the answer to this question ought to be yes. Standard theories of growth, for example, predict a positive relationship between innovation incentives and local factor supplies due to the presence of market size effects. Theories of directed technological change imply that innovation efforts are directed towards abundant factors. And many models of trade, development and economic geography incorporate agglomeration forces, whereby local productivity depends positively on population density, presumably as a reduced form for such considerations.

In this paper I study a particular historical episode to provide direct evidence for the importance of such mechanisms. 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 Middle and Eastern Europe and transferred them to Western 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 Western Germany. Given the population at the time, this amounted to an increase in the total population of about 20%. Importantly, the extent to which the inpouring refugees were settled in different regions was very heterogeneous. While some counties saw their population almost double, other counties were far less affected. It is this cross-sectional variation in refugee inflows across counties which I exploit to study the link between population inflows and the endogenous response of productivity.

Two features of the historical context make this empirical variation particularly interesting for the question of interest. The first is the sheer size of the initial population shock. As many of the theoretical mechanisms above stress the importance of general equilibrium effects, one requires changes in factor supplies, which are sufficiently large to plausibly have aggregate consequences. The second concerns the obvious endogeneity problem that the incoming refugees might have settled in locations with a favorable growth prospects. The specifics of the historical context allow me to address this concern. With millions of refugees being transferred, the Western German population in 1950 exceeded its pre-war level by about 13%. 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 80%. The dominant consideration for the Military Governments of the US and the UK, which were in charge to allocate the inpouring refugees, was therefore the availability of housing rather than economic prospects of the local economy. 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 political allocation rule and by using an instrumental variable strategy, which exploits the distance to the pre-war population centers in Eastern and Central Europe and aspects of the local housing supply.

Using a newly digitized dataset on the regional development for roughly 500 Western German counties between 1930 and 1970, I then analyze the effect of refugee inflows on industrialization and local productivity in the short- and long-run. First of all, I show that the initial allocation of refugees was very persistent - by 1961, i.e. 10 - 15 years after the actual expulsion, counties who received more refugees were still substantially larger and the share of refugees was still higher. Second, I establish a positive relationship between the allocation of refugees and subsequent manufacturing employment growth in the 1950s and 60s. This expansion of the manufacturing sector stems exclusively from a commensurate decline in the local agricultural employment share. Third, I provide evidence for a positive relationship between the inflow of refugees and local GDP per capita in the 50s and 60s. Importantly, the effect on income per capita is considerably larger in the long-run than in the short run. Finally, I use microdata on employment histories of refugees and natives after the war to document that refugees were much more likely to work in the manufacturing sector and hence acted as a shift of the supply of manufacturing workers at the local level.

I then show that this evidence is qualitatively and quantitatively consistent with a parsimonious model of spatial growth. At the heart of the theory is an explicit model of regional productivity, which is determined endogenously and responds to the size of the local manufacturing workforce. More specifically, I embed a standard model of economic geography where individuals are spatially mobile but migration is subject to frictions into a canonical idea-based model of economic growth in the tradition of Romer (1990).[1] The model generates the joint evolution of the spatial distribution of workers and regional productivity as an equilibrium outcome and makes tight predictions for the effects of population shocks on local productivity in the short- and the long-run.

The theory highlights that the long-run effects of a population shock on the spatial distribution of economic activity depends crucially on a single parameter, which I term the *inter-temporal knowledge elasticity.* This elasticity determines how quickly the costs of creating new ideas decline in the existing stock of ideas. If this elasticity is large, in particular if it exceeds unity, the model predicts that population shocks, even if they are transitory, have permanent effects on local productivity and that the spatial distribution of economic activity is not stationary. If in contrast, this elasticity is below one, the model has a unique stationary solution, which is fully determined from parameters. In particular, population shocks will not have permanent effects (but their effects can be very long lived).

I then show that this result is exactly the spatial-analogue of the main result of Jones (1995). Jones (1995) highlights the importance of the returns to scale in knowledge production to distinguish models of endogenous and semi-endogenous growth in the time series. The inter-temporal knowledge elasticity plays the exact same role. If this elasticity is equal to unity, my model has strong scale effect and features fully endogenous growth. If this elasticity is below unity, the model is a model of semi-endogenous growth. Hence, the very same "knife-edge" condition, which determines the presence of strong scale effects in the time series, also determines whether the spatial distribution of economic activity is stationary in the long-run and whether population shocks have permanent effect on the level of local productivity.

A key feature of the model is that the endogenous nature of productivity acts as an amplifying force. This implies that there is 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 past productivity constant. This elasticity depends on the elasticity of substitution across varieties, which parametrizes the strength of variety gains, and is isomorphic to agglomeration externalities commonly used in quantitative spatial models (see e.g. Redding and Rossi-Hansberg (2017)). The long-run elasticity describes the relationship between productivity and the local population along a spatial balanced growth path, where the population distribution is stationary and all regions grow at the same rate. This long-run elasticity always exceeds the short-run elasticity and depends on the inter-temporal knowledge elasticity highlighted above. Moreover, this long-run elasticity is finite if and only if the model is consistent with semi-endogenous growth, reflecting the fact that the spatial productivity distribution is non-stationarity in case growth is fully endogenous.

I estimate the structural parameters of the model using the variation from the natural experiment. In particular, I identify the structural parameters by targeting the regression coefficients via indirect inference. The main moments of interest are the relationship between refugee inflows and income per capita at different time horizons and the persistence of the spatial allocation of the refugee population. The long-run response of income per capita is particularly important because it identifies the inter-temporal knowledge elasticity. My empirical estimates imply that I can comfortably reject the case of fully endogenous growth in favor a parametrization where growth is semi-endogenous. While this implies that a population shock will not have effects on the spatial distribution of economic activity in the very long-run, my estimates show that this long-run can indeed be very long. In particular, I find

[1] While this modeling choice is not essential for the theory and I could have modeled local productivity as stemming from process innovation, I show empirically that the number of manufacturing plants increases in response to the inflow of refugees.

that the effect of refugee inflows on local income per capita is increasing during the 1950s, peaks in the mid 60s and then slowly reverts back to zero over the next decades.

Finally I use the calibrated model to quantify the aggregate effect 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, in particular the extent to which population inflows in one region reduce equilibrium entry in another region, i.e. a form of innovation crowd-out. The quantitative results reflect this ambiguity. In particular, I find that the inflow of refugees reduced income by capita by about 3% in the short-run but increased income per capita by 7% in the long-run, i.e. after about 10 years. The reason is that the dynamic response of entry is sluggish and takes time to adjust to the population inflow.

Related Literature On the theoretical side, the paper is related to a large literature in economic growth, which argues that innovation incentives respond to changes in factor supplies. While this is true for many models of growth (see, for example, the survey articles by Jones (2005) or Akcigit (2017)), this reasoning is at the heart of the literature on directed technological change and the bias of innovation (see e.g. Acemoglu (2002, 2007, 2010)), the relationship between economic integration and growth (Rivera-Batiz and Romer, 1991) or the interaction between market size and specialization (Krugman, 1980a). Of particular relevance is the semi-endogenous growth model by Jones (1995). I show that the distinction between endogenous and semi-endogenous growth has a clear spatial counterpart: the spatial analogue of strong scale effects in the time series is whether or not the spatial distribution of economic activity is stationary in the long-run.

The paper is also related to the recent literature on dynamic models of trade and economic geography. Of particular relevance are the papers by Desmet et al. (2015), Desmet and Rossi-Hansberg (2014) and Nagy (2017), all of which present growth models with a realistic geography, where local innovation incentives respond to local factor supplies. The main difference is that I assume agents to be myopic. This simplification allows me to incorporate mobility frictions so that the regional distribution of the population becomes a slow-moving dynamic state variable and I can estimate the structural parameters from the dynamic response of the model to a population shock. Both these features are of first-order importance for the question at hand. Allen and Donaldson (2018) also study the dynamic interaction between spatial mobility and local productivity but adopt a reduced-form specification between productivity and the local labor force. Caliendo et al. (2019) analyze a dynamic model of trade and migration but assume that regional productivity is exogenous.

There is also a close connection with a large literature on economic geography, which posits the existence of exogenous agglomeration economies - see for example Fajgelbaum and Redding (2014); Ahlfeldt et al. (2015); Allen and Arkolakis (2014); Faber and Gaubert (2016); Kucheryavyy et al. (2016); Ramondo et al. (2016) or the recent survey by Redding and Rossi-Hansberg (2017). These reduced-form specifications suggest a link between local productivity and the size of the population, which is time-invariant. My results on the discrepancies between the long-run and the short-run scale elasticities show that this perspective might be misleading.

There is a large literature which uses the German context as a source of historical experiments. Burchardi and Hassan (2013) use the settlement of refugees coming from the Soviet Occupied Zone and the interaction with the fall of the Berlin Wall in 1989 to measure the importance of social ties. Ahlfeldt et al. (2015) exploit the partition of Berlin as a shock to the distribution of economic activity to estimate the strength of agglomeration forces within city-blocks in Berlin. Redding and Sturm (2008) use the division of Germany as a shift in market access and study how the spatial equilibrium responds. Finally, this historical episode of the population expulsions has also been analyzed in Braun and Mahmoud (2014) and Braun and Kvasnicka (2014). In contrast to my paper, these contributions do not focus on the effect of refugee inflows on local productivity and also do not study the spatial variation across counties.

The paper also connects to the literature on immigration. Burstein et al. (2017) study the effects of immigration on native employment but do not focus on the possibility of immigrants to affect local productivity. Card (1990) uses the unexpected shock of the Miami-Boatlift to study the effects of immigrants on the labor market in Miami. This paper and many other papers in that literature (see e.g. Peri (2016) or Dustmann et al. (2016)) are mainly concerned with the short-run impact of immigrants rather the long-run impact on regional outcomes. Exceptions are Nunn et al. (2017), who study the long-run effects of immigration in the US and Hornung (2014), who uses data on textile plants to analyze the productivity effects of the Huguenot re-settlement for the 18th century. All of these studies are mostly empirical in nature and do not attempt a structural analysis.

The remainder of the paper is structured as follows. In the next section I describe the historical setting, the political environment leading to the population expulsions and the initial allocation of refugees across counties in Western Germany. Section 3 contains the main empirical analysis. Section 4 contains the theoretical model, which I estimate in Section 5. Section 6 concludes. A long Appendix contains a variety of robustness checks and the derivations of the main theoretical results.

2 The Historical Setting

Germans in Eastern and Middle Europe before 1939

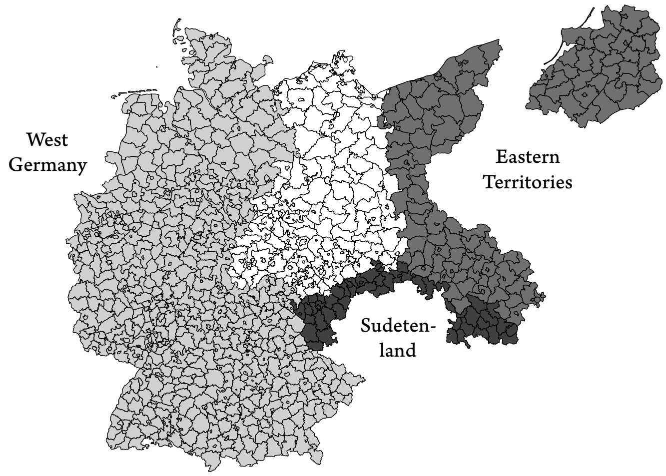
The presence of Germans in Middle and Eastern Europe is by no means a novel phenomenon. In fact, the settlement of ethnic Germans in Eastern Europe dates back to the Middle Ages. At the beginning of the Second World War in the summer of 1939, there are two groups to distinguish. On the one hand, there are large parts of todays' Poland and Russia, which used to be part of the German Reich. This area, known as the East German Territories, encompasses for example the regions of East Prussia and Silesia. On the other hand, there were sizable German minorities in other countries of Eastern Europe, most importantly the so-called Sudetenland in Czechoslovakia. This region in the north of Czechoslovakia has a long tradition of German settlements and was annexed by the Nazi Government in 1938. In Table 1 I report the size of the ethnic German population for different countries in Eastern and Middle Europe in 1939. Naturally, there is a large German population in the East German Territories. In 1939 almost 10m people resided in these areas. In addition, 3.5m million ethnic Germans lived in Czechoslovakia, many in the so-called Sudetenland, and many other countries like Poland, Hungary and Romania had substantial German minorities. Hence, almost 17m Germans lived in Central and Eastern Europe and would be affected by the ensuing expulsion.

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 East and Central 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 Central and Eastern Europe in 1939

These regions are shown on the map in Figure 1, which displays the territory of the German Reich on the eve of the Second World War. I also display the individual counties, which is the source of cross-sectional variation I will be using for this paper. In the West, shown in a light shade, is the area which is going to become West Germany in 1949. These regions form the main part of the analysis in this paper, as I will measure post-war outcomes in the 50s



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 Western 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 to-be GDR.

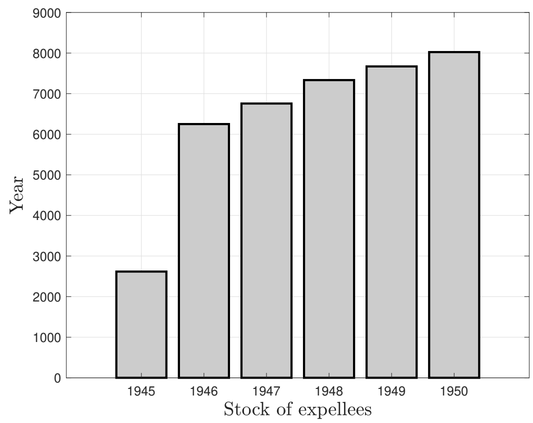
Figure 1: The German Reich in 1939

and 60s in these regions. In the far East, shown in medium dark, are the "Eastern Territories of the German Reich". This is the part of the German Reich, which will no longer be part of Germany after 1945. In the south-east, shown in dark, is the aforementioned Sudetenland in the north of Czechoslovakia. Finally, in the middle is the area of the German Reich, which will become the Soviet Occupied Zone (in 1945) and then turn into the German Democratic Republic (in 1949). This area will not be part of the analysis in this paper. Not shown on the map, there are are additional smaller German minorities living in other countries in Eastern Europe, in particular Poland, Hungary and Romania, reported in Table 1.

The Expulsions and the Potsdam Conference in 1945

The Second World War marks a drastic change in the geography of Europe and Germany in particular as it marks an end to the presence of the German population in Central and Eastern Europe. First of all, Germany lost the Eastern Territories and the Sudetenland as means of war reparations. Secondly, almost the entire German population residing in other countries in Eastern Europe either fled or were expelled in the aftermath of the war. The population transfer is one of the largest transfers in world history as between 1945 and 1950, roughly 12 million ethnic Germans were expelled and 8 million people were allocated to Western Germany (Reichling, 1958, p. 17).

The expulsion can be broadly divided into three phases. The first wave of refugees arrived in Western Germany during the last months of the war. Soviet forces made their appearance at the eastern German border in the summer of 1944. Trying to reach Berlin, soviet soldiers were advancing through the German Eastern Territories at great speed causing the German population to flee westwards. As the Nazi government considered the evacuation of German territories a defeatist act most inhabitants evacuated their homes fully unprepared. Because there were hardly any official evacuation plans as trains and ships were often reserved for the German soldiers, most refugees fled their homes by joining refugee treks, which suffered enormous casualties during the flight. After the German defeat in May 1945, the so-called wild expulsions started. These where mainly taking place in the spring and summer of 1945 before the Potsdam Agreement was signed in August 1945, most importantly in Poland and



Notes: The figure shows the stock of refugees in Western Germany by year. Source: Federal Statistical Office (1953)

Figure 2: Expellees' Arrival in Western Germany

Population May 17 1939	Population Losses 1939-50				Population Gains 1939-50			Population Sept. 13, 1950
	Military Losses	Civilian Losses	Non-military Deaths	Others	Expellees	Refugees from SOZ	Births	
39.3	2	0.4	5.2	0.5	7.9	1.5	7	47.6

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

Table 2: Changes of the Population in Western Germany: 1939 - 1950

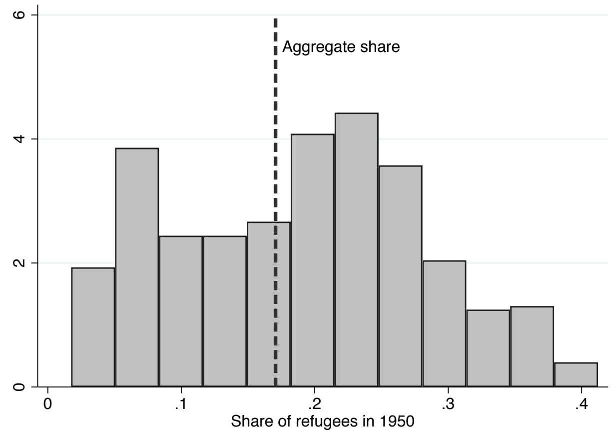
Czechoslovakia, where a substantial German minority resided. Under the backing of the respective governments, both the army and privately organized militias started to systematically expel the German population. It is only the Potsdam agreement which tried put an end to these unorganized expulsions and legalized them ex-post. The official protocol of the Potsdam conference reads:

> *"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 German population was transferred from Middle- and Eastern Europe to Western Germany and the Soviet Occupied Zone. In Figure 2 I depicts the stock of expellees in Western Germany. By 1948 almost 7m expellees were already present in Western Germany.[2] This amounted to roughly 20% of the population living in Western Germany at the time. Despite the casualties during the war, the population of Western Germany had therefore increased substantially.

This is seen in Table 2, which reports the sources of population dynamics in Western Germany between 1939 and 1950. On net, the the Western Germany population increased by 8m people between 1939 to 1950. This increase was mostly due to the 8m expellees and the 1.5m refugees from the Soviet Occupied Zone. The direct war losses can account for roughly 2.5m deaths in Western Germany.

[2] There are additional refugees from the East coming into Germany after 1950. These flows are not only much smaller in magnitude, but most of them moved to Western 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.



Notes: The figure shows the distribution of the share of refugees relative to the respective county population in Western Germany.

Figure 3: The Distribution of Refugees Across Counties

The Initial Allocation of Refugees Across Counties

A key feature of the initial inflow of refugees in the late 1940s was that it was strikingly unbalanced across space. Gerhard Reichling, a German historian who was in charge of the main statistical analysis about the integration of the refugees, concludes that "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 Figure 3, where I show the distribution of the share of refugees across counties. In the aggregate, refugees amounted to roughly 20% of the population. However, this aggregate statistic hides substantial heterogeneity across space. While some counties received hardly any expellees other counties had refugee shares exceeding 40%. What accounts for this unequal distribution of refugees across space? The historical sources overwhelmingly agree on the fact that an orderly allocation of the refugees across localities in Western Germany was impossible at the time and that the search for housing rather than employment opportunities was the predominant concern. Werner Nellner, one of the leading post-war economic historians, describes the situation as follows: "In the midst 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, even though in the majority of cases the situation was very primitive and many had to dwell in the tightness of refugee camps for years" (Nellner, 1959, p. 73).

This uncoordinated assignments of refugees was known to the military government and was 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) concludes 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 of supporting the poor." (Grosser and Schraut, 2001, p. 85).

This focus on the availability of housing supply is also directly seen in the data as refugees were naturally sent to

Notes: The figure depicts the relationship between the share of refugees and the extent of war-time destruction, the the population density in 1939 and the population-weighted distance to the expulsion region calculated according to (1). All plots show the residual variation after controlling for state fixed effects.

Figure 4: Regional housing supply and the Allocation of Refugees

places where it was easies to find shelter, i.e. places with historically low population density. In addition, data from the German housing census at the county-level show that about 23% of the aggregate housing stock was damaged during the Allied bombing campaign. Moreover, there is considerable heterogeneity in the extent to which Western German counties were affected as a large share of counties saw more than 60% of their housing stock damaged during the war (see Section 8.1 in the Appendix). In the two panels on the right I depict the correlation of the share of refugees with the population density in 1939 and the extent of war-time destruction - there is a strong negative correlation, indicating that rural location with little war-time destruction were natural places for refugees to be allocated to. The left panel shows that there is also a strong geographical element in the initial settlement. In particular, I depict the correlation of the share of refugees with the population-weighted distance to the expulsion region, which I calculate as

$$\text{exp\_dist}_c = \ln\left(\sum_{r \in ER} d_{c,r} \times pop_r^{1930}\right), \tag{1}$$

where $d_{c,j}$ is the geographical distance between county j and r and pop_r^{1930} is the size of the population in 1939. Again there is a strong negative correlation that regions, which were closer to the population centers in the pre-war period, saw more refugee inflows.

Migration and Persistence of the Initial Shock

In this paper I exploit population inflow as a shock to spatial labor supply. Hence, the consequences of this shock depend crucially on its persistence (i.e. how quickly did refugees move out) and to what extent native workers are crowded out. Regarding the first point, the initial distribution of refuges showed a lot of persistence. First of all, it is important to note that labor mobility was severely restricted in the post-war period. In 1945, the refugee committee of the Occupying Forces decided to deploy armed forces at the state boundaries to prevent internal migration. William H. Draper, Director of the Economic Division of the OMGUS, notes 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). Additionally, the incentives to migrate were also arguably low as the political support for the housing allocation was only provided in the locations refugees were initially assigned to. Similarly, there were restriction to receive food stamps without being officially registered (see e.g. Grosser and Schraut (2001, p. 83) or Nellner (1959, p. 75)). This absence of spatial

	Share of refugees in ...		Population growth ...	
	1955	1961	1939-1950	1939 - 1961
Share of refugees in 1950	0.735***	0.601***	1.569***	1.315***
	(0.028)	(0.035)	(0.117)	(0.255)
ln pop dens. 1939			-0.043***	0.007
			(0.007)	(0.007)
State FE	✓	✓	✓	✓
Observations	536	484	536	488
R^2	0.859	0.733	0.651	0.174

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively.

Table 3: Refugees: Persistence and Population Growth

mobility is in fact often alluded to in the contemporary sources. For example the economic reports of the OMGUS themselves argue that high levels of unemployment are accompanied by labor shortage because "the mobility of labor is limited. Hence there is little possibility of an early change in the distribution of labor. For example. 46% of the job openings in Bavaria in March 1947 were in the major cities of Munich, Nuremberg and Augsburg, while the majority of immigrant labor resided in rural districts" (Office of the Military Government for Germany, 1947, p. 10).

In Table 3 I provide direct evidence that the initial allocation of refugees was highly persistent and that it presented a main source of regional population growth in the post-war period. In columns 1 and 2 show that there is a strong correlation between the share of refugees in 1950, 1955 and 1961. Columns 3 and 4 document a large positive correlation between the share of refugees and subsequent population growth. Hence, the initial initial refugee shock had a sizable effect on the spatial distribution of the population across Western German counties and there was only limited crowding-out of the native population.

3 Economic Consequences in the 50s and 60s

I now turn to the economic consequences of this historical experiment, i.e. how did the local economy respond to this inflow of people. My main analysis focuses on the effect on regional industrialization and income per capita. In Section 3.2 I use the historical context that the initial allocation was based on the availability of housing for an analysis based on OLS. In Section 3.3 I report a complementary instrumental variable strategy. In Section 3.4 I use micro-data to provide direct evidence that refugees' labor supply was indeed sharply directed to the manufacturing sector.

3.1 Data

In this paper I use a variety of datasets. The majority of the analysis exploits the spatial variation across 500 counties in Western Germany and links refugee inflows in the late 1940s to economic outcomes in the 1950s and 1960s. To perform this analysis, I constructed a panel dataset at the county-level spanning the time-period from 1933 to 1970. The dataset was constructed by digitizing a host of historical publications. In contrast to many other countries there are, to the best of my knowledge, no records of the historical micro census data with sufficient regional breath to calculate outcomes at the level of the roughly 500 Western German counties. However, the local statistical offices did publish summary statistics of the respective census at the county-level at the time, which I could digitize and harmonize to a consistent and time-invariant border definition.

The basis of the dataset is comprised of the population censuses for the years 1933, 1939, 1950, 1961 and 1970, which are published individually for each of the 9 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 at the county-level. I then augmented this dataset with five additional pieces of information. The first concerns the regional allocation of refugees, which I digitized from a special statistical publication published in 1953. Secondly, in the 60s and 70s, the different statistical offices from the respective German states instituted a commission to construct measures of GDP at the county-level. These results were published and could be digitized. Third, to obtain a county-level measure of GDP in the pre-war and intermediate post-war period, I exploit information from tax record, which report measures on value added taxes at the county-level. Because tax rates do not vary across space in Germany, I take these measures as being proportional to local GDP. Fourth, I also digitized the county-level results for three waves of the manufacturing census in 1933, 1939 and 1956. The manufacturing census reports the number of plants by industry at the county-level. This allows me to measure the entry of manufacturing plants at the regional level. Finally, I also provide new measures of the extent of war time destruction and regional housing supply, which I digitized from the historical housing census conducted in 1950. This census contains information on the extent of war damages for each county and detailed information on living conditions of refugees and natives. I want to stress that this data is different from the one used in Brakman et al. (2004) and Burchardi and Hassan (2013). These papers mostly focus on the extent of war-time destruction in 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.

Finally, I also use microdata to shed light on the specific mechanism of reallocation after the initial expulsion. The most important dataset is the *Mikrozensus Zusatzerhebung 1971 (MZU 71),* a special appendix to the census conducted in 1971. The purpose of this appendix was to study the "social and economic mobility of the German population" and fortunately it includes identifiers about individuals' refugee status. Most importantly, the data contains retrospective information about employment characteristics in 1939, 1950, 1960 and 1971 at the individual level. This allows me to observe the whole employment history of individuals for a 40 year window, which includes both pre- and post-expulsion outcomes. The MZU 71 has roughly 200.000 observations, 40.000 of which are refugees. The MZU 71 data does not contain information about historical wages nor does it contain regional identifiers at the county level. To provide some information on earnings, I use additional micro data that contains information on both wages and the refugee status of respondents. The *Einkommens-und Verbrauchsstichprobe 1962/63 (EVS 62)* is a micro dataset conducted in 1962 to measure household income and expenditure and is hence similar to the Consumer Expenditure Survey in the US. The 1962/63 wave of the survey has about 32.000 observations.

3.2 Refugees, Industrialization and Spatial Growth

In this section I relate the inflow of refugees to the development of local manufacturing sector and to the rate of spatial productivity growth both in the short- and the long-run. These cross-sectional estimates are also the backbone of my structural analysis because I will calibrate the structural parameters to these regression coefficients using indirect inference.

There are two theoretical reasons for why I expect the inflow of refugees to be associated with the expansion of the manufacturing. First, if the agricultural sector is subject to decreasing returns, an expansion of the size of the local workforce will be associated with an increase in the manufacturing employment share. Second, if the inflowing refugees had a comparative advantage in the manufacturing sector within the local labor market, they represent a shift in labor supply, which is biased towards the manufacturing sector. The effect of these population inflows

	Manufacturing				Agriculture	Services	Manufacturing	
	1939	1950			1950		1961	
Share of refugees in 1950	-0.077	0.235^{***}	0.302^{***}	0.315^{***}	-0.264^{***}	-0.042	0.248^{***}	0.249^{***}
	(0.089)	(0.083)	(0.045)	(0.047)	(0.072)	(0.069)	(0.057)	(0.058)
Manufacturing share in 1939			0.865^{***}	0.875^{***}	-0.534^{***}	-0.341^{***}	0.884^{***}	0.885^{***}
			(0.035)	(0.036)	(0.044)	(0.040)	(0.038)	(0.040)
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓	✓
Geography				✓	✓	✓		✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	535	536	535	535	535	535	535	535
R^2	0.461	0.486	0.895	0.896	0.863	0.702	0.794	0.794

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The regression is at the county level. The dependent variable is the manufacturing employment share in 1939 (column 1), in 1950 (columns 2 - 4), the agricultural employment share in 1950 (column 5), the service employment share in 1950 (column 6) and the manufacturing employment share in 1961 (columns 7- 8). All specifications control for the log of population density in 1939, the extent of wartime destruction through the share of the housing stock, which was damaged during the war, and state fixed effects. Specifications with geography controls also control for the log of the distance to the inner german border and a fixed effect for whether a county is a border county.

Table 4: Refugees and Manufacturing Employment in 1950

on GDP pc is in principle ambiguous. If regional economies are subject to decreasing returns to scale, we would expect that the inflow of refugees to reduce income per capita. If on the other hand population inflows increase spatial productivity through agglomeration economies, we would expect income per capita to increase, especially in the long-run.

To estimate the relationship between refugee inflows and manufacturing employment, I consider a specification of

$$\pi_{rt}^{M} = \delta_s + \beta\mu_{r1950} + \alpha\pi_{r1939}^{M} + \phi \ln l_{r1939} + \varphi \text{wd}_r + x_r'\zeta + u_r, \quad (2)$$

where π_{rt}^{M} denotes the manufacturing employment share in time t and μ_{r1950} is the share of refugees in 1950. Furthermore I control for a set of state fixed effects (δ_s), the population density in 1939 (l_{r1939}) and the extent of wartime destruction (wd_r), which are the main determinants of the housing supply (and hence refugee flows) displayed in Figure 4, and a set of additional spatial controls (x_r). Because I explicitly control for π_{r1939}^{M}, the coefficient of interest β captures the effect of refugees on the growth of manufacturing employment. In Table 4 I report the results when I estimate this specification using OLS. This strategy is valid if the residual variation in refugee flows is uncorrelated with other local characteristics that affect the growth in the local manufacturing share. In Section 3.3 below I consider an alternative strategy based on an instrumental variable strategy. In column 1 I first consider a "Placebo"-like specification and show that there is no relationship between the share of refugees in 1950 and the manufacturing share in 1939. Columns 2 runs the exact same specification using the 1950 manufacturing employment share as the dependent variable. Now there is a sizable positive effect: an increase in the share of refugees by 10 percentage points increases the manufacturing employment share by almost 2.5 percentage points. In columns 3 and 4 I control explicitly for the share of manufacturing employment in 1939 and for additional geographical determinants, in particular a fixed effect for border counties and the distance to the inner german border. This leaves the coefficient qualitatively unchanged but increases the precision substantially. In columns 5 and 6 I show that the sectoral reallocation takes place exclusively between the manufacturing and the agricultural sector. In contrast, the employment share of the regional service sector is unrelated to the inflow of refugees. This is exactly what one would expect if services are non-traded and agricultural production is subject to decreasing returns. Finally, the last two columns repeat the same regression for the manufacturing employment share in 1961

	ln GDP pc							
	1950		1957-1966		1957	1961	1964	1966
Share of refugees in 1950	0.098	0.178	0.495***	0.617***	0.505**	0.643***	0.679***	0.642***
	(0.321)	(0.322)	(0.101)	(0.100)	(0.200)	(0.185)	(0.215)	(0.198)
$\ln y_{r1935}$	0.362***	0.365***	0.115***	0.102***	0.110***	0.099***	0.095***	0.104***
	(0.056)	(0.059)	(0.009)	(0.010)	(0.017)	(0.016)	(0.021)	(0.020)
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓	✓
Prewar Industrial Structure		✓		✓	✓	✓	✓	✓
Year FE			✓	✓				
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	524	523	2080	2076	519	519	519	519
R^2	0.699	0.702	0.727	0.750	0.656	0.570	0.432	0.423

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The dependent variable is log GDP pc in 1950 columns 1,2), log GDP pc in 1957, 61, 64 and 66 in columns 5-8. In columns 3 and 4 I pool the data for the years 57-66 and estimate the regressions at the county-year level and include a full set of year fixed effects. All specifications control for the log of population density in 1939, the extent of wartime destruction through the share of the housing stock, which was damaged during the war, state fixed effects, the log of the distance to the inner german border and a fixed effect for whether a county is a border county. Specifications 2, 4 and 5-8 also also control for the pre-war industrial structure via the manufacturing employment share in 1933 and 1939.

Table 5: Refugees and GDP per capita

and show that the inflow of refugees is still strongly related to manufacturing specialization in 1961. Together with the result shown in Table 3 that the inflow of refugees had persistent effects on spatial population growth this is again consistent with the presence of decreasing returns in the agricultural sector and and an increase in manufacturing productivity.

In Section 8.2 in the Appendix I report a battery of robustness checks for these results. In particular, I include additional controls for the sectoral composition of employment in 1933, focus on particular states and estimate (2) using refugee share in 1946 (instead of 1950), which is observed for a subset of counties. I also include additional controls for post-war changes in the housing stock (to control for direct construction needs) and for cross-county differences in demographics and employment-population ratios to control for differences in labor supply. The results reported in 3 are generally robust to these considerations.

I now estimate the effect of the initial refugee settlement on regional income per capita. As discussed in Section 3.1, I observe GDP per worker for the years 1935, 1950, 1957, 1961, 1964, and 1966.[3] My main empirical specification takes a similar form as the specification in (2):

$$\ln y_{rt} = \delta_s + \beta\mu_{r1950} + \alpha \ln y_{r1935} + \phi \ln l_{r1939} + \varphi \text{wd}_r + x_r'\zeta + u_r, \tag{3}$$

where y_{ct} denotes GDP per capita in region r at year t and $\ln y_{r1935}$ denotes the log of GDP per capita in 1935. All remaining variables are defined as above. The results are contained in Table 5.

The first two columns show that income per capita growth between 1950 and 1935 is essentially unrelated to the inflow of refugees. While the coefficient is positive, the standard errors are large. Hence in the short-run, the countervailing forces between decreasing returns and agglomeration economies seem to cancel each other out. Hence, if agricultural production is subject to decreasing returns (as is suggested by the results in Table 2), this implies that some other sector is subject to some form of aggregate increasing returns.[4] The remaining columns of Table 5 show a large positive effect between the inflow of refugees and income per capita in the long run. In columns

[3]Recall that I only observe value added taxes per worker in 1935 and 1950. For simplicity I also refer to these measures as "GDP per worker". This is correct if value added taxes are proportional to regional income.

[4]I also want to note that the inflow of refugees was unlikely to increase the regional supply of human capital. Below I present evidence from microdata that refugees earned less than natives.

3 and 4 I pool the data between 1957 and 1966 and estimate (3) while controlling for a full set of year fixed effects. According to these estimates, an increase in the share of refugees by 10% increases income per capita by roughly 5% after 15 years. In the remaining columns I estimate the long-run effect of refugees on income per capita separately for each year. The results are quantitatively consistent with the pooled regression. Interestingly the point estimates are increasing. However, these differences are too small to detect them statistically. In summary, Table 5 suggests that the population inflow has a positive effect on local productivity, which however needs a decade to materialize. In Sections 4 and 5 I show that these results are quantitatively consistent with a parsimonious model of spatial growth. In the theory, the endogenous productivity response is rationalized through the entry of manufacturing plants and there I also present direct evidence for this mechanism.

Section 8.2 in the Appendix shows that these results a robust to the inclusion of a host of other controls, e.g. fixed effects for being a city-state or additional controls for the industrial structure pre-war. All specification show that the effect of refugees on GDP per capita is insignificant in the short-run and positive in the long-run.

3.3 Instrumental Variable Estimates

For the structural estimation of the theory below, I use the OLS estimates reported in Tables 4 and 3 as identified moments and estimate the structural parameters of the theory via indirect inference by targeting the regression coefficients. As complementary evidence that the results reported above reflect to a large extent the causal effect of changes in local market size on regional industrialization and local productivity growth, this section presents an instrumental variable strategy to estimate specifications (2) and (3).

The identification strategy is based on the striking geographical variation presented in Figure 3: even within states there is a strong negative relationship between the share of refugees and the population-weighted distance to the expulsion regions. This expulsion distance can therefore be used as an instrument for the share of refugees if it only affects the growth in manufacturing employment or the growth in income per capita via the inflow of refugees. 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. Hence, if regions closer to the border are directly affected by the German division through for example political uncertainty or - as argued by Redding and Sturm (2008) - through a larger loss in market access, the identification assumption would be violated.

I address these concerns in three ways. First of all, 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. Secondly, I note that both of these arguments would imply a negative correlation between the instrument and regional income growth or the growth of the manufacturing sector, which produces tradable goods. Hence, such concerns would induce a negative bias for the coefficients on the share of refugees. Third, in the Appendix I also offer an additional instrumental variable strategy, which is less subject to these concerns. The results are contained in Table 6. The first column contains the first stage relationship and confirms that - as suggested by Figure 3 - that there is still a strong negative relationship between the expulsion distance and the share of refugees even if the controls, which are included in the second stage, are controlled for. The remaining columns then contain the results for manufacturing employment and income per capita. Columns 2 and 3 show the results for manufacturing employment and correspond to columns 4 and 8 in the OLS specification of Table 4. The IV estimates are very similar to the OLS estimates. Columns 5 and 6 focus on income per capita (the corresponding OLS specifications are shown in columns 2 and 4 of Table 5). The results are again quite comparable to the OLS results. The IV estimate for 1961 is a tad lower than the OLS estimate but in the same ballpark. Importantly, the IV estimates also show the qualitative pattern of a small positive but insignificant effect in 1950 and a positive effect in the long-run.

	First stage	Man. empl. share		ln GDP pc	
	Share of refugees in 1950	1950	1961	1950	1957 - 1966
Dist. to Expulsion Population	-0.118***				
	(0.033)				
Share of refugees in 1950		0.286***	0.258**	0.162	0.379**
		(0.078)	(0.104)	(0.685)	(0.168)
$\ln y\_1935$				0.365***	0.104***
				(0.059)	(0.010)
ln pop dens 1939	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓
Prewar Industrial Structure	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓
Observations	523	523	523	523	2076
R^2	0.784	0.909	0.803	0.702	0.749

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: IV Estimates: Distance to the expulsion population

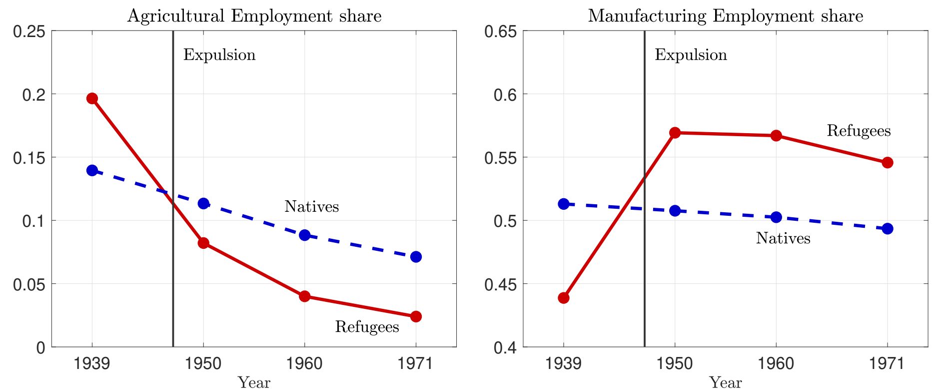
Given that if anything I would expect the OLS to be upward biased and the IV estimate to be downward biased, it is reassuring that the results are quite similar and hence informative about the economic effect of the refugee inflow on regional income per capita and industrialization. I nevertheless discuss an alternative instrumental variable strategy in the Appendix. This strategy exploits the historical fact that the inflowing refugees were often housed within the apartments of natives whenever housing was particularly scarce. Because such an allocation was easier if native homes were multi-room apartments, this suggests that the interaction between the expulsion distance and the supply of multi-room houses should predict the allocation of refugees as this margin of housing supply was only tapped into when other options to house refugees were exhausted (which is more likely to occur in regions close to the expulsion regions as the stream of refugees was higher). In the Appendix I report the results from this exercise and show that the results are qualitatively similar but less precisely estimated. The estimated relationship with the manufacturing share (both in 1950 and 1061) is somewhat larger and the effect on GDP per capita in 1961 is very similar to the one reported above. The relationship with GDP per capita in 1950 is very noisy with a large standard error is sometimes negative and statistically significant in some specifications.

3.4 Refugees and Manufacturing Labor Supply

The above analysis showed a positive relationship between refugee-inflows and manufacturing employment. In this section I provide some direct micro-evidence on that refugees were indeed “pushed” into the manufacturing sector and plausibly acted as a local supply shifter.

Recall that refugees came from areas, which - historically - tended to specialize in agriculture. Also recall that refugees were predominantly allocated to rural areas, which - in 1939 - had a low employment share in manufacturing.[5] Upon their arrival in Western Germany, however, refugees did not enter the agricultural sector but predominantly worked in the manufacturing sector. This reallocation pattern is clearly visible in a unique special supplement to the 1971 census, which contains micro-data on long-run employment histories. In particular, that census asked every respondent in 1971 where he/she lived in 1939 and in which occupation/sector cell he/she worked in 1939, 1950, 1960 and 1971. By analyzing this time-series of retrospective questions, I can therefore measure the life-cycle of employment patterns for both refugees and natives. Importantly, the data spans the time

[5]Recall that the result in column 1 of Table 4 that the share of refugees is uncorrelated with the manufacturing employment share in 1939 is a conditional statement. The unconditional correlation between the share of refugees and the manufacturing employment share in 1939 is strongly negative.



Notes: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. The experience of refugees (natives) is depicted in solid (dashed) lines. The expulsion, taking place in 1947, is drawn as the red, vertical line. The data stems from the MZU 71 (Mikrozensus Zusatzerhebung 1971), i.e. the data for the years prior to 1971 is based on individuals' responses in 1971.

Figure 5: The Life-Cycle of the 1915-1919 Cohort

of the expulsion in the mid 1940s. Hence, I can exactly see how refugees' employment patterns change *relative* to natives between 1939 and 1950, i.e. pre- and post expulsion.

In Figure 5 I depict the sectoral life-cycle profile for the cohort of workers born between 1915 and 1919. Hence, 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 in the figure show the agricultural employment share (left panel) and the manufacturing employment share (right panel). The vertical line indicates the time of the expulsion.

Figure 5 vividly displays the process of biased reallocation. Among refugees, 20% used to work in the agricultural sector in 1939.[6] After the expulsion and their resettlement to Western Germany, only 8% still did so. In contrast, the share of manufacturing employment within the same cohort of individual increases from 44% to almost 60%. This is very different for the cohort of natives. For them, the time period of the expulsion is hardly noticeable and the slight secular decline in agricultural and manufacturing employment reflect the process of structural change towards the service sector.

One interpretation of the patterns in Figure 5 is that it is a consequence of spatial sorting. Even though refugees were on average specialized in agriculture, they might have have a comparative advantage in the manufacturing sector within the actual rural labor market they were sent to. If native workers were spatially sorted the average native working in a rural area is likely to have more agricultural skills than an average refugee. Intuitively: the only chance for a rural location to ever get a handful of engineers might be as part of the refugee trek.

Another interpretation of these results is that they are the result of frictions in the agricultural labor market, i.e. refugees might have had difficulties to find agricultural work despite their agricultural human capital. In 1950 Germany, the majority of agricultural employment was still very much concentrated in small, family-run farms. According to the agricultural census, which reports the average farm size for each county in Germany, the average farm size is on the order of magnitude of 10-15 hectares.[7] Hence, the demand for outside agricultural workers was quite limited. Even the Military Government of the US pointed out in 1947 that of the immigrants "well over half

[6]Note that this number is substantially smaller 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.

[7]As a point of comparison: the average farm in the US today is about 180 hectares large, i.e. ten times that size. And even in 1900, US farms already had a size of 60 hectares.

a million, were farmers. But agricultural acreage [...] cannot be expanded significantly. Within the US Zone the possibility of increasing settlement by changing the size and structure of farms is very small". In Section 8.1 in the Appendix I discuss this in more detail and show direct evidence that refugees were indeed much less likely to be independent farmers. For the empirical results above, it is less important to distinguish why refugees might have acted as a supply shifter of manufacturing human capital. In the structural analysis below I assume that it is due to spatial sorting.

4 Theory: A Semi-Endogenous Model of Spatial Growth

The evidence presented above showed that the inflow of refugees was associated with regional industrialization and with productivity gains in the long-run. I also showed that initial population inflow was highly persistent and increased the regional population in the long-run. In this section I develop a theory that can speak to this evidence and show that it is quantitatively consistent with this evidence. The model combines elements from the literature on growth theory, particular the distinction between models of endogenous and semi-endogenous growth of Jones (1995), and from the recent literature on model of economic geography as summarized in Redding and Rossi-Hansberg (2017). In Section 5 I estimate the structural parameters of the theory by targeting the empirical regression results as "identified moments" (Nakamura and Steinsson, 2018).

4.1 Environment

I consider an economy with R regions ("counties"). Consumers face a consumption choice, i.e. how to allocate their expenditure across different goods, 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 their per-period utility. They derive utility from consuming both agricultural and manufacturing goods according to a Cobb-Douglas utility function

$$u\left(c_A, c_M\right) = c_A^{\alpha} c_M^{1-\alpha}. \tag{4}$$

Both goods $j = A, M$ are in turn CES aggregates from a set of differentiated, regional varieties. In particular, there are R regions, denoted by r, which produce a differentiated variety Y_{jrt}, which are tradable across space and aggregated according to

$$Y_{jt} = \left(\sum_{r=1}^{R} Y_{jrt}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}. \tag{5}$$

Here $\sigma > 1$ denotes the elasticity of substitution across varieties, which for simplicity I assume to be the same between agricultural and manufacturing goods. For expositional simplicity I abstract from the presence of trade costs. This simplifies some expressions. For the quantitative application I assume that the trade costs take the usual iceberg form.

Production The agricultural good is produced using labor, measured in efficiency units, and land, i.e.

$$Y_{Ar} = Q_{rt} T_r^{1-\gamma} H_{Art}^{\gamma},$$

where T_r denotes the fixed total supply of land in region r, H_{Art} denotes the total amount of labor employed for agricultural production and Q_{rt} is total productivity in region r at time t. Evidently, agricultural production is subject to decreasing returns to scale.

The production of the manufacturing good, in contrast, is subject to variety gains as in Romer (1990). In particular, the tradable manufacturing goods is produced using production labor and intermediate inputs according to

$$Y_{Mrt} \quad = \quad Q_{rt} X_{rt}^{1-\beta} H_{Prt}^{\beta}. \tag{6}$$

The intermediate input X_{rt} is in turn composed of a CES bundle of a continuum of differentiated input varieties

$$X_{rt} \quad = \quad \left(\int_{i=0}^{N_{rt}} x_{it}^{\frac{\rho-1}{\rho}} di \right)^{\frac{\rho}{\rho-1}}, \tag{7}$$

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

The regional productivity term Q_{rt}, which for parsimony I assume to be common across the two sectors, evolves according to

$$\ln Q_{rt} = (1 - \varrho) \ln Q_r + \varrho \ln Q_{rt-1} + \varpi u_{rt}, \tag{8}$$

where Q_r is a fixed, region-specific level of innate productivity, ϱ governs the regional auto-correlation of the stochastic productivity process and u_{rt} is spatial productivity shock, which is distributed iid. Hence, the county fixed effect Q_r determines the long-run value of productivity for region r and ϖ governs the variance of regional productivity shocks.

Entry The number of input varieties N_{rt} is determined endogenously and provides the link between local productivity growth and local labor supply. I assume that at the end of each period (after production has taken place) an exogenous fraction δ of firms exits. Firm entry takes place in the beginning of the period. The labor requirement to start a new firm, h_{rt}^E, is given by

$$h_{rt}^E = f_E N_{rt-1}^{-\lambda}. \tag{9}$$

The parameter f_E parametrizes the size of entry costs and λ governs the extent of spillovers as in Jones (1995). As I will show below, λ is the crucial parameter to determine the long-distribution of economic activity across space and whether population shocks have persistent or transitory effect. I will refer to λ as the *inter-temporal knowledge elasticity* or simply the *knowledge elasticity*. This terminology is motivated by the fact that λ determines how the existing state of knowledge affects the costs of creating new knowledge.

If $\lambda = 0$ and $\delta = 1$, the model is essentially the static model of Krugman (1980a): firms only live for a single period and the cost of entry are linear in labor and do not depend on the number of varieties, which are already available. The case of $\lambda = 1$ and $\delta = 0$ is the polar opposite case of Romer (1990), where ideas do not depreciate and the inter-temporal knowledge elasticity is large enough to lead to endogenous growth. The intermediate case of $0 < \lambda < 1$ is the semi-endogenous growth model of Jones (1995). I show below that the endogenous growth case of $\lambda = 1$ and the semi-endogenous case of $\lambda < 1$ have strikingly different implications for the long-distribution of economic activity across space and the persistence of the consequences of refugee inflows.

Sectoral Labor Supply I model the sectoral supply of human capital using the usual Roy-type machinery common in models of economic geography (Redding and Rossi-Hansberg, 2017). 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 at time t. For tractability I assume that z_{ijt} is drawn from a Fréchet distribution, i.e. $F_j(z) = e^{-\phi_j z^{-\theta}}$, where ϕ_j parametrizes the average number of efficiency units individuals can provide in sector j and θ governs the dispersion of talent draws and hence the elasticity of sectoral labor supply.

To meaningfully talk about spatial sorting, I allow for persistent differences in the supply of skills, i.e. in the parameters ϕ_j. In particular, I assume that there exist two latent types, who have a comparative advantage in the respective sectors. A share χ of individuals are industrial workers (I) and a share $1-\chi$ are "farmers" (F) and I let ϕ_j^I and ϕ_j^F parametrize the respective distribution of skills in sector j.

Standard arguments (see Section 7.3 in the Appendix), show that the share of individuals of type v working in sector j in region r is given by

$$\pi_{rjt}^v = \frac{\phi_j^v w_{rjt}^\theta}{\phi_A^v w_{rAt}^\theta + \phi_M^v w_{rMt}^\theta} = \phi_j^v \left(\frac{w_{rjt}}{\overline{w}_{rt}^\nu}\right)^\theta,$$

where the average earnings of type ν in region r at time t, $\overline{w}_{rt}^\nu$, is given by

$$\overline{w}_{rt}^\nu = \left(\phi_A^v w_{rAt}^\theta + \phi_M^v w_{rMt}^\theta\right)^{1/\theta}. \tag{10}$$

Hence, if industrialist workers have a comparative advantage in the manufacturing sector (i.e. $\phi_M^I/\phi_A^I > \phi_M^F/\phi_A^F$), they have a relative higher employment share in manufacturing within a labor market and their average earnings $\overline{w}_{rt}^I$ put a higher weight on the manufacturing wage w_{rMt}. Finally, if region r is inhabited by L_{rt}^v workers of type v, the total supply of human capital in sector j is given by

$$H_{rjt} = \Gamma_\theta L_{rt}\left(\omega_{rt}^I \phi_j^I \left(\frac{w_{rjt}}{\overline{w}_{rt}^v}\right)^{\theta-1} + \omega_{rt}^F \phi_j^F \left(\frac{w_{rjt}}{\overline{w}_{rt}^F}\right)^{\theta-1}\right) \tag{11}$$

where $\Gamma_\theta = \Gamma\left(1-\theta^{-1}\right)$ is the gamma function, $L_{rt} = L_{rt}^I + L_{rt}^F$ is the total population in county r and $\omega_{rt}^v = L_{rt}^v / L_{rt}$ is the share of v-types in the population. Equation (11) shows that the total sectoral labor supply depends on relative factor prices, the total population size and the degree of spatial sporting encapsulated in ω_{rt}^v.

To connect this model of labor supply to the historical evidence, I assume that the population consists of natives and refugees and I denote the number of natives and refugees at time t by L_t^N and L_t^R. For parsimony, I assume that refugees and natives are exactly identical in terms of their skills and I denote the number of natives (refugees) of type $v = I, F$ in region r at time t by $L_{rt}^{N\nu}$ ($L_{rt}^{R\nu}$). Note that even though the *aggregate* share of industrial and rural workers among natives and refugees is identical, the spatial distribution is not. While native workers had a long time to sort spatially according to their skills (see below), the empirical evidence presented above showed that refugees are allocated mostly to rural areas and that the initial placement is uncorrelated with their latent skill. Hence, the historical refugee shock is not only a shock to population size but has a sector-specific component: by allocating a random sample of the population to rural areas, the treated areas received population inflows whose share of industrial workers exceeded the population average in rural communities.

Spatial Mobility Individuals are partially mobile across space. While individuals know their type $v = I, F$, they do not observe their skill realization z_{it}. The value for individual i living in region r at time t is hence given by

$$\mathcal{U}_{rt}^i = V_r \overline{u}_{rt}^i \xi_{rt}^i,$$

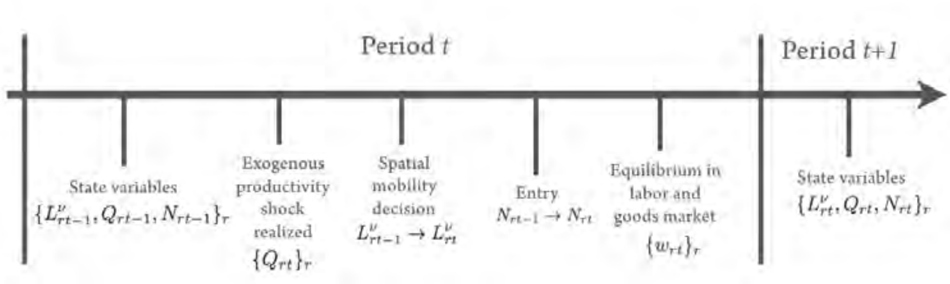


Figure 6: Timing of events

where V_r is a constant region-specific amenity, $\overline{u}^i_{rt}$ is the expected utility individual i achieves in region r at time t and ξ^i_{rt} is a regional taste shock which is independent both across individuals and across locations for any given individual and distributed according to a Fréchet distribution with shape parameter ε, i.e. $F(\xi) = e^{-\xi^{-\varepsilon}}$. Given the utility function in (4), expected utility of individual i in region r is given by

$$\overline{u}^i_{rt} \propto \frac{\overline{w}^{\nu}_{rt}}{P^{\alpha}_{At} P^{1-\alpha}_{Mt}}.$$

I make two simplifying assumptions in regards of individuals' moving choices. First of all, as highlighted above, I assume that individuals maximize static utility and hence abstract from forward looking mobility decisions. This simplifies agents' mobility choices considerably and allows me to treat the spatial population distribution as a dynamic endogenous state variable. Secondly, I model agents' mobility frictions in a reduced form way a la Calvo: at each point in time individuals have the option to move with probability ψ. Under these assumption, the evolution of the spatial distribution of population takes the simple form

$$L^{\nu}_{rt} = (1-\psi) L^{\nu}_{r,t-1} + \psi L^{v}_{t-1} m^{v}_{rt}, \tag{12}$$

where L^v_{rt} denotes the number of people of type v in region r at time t, $L^v_{t-1} = \sum_j L^v_{jt-1} = \sum_j \left(L^{Nv}_{jt-1} + L^{Rv}_{jt-1}\right)$ is the total mass of v-types in the population and m^v_{rt} is the probability for v-types of moving to location r in period t[8]

$$m^{v}_{rt} = \frac{\left(V_r \overline{w}^{\nu}_{rt}\right)^{\varepsilon}}{\sum_k \left(V_k \overline{w}^{\nu}_{rt}\right)^{\varepsilon}}.$$

4.2 Equilibrium

The framework above is a simple environment where both the spatial distribution of productivity and the population is both persistent and endogenously determined. The timing of events, which is displayed in Figure 6, is as follows. In the beginning of the period, the set of state variables in region r is given by $\left(Q_{rt-1}, N_{rt-1}, L^v_{rt-1}\right)$. Then the exogenous productivity shock Q_{rt} is realized according to (8). Given $\{Q_{rt}, N_{rt-1}\}_r$, individuals make their mobility decision, potential entrants decide on their entry activities and production and consumption takes place. Individuals' mobility decisions and firms' entry decisions determine $\{L^v_{rt}, N_{rt}\}_r$, i.e. the future set of state variables.

[8]The fact that m^v_{rt} does not depend on the prices of goods is a consequence of the absence of trade costs. In the quantitative model, which allows for trade costs, the real wage in location r depends on location r's position in the trading network.

Static Allocations Consider first the manufacturing sector. Because the market for intermediate inputs is monopolistically competitive, the N_{rt} suppliers of differentiated input varieties charge a constant markup $\frac{\rho}{\rho-1}$. Their profits are hence a share $1/\rho$ of firm revenue. The Cobb Douglas structure of production in (6) implies that a share $1-\beta$ of total manufacturing revenue $P_{rMt}Y_{rMt}$ is allocated to the purchases of aggregate intermediate inputs X_t and hence among the N_{rt} firms. This implies that profits of firm i in region r at time t are given by

$$\pi_r = \frac{1}{\rho}\frac{(1-\beta)\,P_{rM}Y_{rM}}{N_r}. \tag{13}$$

Similarly, workers in the production of the tradable manufacturing good, H_{rP}, receive a share β of total revenue, i.e.

$$w_{rM}H_{rP} = \beta P_{rM}Y_{rM}. \tag{14}$$

One can also show that total labor demand of intermediate firms is given by

$$H_{rX} = \int_{i=0}^{N_{rt}} x_i di = \frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right)H_{rP}, \tag{15}$$

i.e. independently of local characteristics or equilibrium prices, total intermediate employment in region r, H_{rX}, is always proportional to total production employment H_{rP}. Finally, this also implies that aggregate manufacturing output Y_{rMt} and the price of the manufacturing composite P_{rMt} are given by

$$Y_{rMt} = \varsigma_1 Q_{rt} N_{rt}^{\frac{1-\beta}{\rho-1}} H_{rPt} \qquad \text{and} \qquad P_{rMt} = \frac{1}{\beta}\frac{1}{\varsigma_1}\frac{1}{Q_{rt}} w_{rt} N_{rt}^{-\frac{1-\beta}{\rho-1}}, \tag{16}$$

where $\varsigma_1 = \left(\frac{\rho-1}{\rho}\frac{1-\beta}{\beta}\right)^{1-\beta}$. Equation (16) shows the usual variety gains: a larger mass of varieties N_{rt} increases productivity and these productivity gains are higher the lower ρ.

In the agricultural sector, cost-minimization implies that the agricultural price is given by

$$P_{rAt} = \frac{1}{Q_{rt}}\frac{w_{Art}}{\gamma}\left(\frac{H_{rAt}}{T_r}\right)^{1-\gamma}.$$

Hence, holding the agricultural wage constant, prices are rising in H_{rA}/T_r because of decreasing returns.

Entry, Market Size and Spatial Growth As for workers, I also assume that entering firms act myopically only consider static profits as part of their entry decision.[9] Free entry therefore requires that[10]

$$\pi_r = w_{rMt}h_{rt}^{E} = w_{rMt} f_E N_{rt-1}^{-\lambda}. \tag{17}$$

Using (13) and (14), this condition implies a simple expression for the rate of spatial productivity growth

$$\frac{N_{rt}}{N_{rt-1}} = \frac{1-\beta}{\rho\beta}\frac{1}{f_E}\underbrace{H_{rPt}}_{\text{Market size}} \times \underbrace{N_{rt-1}^{\lambda-1}}_{\text{Spatial catch-up growth}}. \tag{18}$$

[9]See Walsh (2019) for a related model where entrepreneurs are fully forward looking.

[10]The free entry condition only holds if new firms are actually created, i.e. if $N_{rt} > (1-\delta)\,N_{t-1}$. While this condition will always be satisfied in the steady-state, it might not hold during the transitional dynamics. In that case, the free entry condition in (17) holds with inequality, i.e. $\pi_{irt} < w_{rMt} f_E\left((1-\delta)\,N_{rt-1}\right)^{-\lambda}$. As I show in Section 7 in the Appendix, this condition can be written as $H_{rMt} < \frac{\rho-(1-\beta)}{1-\beta} f_E\left((1-\delta)\,N_{rt-1}\right)^{1-\lambda}$. To avoid a taxonomic presentation of the results, I focus on the case where (17) holds with equality in the main text. In the quantitative application I of course allow for the general case where (17) might be slack.

Equation (18) is the key equation of the model as it highlights the two determinants of the rate of variety and hence productivity growth at the local level. The first term is the usual market size effect present in most models of growth: a larger workforce H_{rPt} increases innovation incentives, because it goes hand in hand with larger profits. Note that H_{rPt} emerges as a sufficient statistic which summarizes all equilibrium effect on sectoral wages and aggregate demand, which are determined as part of the trade equilibrium. The second term captures the extent of "fishing out". As in Jones (1995), if $\lambda < 1$ there is a negative relationship between the rate of variety growth and the number of variety N_{rt-1}. As I will discuss in detail below, this has important implications for both the time-series and the spatial dimension. First of all, this congestion force makes the model a model of semi-endogenous growth, where the growth rate is not a function of the size of the workforce but only determined from the rate of population growth. Secondly, the very same congestion term also has strong implication for the long-run distribution of economic activity across locations on a BGP and for whether population shocks are permanent or transitory.

Static Agglomeration Even though (16) shows that regional manufacturing output is a linear function of the number of production workers, the number of active firms N_{rt} is itself a function of market size. Combining (16) and (18) therefore yields the equilibrium production function

$$Y_{rMt} = \varsigma_1 Q_{rt} N_{rt}^{\vartheta} H_{rPt} = \varsigma_2 Q_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{1+\vartheta}, \tag{19}$$

where

$$\vartheta = \frac{1-\beta}{\rho - 1}, \tag{20}$$

and $\varsigma_2 = \left(\frac{\rho-1}{\rho}\frac{1-\beta}{\beta}\right)^{1-\beta}\left(\frac{1-\beta}{\rho\beta}\frac{1}{f_E}\right)^{\vartheta}$ is an inconsequential constant. Equation (19) shows that - holding a location's state variable (Q_{rt}, N_{rt-1}) fixed - the aggregate production function in the manufacturing sector has increasing returns to scale as productivity is increasing in the size of the local workforce if $\vartheta > 0$. In particular, the production structure is isomorphic to a production function with a static agglomeration force in many models of economic geography (Redding and Rossi-Hansberg, 2017). I therefore refer to ϑ as the "short-run" scale elasticity.

Dynamic Equilibrium Given the environment above, I now characterize the dynamic equilibrium of this economy. The dynamics are induced because the number of firms N_{rt} and the population distribution L_{rt} are endogenous dynamic state variables which evolve jointly. Such an equilibrium is defined in the usual way.

Definition 1. *An dynamic equilibrium is a path of wages and land rents* $\{w_{rAt}, w_{rMt}, R_{rt}\}_{r,t}$, *intermediate input prices and quantities* $\{[p_{irt}, x_{irt}]_i\}_{rt}$, *local populations for both types for refugees and natives* $\left\{L_{rt}^{Nv}, L_{rt}^{Rv}\right\}_{vrt}$, *employment allocations* $\{H_{rAt}, H_{rPt}, H_{rXt}, H_{rEt}\}_{rt}$, *regional manufacturing firms* $\{N_{rt}\}_{rt}$ *and quantities of tradable goods* $\{Y_{rAt}, Y_{rMt}\}_{rt}$ *such that*

1. *Firms and consumers behave optimally,*
2. *The labor market in all locations clears at each point in time,*
3. *The goods market clears at each point in time,*
4. *The evolution of the local population is consistent with individuals' optimal location decisions,*
5. *The evolution of the number of industrial plants is consistent with free entry.*

As already suggested by the aggregation results above, the equilibrium can be represented in a parsimonious way. First of all, the number of workers who are engaged in entry activities in region r is given by

$$H_{rEt} = f_E N_{rt-1}^{-\lambda} \left(N_{rt} - (1-\delta) N_{rt-1}\right) = \frac{1-\beta}{\rho\beta} H_{rPt} - (1-\delta) f_E N_{rt-1}^{1-\lambda}. \tag{21}$$

This implies that aggregate labor demand in the manufacturing sector is given by

$$H_{rMt} = H_{rPt} + H_{rXt} + H_{rEt} = \frac{1}{\beta} H_{rPt} - f_E N_{rt-1}^{1-\lambda}. \tag{22}$$

Hence, given the state variables N_{rt-1}, labor demand for the manufacturing sector is fully determined from the mass of production workers H_{rPt}. This implies that the dynamic equilibrium has a tractable formulation.

Proposition 2. *Let $\left\{L_{r0}^{I}, L_{r0}^{F}, N_{r0}, Q_{r0}\right\}_{r=1}^{R}$ be given. The equilibrium is fully characterized by the spatial goods market clearing equation*

$$\frac{w_{rMt}H_{rPt}}{\beta} = \left(\frac{P_{rM}}{P_M}\right)^{1-\sigma} (1-\alpha) \sum_{j=1}^{R} \left(\frac{w_{jAt}H_{jAt}}{\gamma} + \frac{w_{jMt}H_{jPt}}{\beta}\right) \tag{23}$$

$$\frac{w_{rAt}H_{rAt}}{\gamma} = \left(\frac{P_{rA}}{P_A}\right)^{1-\sigma} \alpha \sum_{j=1}^{R} \left(\frac{w_{jAt}H_{jAt}}{\gamma} + \frac{w_{jMt}H_{jPt}}{\beta}\right), \tag{24}$$

where $P_{rM} = \frac{1}{Q_{rt}N_{rt-1}^{\lambda\vartheta}H_{rPt}^{\vartheta}} w_{rMt}$, $P_{rA} = \frac{1}{Q_{rt}} w_{rAt} \left(\frac{H_{rAt}}{T_r}\right)^{1-\gamma}$ and $P_s = \left(\sum_r P_{rs}^{1-\sigma}\right)^{1/(1-\sigma)}$, the labor market clearing conditions

$$H_{rAt} = \Gamma_\theta \sum_{v=I,R} L_{rt}^{\nu} \left(\phi_A^v\right) \left(\frac{w_{rAt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1} \tag{25}$$

$$\frac{1}{\beta} H_{rPt} - f_E N_{rt-1}^{1-\lambda} = \Gamma_\theta \sum_{v=I,R} L_{rt}^{\nu} \left(\phi_M^v\right) \left(\frac{w_{rMt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1}, \tag{26}$$

evolution of the dynamic state variables L_{rt}^{ν} and N_{rt}

$$L_{rt}^{\nu} = (1-\psi) L_{r,t-1}^{\nu} + \psi L_{t-1}^{v} \frac{\left(V_r \overline{w}_{rt}^{\nu}\right)^{\varepsilon}}{\sum_k \left(V_k \overline{w}_{kt}^{\nu}\right)^{\varepsilon}} \text{ for } v = I, F, \tag{27}$$

$$N_{rt} = \frac{1-\beta}{\rho\beta} \frac{1}{f_E} H_{rPt} \times N_{rt-1}^{\lambda}, \tag{28}$$

and the low of motion for Q_{rt} in (8).

Proof. See Section 7.5 in the Appendix. □

Proposition 2 illustrates the link between this model and the commonly used static models of economy geography. Equations (23) and (24) are the usual trade balance equation, which determine (w_{rMt}, w_{rAt}) as a function of labor supplies (H_{rAt}, H_{rMt}). As highlighted already above: the short-run scale elasticity ϑ captures the strength of static agglomeration forces, whereby local productivity rises as the local population expands. Equations (25) and (26) are the regional labor market clearing conditions.

The crucial equations are the laws of motion for the dynamic state variables $\{L_{rt}, N_{rt}\}$ in (27) and (28). Whether the equilibrium allocations show persistence depends crucially on the extent of spatial mobility governed by ψ and the knowledge elasticity λ. With free mobility, $\psi = 1$, the distribution of people across space ceases to be a state variables. If in addition existing knowledge is not an input in the production of new knowledge, i.e. $\lambda = 0$, the model is a static model of economic geography with agglomeration forces as in Ahlfeldt et al. (2015). If in addition $\varepsilon = \infty$ so that individuals' spatial labor supply is perfectly elastic, the (amenity adjusted) real wages are equalized in equilibrium. This is the case analyzed in Allen and Arkolakis (2014).

These parametrizations of the model are not going to be consistent with the empirical findings reported above as the population shock was highly persistent and the effect on income per capita was increasing over time and hence showed a dynamic pattern. In contrast, as I will show now, these patterns are consistent with the model if spatial mobility is subject to frictions, i.e. $\psi < 1$, and the knowledge elasticity λ exceeds zero.

4.3 Entry, Population Inflows and Spatial Growth

At the heart of the theory is the joint evolution of the number of firms N_{rt} and the population L_{rt}. In particular, note that (28) implies that

$$\ln N_{rt} = \alpha_0 + \underbrace{\lambda \ln N_{rt-1}}_{\text{Mean reversion}} + \underbrace{\ln H_{rPt}}_{\text{Labor Supply Shocks}}, \tag{29}$$

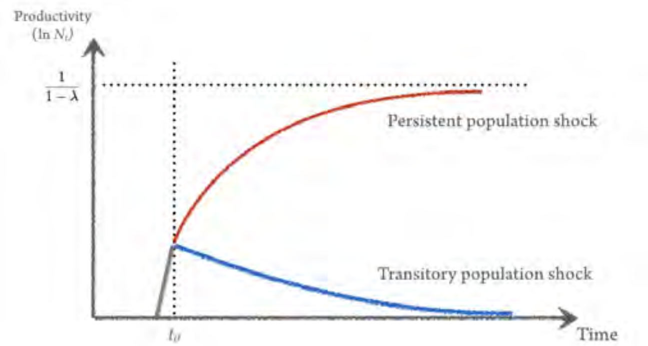
where $\alpha_0 = \ln\left(\frac{1-\beta}{\rho\beta}\frac{1}{f_E}\right)$. Hence, local productivity N_{rt} endogenously follows an AR(1) process, with two ingredients. First of all, spatial productivity N_{rt} is subject to shocks, which are fully summarized by the size of the local production workforce H_{rPt}. Crucially, such shocks are endogenous as they depend directly on the extent to which factors are mobile across space, how the spatial distribution of factor prices responds and the allocation of workers across sectors within a location. Secondly, the process is characterized by mean reversion if $0 < \lambda < 1$. In this case, shock to local labor supply will have persistent productivity effect but will not be permanent. In contrast, if $\lambda = 1$, local productivity is a random walk, transitory shock to local market size will have permanent effects on the level of productivity and the distribution of productivity across space will not be stationary.

To see this interaction between the amplifying forces of endogenous productivity and spatial mobility with frictions more clearly, note that (29) implies that for any period t_0 and $\tau \geq t_0$

$$\ln N_{r\tau} = \alpha_0 \sum_{j=t_0}^{\tau} \lambda^{j-(t_0-1)} + \lambda^{\tau-(t_0-1)} \ln N_{rt_0-1} + \sum_{j=t_0}^{\tau} \lambda^{-(j-(t_0+\tau))} \ln H_{rPj}, \tag{30}$$

Hence, given the initial state N_{rt_0-1}, productivity in any time period $\tau \geq t_0$ depends on the entire *history* of the manufacturing workforce $\{H_{rPj}\}_{j=t_0}^{t_0+\tau}$, appropriately discounted by inter-temporal knowledge elasticity λ. Intuitively: local productivity at any point in time encapsulates the entire history of local market size in the past, as historical market size led to plant entry, which has persistent effects because it makes the creation of future varieties easier. If $\lambda < 1$, proportional increases in the number of plants are getting more and more costly as the number of plants grows. If such a form of congestion is absent, i.e. $\lambda = 1$, there is no mean reversion and past population size does not have to be discounted at all, population shocks have persistent productivity effects and the distribution of productivity across space will not be stationary in the long-run.

Equation (30) is useful to think about the productivity effects of an exogenous change in labor supply at time t_0. The effect of such a shock on local productivity depends crucially on how quickly the shock subsides - either because people move spatially or reallocate to other sections. For concreteness, suppose there is a positive shock at t_0, which subsides at rate $p < 1$. Hence, for $j \geq t_0$ we have $d \ln H_{rPj} = d \ln H_{rPt_0} \times p^{j-t_0}$. If $p = 0$, the manufacturing workforce H_{rPj} returns to its pre-shock level immediately. If $p = 1$, the shock is persistent. As I show in Section 7 in the Appendix, Equation (30) implies that the elasticity of productivity at time τ with respect



Note: The figure displays the impulse response of local productivity to a one time increase in the local population ("transitory population shock") and a permanent increases in the local population ("persistent population shock").

Figure 7: Local productivity and Population size: Impulse response

to the shock at time t_0 is given by the impulse response function

$$\frac{d\ln N_{r\tau}}{d\ln H_{rPj}} \equiv \Psi_\tau\left(p,\lambda\right) = \frac{\lambda^{\tau+1} - p^{\tau+1}}{\lambda - p}.$$

If the shock is purely transitory, i.e. $p = 0$, we get that

$$\Psi_\tau\left(0,\lambda\right) = \lambda^\tau \underset{\tau\to\infty}{\longrightarrow} 0. \tag{31}$$

Hence, a transitory population inflow shows persistence but the effect is *declining* over time. Moreover, the higher the knowledge elasticity λ, the more prolonged the productivity response.

On the other extreme, if the shock was purely persistent, i.e. $p = 1$, we have

$$\Psi_\tau\left(1,\lambda\right) = \frac{1-\lambda^{\tau+1}}{1-\lambda} \underset{\tau\to\infty}{\longrightarrow} \frac{1}{1-\lambda}, \tag{32}$$

i.e. the effect is *increasing* over time and the endogenous creation of varieties acts as an amplifying force. This is exactly the pattern, which I documented empirically: the short-run effect of the inflow of refugees on GDP pc in 1950 was small, the long-run effect on GDP pc the 1960s was large. If $\lambda = 1$, the productivity effect keeps accumulating and $\Psi_\tau\left(1,1\right) \to \infty$. Finally, if $0 < p < 1$, the effect of the shock will always subside in the very long-run, but the impulse response can be hump-shaped. The two polar cases of full and no persistence are displayed in Figure 7.

4.4 A Spatial Balanced Growth Path: Endogenous vs Semi-Endogenous Growth

I now characterize the behavior of the economy in the long-run, i.e. along a non-stochastic spatial balanced growth path (SBGP), which I define as an allocation where the population distribution is constant across space and regional wages and incomes grow at a common rate. Along a SBGP innate productivity Q_{rt} has to be constant and equal to its long-run level Q_r. Hence, locations are heterogeneous in three fundamentals: innate productivity Q_r, amenities V_r and the endowment of land T_r.

With a stationary population, goods market clearing implies that regional productivities N_{rt} need to grow at a common rate g_N in all counties. Using (28) this implies that

$$g_N = \frac{N_{rt}}{N_{rt-1}} = \frac{1-\beta}{\rho\beta}\frac{1}{f_E}H_{rPt}N_{rt-1}^{\lambda-1}. \tag{33}$$

Equation (33) has a very similar structure to the growth equation analyzed in Jones (1995). For g_N to indeed be constant across space, (33) requires $H_{rPt}N_{rt-1}^{\lambda-1}$ to be constant across space.

Consider first the case where $\lambda < 1$. (33) then implies that the level of productivity across regions along a SBGP is given by

$$N_{rt} = \left(\frac{1}{g_N}\frac{1-\beta}{\rho\beta}\frac{(1-\delta)^{\lambda}}{f_E}\right)^{\frac{1}{1-\lambda}} H_{rPt}^{\frac{1}{1-\lambda}}. \tag{34}$$

Hence, the equilibrium growth rate in productivity g_N is given by $g_N = n^{\frac{1}{1-\lambda}}$, where n is the growth rate of the manufacturing workforce. This secular growth could either be driven by population growth or by the accumulation of efficiency units though increases in human capital. Hence, if $\lambda < 1$ this is a model of semi-endogenous growth, where the growth rate is endogenous but in equilibrium determined directly from exogenous parameters. More importantly, (34) has strong implications for the spatial distribution of economic activity in the long-run. First of all, the spatial distribution of productivity in the long-run exactly reflects local scale effects: regions where H_{rPt} is large will have high productivity. Hence, regional scale has *level* effects, not *growth* effects.[11] Secondly, the long-run distribution of productivity N_{rt} and population is stationary and fully determined from regional fundamentals. Crucially, the mapping from these fundamentals to the endogenous levels of productivity N_{rt} and the spatial distribution of economic activity depends directly on the inter-temporal knowledge elasticity λ, which acts as an amplifying force: small differences in innate productivity Q_r or amenities V_r can lead to sizable spatial income differences through the endogenous increase in local productivity. Below, I derive this mapping explicitly. Note that this determinacy of the equilibrium as a function of fundamentals also implies that the effect of population shocks will necessarily vanish in the long-run as long as there is some mobility, i.e. $\psi > 0$.

The case of $\lambda = 1$ is qualitatively different. First of all, it is apparent from (33) that - generically - there does not exists a SBGP because this would require the amount of human capital to be equalized across space. This will only be the satisfied in knife-edge cases as differences in amenities V_r will cause differences in the local population or differences in productivity Q_r or land endowments T_r will affect relative factor prices and therefore the sectoral allocation of resources and H_{rPt}. Hence, if $\lambda = 1$, the economy will not permit a stationary distribution of economy activity and initial conditions matter: regions that start "large" will grow faster, which might draw in even more people. In the same spirit, population shocks like the inflow of refugees have persistent effects on the distribution of economic activity. The linear relationship between spatial 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 (34) therefore highlights the spatial analogue of the distinction between endogenous and semi-endogenous models of growth: the spatial distribution of economic activity will be stationary in the latter but not stationary in the former.

Equation (34) also highlights that the long-run relationship between productivity and population size is fundamentally different from the short-run relationship. Equations (19) and (20) showed that the labor productivity in the manufacturing sector is given by

$$Y_{rMt}/H_{rPt} \propto Q_{rt}N_{rt}^{\lambda\vartheta}H_{rPt}^{\vartheta}.$$

Hence, the impact elasticity between local productivity and marker size is given by the short-run scale elasticity $\vartheta = \frac{1-\beta}{\rho-1}$. Equation (34) shows that this elasticity along a SBGP is quite different. In particular, substituting for

[11] Note that regional scale in the manufacturing sector can be achieved in three ways: through a large population L_{rt}, through sorting whereby the share of industrial types ω_r^I is large or through within-region labor supply whereby relative manufacturing wages w_{rMt}/ϖ_{rt}^I are large. Hence, the long-run distribution of productivity is jointly determined with the distribution of individuals and their sorting behavior across sectors and space.

the BGP value of N_{rt}, we get that

$$Y_{rMt}/H_{rPt} \propto Q_r H_{rPt}^{\frac{\lambda\vartheta}{1-\lambda}} H_{rPt}^{\vartheta} = Q_r H_{rPt}^{\frac{\vartheta}{1-\lambda}}.$$

Hence, the long-run elasticity between labor productivity is given by $\vartheta/(1-\lambda)$. Importantly, the knowledge elasticity λ acts as an amplifying force so that the long-run elasticity is larger than the short-run elasticity. If $\lambda = 0$ as in Krugman (1980b), there is no amplification as the entry costs are independent of the number of firms located in a particular location. As λ approaches the endogenous growth limit of $\lambda = 1$, the long-run elasticity becomes arbitrarily large as a higher labor force increases the growth rate of productivity.

Two Special Cases To highlight how the long-run regional distribution of productivity is determined in equilibrium and how the knowledge elasticity λ amplifies existing differences in regional fundamentals, it is useful to consider two polar cases for individuals' sorting behavior.

Suppose first that there is *no* substitution across skill groups, i.e. industrial workers can only work in the manufacturing sector and farmers can only provide their skills to the agricultural sector. Formally, $\phi_M^R = \phi_A^I = 0$ so that $\overline{w}_{rt}^I = w_{rMt}$ and $\overline{w}_{rt}^F = w_{rAt}$. This implies that total human capital in the production sector is proportional to the number of industrial workers, i.e. $H_{rPt} \propto L_{rt}^I$.[12] Equation (34) then directly implies that

$$N_{rt} \propto \left(L_{rt}^I\right)^{\frac{1}{1-\lambda}}. \tag{35}$$

The distribution of industrial workers in a SBGP is then described by the spatial mobility condition

$$L_{rt}^I \propto \left(V_r w_{rMt}\right)^{\varepsilon}. \tag{36}$$

Finally, good market clearing requires that

$$w_{rMt} L_{rt}^I \propto \left(\frac{w_{rMt}}{Q_r N_r^{\lambda\vartheta}}\right)^{1-\sigma}. \tag{37}$$

Equations (35), (36) and (37) fully summarize the long-run distribution of endogenous productivity N_{rt}, manufacturing wages w_{rMt} and individuals L_{rt}^I as a function of the underlying fundamentals: amenities V_r and innate productivity Q_r. In particular, they imply that (see Section 7.2 in the Appendix)

$$\ln L_{rt}^I = \kappa_L + \frac{\sigma-1}{\zeta} \times \ln Q_r + \frac{\sigma}{\zeta} \times \ln V_r \tag{38}$$

$$\ln w_{rMt} = \kappa_w + \frac{1}{\varepsilon}\frac{\sigma-1}{\zeta} \ln Q_r + \frac{1-\zeta}{\zeta} \times \ln V_r, \tag{39}$$

where

$$\zeta \equiv 1 + \frac{\sigma}{\varepsilon} - \frac{(\sigma-1)\vartheta}{1-\lambda} > 0$$

and κ_L and κ_w summarize terms, which are constant across space.[13]

Equation (38) shows - as expected - that the distribution of employment reflects both differences in amenities and in innate productivity.[14] The endogenous nature of productivity *amplifies* such differences, as N_{rt} is increasing

[12] Along a SBGP, the amount of entry labor is given by $H_{rEt} = \frac{1-\beta}{\rho\beta}\left(1 - \left(\frac{1-\delta}{\gamma}\right)\right) H_{rPt}$. Hence, $H_{rPt} \propto H_{rMt} = \phi_M^I L_{rt}^I$.

[13] The restriction that $\zeta > 0$ is a stability condition. See Section 7.2 in the Appendix for details.

[14] Note that the land endowment T_r does not feature in these equations. This is an implication of the assumptions that individuals

in both Q_r and V_r. The parameter ζ summarizes the amplifying forces of the possibility of spatial growth. In particular, ζ is decreasing in both the short-run elasticity ϑ and the knowledge elasticity λ. Hence, higher levels of ϑ or λ increase the elasticity of regional population size with respect to productivity or amenities and hence amplify exogenous sources of spatial heterogeneity. The case of exogenous productivity is nested as $\beta = 1$, which implies that $\vartheta = 0$. In that case, $\zeta = 1 + \frac{\sigma}{\varepsilon}$ as in the standard economic geography model with idiosyncratic location preferences.

Equation (39) concerns the distribution of regional wages. As expected, wages are increasing in innate productivity and - as for the level of population - such differences are amplified if ϑ and λ are large. More interestingly, equation (39) also shows that regional wages might be *increasing* in local amenities if $\zeta < 1$, i.e. if the long-run scale elasticity is large enough

$$\frac{\vartheta}{1-\lambda} > \frac{\sigma}{\varepsilon(\sigma-1)}.$$

The relationship between amenities and equilibrium wages, depends on two countervailing forces. Holding local technologies fixed, wages have to compensate individuals to live in unfavorable locations. This induces a negative correlation between wages and amenities. At the same time, higher amenities have a positive effect on productivity by increasing the local population. As long as the long-run scale elasticity is large enough, this second effect dominates in the long-run, making "nicer" places also high-paying locations. If the endogenous technological response is absent, i.e. if $\vartheta = 0$, wages are necessarily decreasing in local amenities holding innate productivity Q_r fixed.

The theory stresses the complementarity between the local market size and productivity. Note that the manufacturing sector in particular location competes for workers spatially through individuals' migration decision and with the agricultural sector within the same location. This implies that the agricultural sector plays a dual role for technological development of the manufacturing. To see this clearly, consider the case where there is only a single worker type, i.e. $\phi_j^I = \phi_j^R$. In that case the long-run level of manufacturing technology in region r is given by[15]

$$N_{rt} \propto \left(L_{rt} \left(\frac{w_{rM}}{\overline{w}_r} \right)^{\theta-1} \right)^{\frac{1}{1-\lambda}} \text{ and } L_{rt} \propto \left(V_r \overline{w}_r \right)^{\varepsilon}. \tag{40}$$

Expectedly, local productivity is increasing in regional amenities V_r and in the manufacturing wage w_{rMt} as both contribute to a rise in market size - the former through attracting workers spatially, the latter though a sectoral reallocation within regions. The effect of *average* earnings $\overline{w}_{rt}$ is more subtle. Holding the population L_{rt} and the manufacturing wage w_{rMt} constant, a higher $\overline{w}_{rt}$ reflect a higher wage in the agricultural sector. This reduces manufacturing labor supply and hence long-run productivity by making the competing agricultural sector more attractive. The strength of this force depends on θ. At the same time, a higher average income $\overline{w}_{rt}$ also increases the spatial population though the spatial equilibrium and some of these workers will end up working in the manufacturing sector. The latter effect dominates if $\varepsilon > \theta - 1$, i.e. if the spatial labor supply elasticity ε is large relative to the sectoral supply elasticity elasticity θ. If that is the case, spatial mobility makes different sectors within a region complements, i.e. the local manufacturing sector benefits from having an efficiency agricultural sector near by. The elasticity λ is again the key determinant of how potent these effects are.

are fully specialized by skill (so that industrialists' labor supply does not depend on w_{rAt}) and the Cobb-Douglas utility function, which implies that total spending on manufacturing products does not depend on the agricultural wage.

[15]To derive (40), note that the amount of human capital in production H_r^P is proportional to the total supply of manufacturing human capital H_r^M, which in turn is proportional to $L_{rt}\pi_{rM}^{\frac{\theta-1}{\theta}} \propto L_{rt}\left(\frac{w_{rM}}{\overline{w}_r}\right)^{\theta-1}$. Hence, $N_{rt} \propto \left(\left(\frac{w_{rs}}{\overline{w}_r}\right)^{\theta-1} L_{rt}\right)^{\frac{1}{1-\lambda}}$. The spatial mobility condition is simply $L_{rt} \propto (V_r\overline{w}_r)^{\varepsilon}$.

5 Quantitative Analysis: The Economic Effects of the Refugee Settlement

The model above is a parsimonious framework to quantify the short- and long-run effects of population inflows. In this section I estimate the structural parameters of the theory to the empirical results established in Section 3. This exercise has two purposes. First, it shows that the empirical results presented in Section 3 are quantitatively consistent with this theoretical framework. In particular, I show that the structural estimates imply an estimate of the knowledge elasticity λ, which is way below unity and hence leads me to conclude that the German post-war development is consistent with a model of semi-endogenous growth. Secondly, the model allows me to answer questions, which cannot be answered from the empirical findings alone. Most importantly, it allows me to quantify the effect of the refugee-settlement on aggregate growth in Germany.

Before calibrating the model, let me provide some direct evidence on the main mechanism underlying the theory. The theory highlights the endogenous link between population size and manufacturing productivity through the entry of new plants. 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 notes 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) From the digitized historical manufacturing census files for 1933, 1939 and 1956, I can measure the number of manufacturing plants at time t. This allows me to relate the growth rate of manufacturing plants $g_{rNt} = ln\left(N_{rt}/N_{r1939}\right)$ to the inflows of refugees.

Table 7 reports the results of regressing the growth rate of plants g_{rNt} on the share of refugees in 1950 in a specification akin to (2). In the first column I show that - reassuringly - there is no relationship between the allocation of refugees and plant entry between 1933 and 1939. Running the same specification for the growth rate between 1939 and 1950 yields a sizable positive effect, whereby an increase in the share of refugees by 10 percentage points increases the number of manufacturing plants in 1950 by 4%. Column 3 shows that this effect is even more pronounced by 1956. Hence, as implied by the theory, there is a positive relationship between an inflow of refugees and the number of active plants and this increase is more pronounced in the long-run compared to the short-run.

5.1 Estimation and Identification

The model is fully parametrized by 14 structural parameters and a tupel of fundamentals $[Q_r, V_r, T_r]$ per region. The parameters and fundamentals are summarized in Table 8. I calibrate 6 parameters externally. In particular, I set the trade elasticity σ to 5, set the labor share in agricultural (γ) and manufacturing (β) to 0.5 and 0.6 respectively, assume values of the spatial and sectoral labor supply elasticity ε and θ, which are in the ballpark of the earlier literature and set the depreciation rate δ to 0.2. This leaves me with 8 parameters, which are particular to the current context: the share of industrialists χ and their human capital in the manufacturing sector ϕ_M^I, the expenditure share on agriculture α, the extent of labor mobility as parametrized by the Calvo-type mobility shock ψ, the parameters of the productivity process (ϖ, ϱ) and - most importantly - the short-run scale elasticity ϑ and the knowledge elasticity λ.[16] My strategy to estimate these parameters and regional fundamentals is as follows. I assume that the economy is in a steady-state in 1933. Using data on the allocation of population L_{r1933}, sectoral employment shares s_{r1933}^M and income per capita y_{r1933}, I can identify the fundamentals $[Q_r, T_r, V_r]$ for a given set of parameters. Formally, there is a 1-1 mapping between the three fundamentals and three moments for each region.

[16]Note that $\vartheta = \frac{1-\beta}{\rho-1}$ so that (for given β) ϑ is a one-to-one function of ρ.

	ln num of manufacturing plants		
	1939	1950	1956
Share of refugees in 1950	0.217	0.407**	0.873**
	(0.290)	(0.203)	(0.434)
ln pop dens 1939	✓	✓	✓
Wartime destr.	✓	✓	✓
ln plants in 1933	✓		
ln plants in 1939		✓	✓
Manufac. share in 1933	✓		
Manufac. share in 1939		✓	✓
Geography	✓	✓	✓
State FE	✓	✓	✓
Observations	506	519	519
R^2	0.797	0.864	0.770

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The regression is at the county level. The dependent variable is the log of the number of manufacturing plants in 1939 (column 1), 1950 (column 2) or 1956 (column 3). All specifications control for the log of population density in 1939, the extent of wartime destruction through the share of the housing stock, which was damaged during the war, state fixed effects. Specifications with geography controls also control for the log of the distance to the inner german border and a fixed effect for whether a county is a border county.

Table 7: Refugees and the Entry Manufacturing Plants

	Parameters	Value	Moment
	Estimated parameters from post-war data		
χ	Share of industrial workers	0.39	$\ln y_i = \beta \times Refugee_i + x_i'\gamma + u_i$
ϕ_M^I	Comp. adv. of industrialist	54.7	$s_{r,1950}^M = \beta \times \mu_{r,1950} + x_r'\gamma + u_r$
ρ	Elasticity of substitution	2.4	$\ln y_{r,1950} = \beta \times \mu_{r,1950} + x_r'\gamma + u_r$
λ	Long-run DRT in entry	0.69	$\ln y_{r,1961} = \beta \times \mu_{r,1950} + x_r'\gamma + u_r$
ψ	Frequency of mobility shocks	0.03	$\mu_{r,1961} = \beta \times \mu_{r,1950} + \delta_{State} + u_r$
ϱ	Correlation of productivity	0.98	$\ln y_{r,t} = \alpha_r + \beta \times \ln y_{r,t-1} + v_{rt}$
ϖ	Dispersion of productivity shocks	0.064	$\sigma\left(\hat{v}_{rt}^2\right)$
	Inferred fundaments from data in 1933		
α	Spending share on agricult.		Agg. Agricult. empl. share (1933)
$[V_r]$	Spatial amenities	See	Population share (1933)
$[Q_r]$	Spatial innate productivity	Section 9	GDP pc (1933)
$[T_r]$	Spatial supply of land	in Appendix	Agricult. empl. share (1933)
	Set exogenously		
σ	Trade elasticity	5	
(γ, β)	Production function	$(0.5, 0.6)$	
ε	Spatial labor supply elasticity	1.5	
θ	Sectoral labor supply elasticity	2	
δ	Frequency of exit	0.2	

Notes: The table reports the structural parameters of the model. The elasticity of labor supply ε stems from Fajgelbaum and Gaubert (2018). The sectoral labor elasticity θ stems from Eckert (2019).

Table 8: Structural parameters

	Manufac. share	Refugee share		ln GDP pc		ln earnings
	1950	1961	1950	1961	Panel	
Share of refugees in 1950	0.264^{***}	0.601^{***}	0.098	0.502^{***}		
	(0.072)	(0.035)	(0.321)	(0.191)		
$\ln y_{r,t-1}$					0.555^{***}	
					(0.041)	
Refugee						-0.098^{***}
						(0.008)
log GDPpc 1935			✓	✓		
ln pop dens 1939	✓		✓	✓		
Wartime destr.	✓		✓	✓		
Manufac. share 1939	✓					
Geography	✓		✓	✓		
Demographics						✓
State FE	✓	✓	✓	✓		✓
County FE					✓	
Observations	535	484	524	520	1650	32584
R^2	0.863	0.733	0.699	0.500	0.894	0.316

Table 9: Regression evidence for model to match

Note also that the steady-state implies the distribution of the endogenous number of firms N_{r1933} and the extent of spatial sorting, i.e. the allocation of industrial types ω^I_{r1933} across counties (see Section 9 in the Appendix).

I then identify the structural parameters using indirect inference by running the regressions of Section 3 in the model. In particular, starting from the BGP allocations in 1933 I simulate the equilibrium allocations of the model at the yearly level and "shock" the economy with the inflow of refugees in the post-war period. Because the majority of refugees arrived around the year 1947 (see Figure 6), I assume that all refugees arrived in 1947 and I allocate them according to empirically observed share of refugees in 1950.[17] I then simulate the evolution of the economy until 1966 and run five cross-county regressions in the data and the model and target the coefficients.

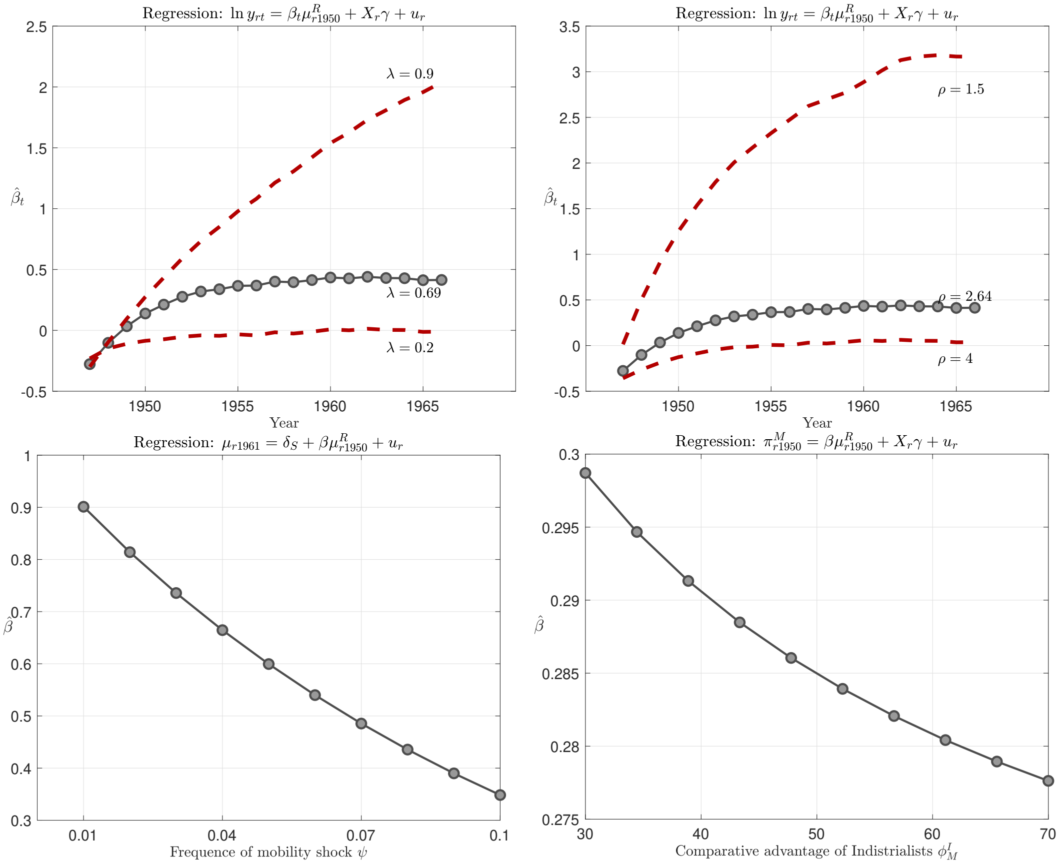
First I use the relationship between the manufacturing employment share in 1950 and the share of refugees in 1950 reported in Table 4. Second I target the long- and short-run relationship between income per capita and the share of refugees reported in Table 5. Third I force the model to replicate the spatial persistence of refugees between 1950 and 1961 shown in Table 3. Fourth I estimate the process of spatial income per capita for the years 1957, 1961, 1964 and 1966 from the specification

$$\ln y_{rt} = \alpha_r + \beta \ln y_{rt-1} + v_{rt}, \tag{41}$$

and target the point estimate for β. Note that α_r is a county fixed effect so that this specification mirrors the structural productivity process for Q_{rt} (see (8)). For convenience I report these specifications in the first 5 columns of Table 9 and also report the exact controls, which I use.

Even though the parameters are calibrated jointly, the regression coefficients reported in Table 9 have a clear mapping to the structural parameters of the theory. Consider first the long and short-run effects of the allocation of refugees on GDP per capita reported in columns 3 and 4. As highlighted in the theory: the long-run elasticity between population inflows and income per capita is mostly affected by λ, the short-run elasticity is mostly affected by ϑ. Hence, the regression coefficient of 0.1 for income per capita in 1950 and 0.5 in 1961 are informative about ϑ and λ respectively. To see this, the top panels in Figure 8 depict regression coefficients from the regression of $\ln y_{rt}$

[17] 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 is small (see Section 9 in the Appendix).



Note: The top-left (top-right) panel plots the regression coefficient β_t from specifications 3 and 4 of Table 9 for different choices of λ (ρ). The bottom left panel reports the regression coefficient β from specifications 2 of Table 9 as a function of ψ. The bottom right panel reports the regression coefficient β from specifications 1 of Table 9 as a function of ϕ^I_M.

Figure 8: Identification of Structural Parameters

on the share of refugees in 1950 (and controls) for different levels of λ (left panel) and ρ (right panel). The left panel shows that λ governs the long-run effect of population inflows on income per capita. The larger λ, the bigger the impact of the initial shock on income per capital in the long-run. The right panel shows that ρ is particular informative about the short-run impact. The lower ρ, the more potent the static agglomeration force and the higher the impact of the inflow of refugees on income per capita in the short-run.

The correlation between the share of refugees in 1950 and in 1961, i.e. column 2 in Table 9, can be used to identify the extent of mobility frictions captured by ψ. Intuitively, ψ governs how quickly regions with an initial "excess" of refugees revert back to the mean. If $\psi = 0$, refugee inflows are naturally persistent. If $\psi = 1$, the model predicts that the share of refugees should be equalized across space within a single period.[18] In panel 3 of Figure 8 I show the implied regression coefficient for different values of ψ. There is a strong negative relationship. To match a coefficient of 0.6 as estimated in the data, the model requires a value of ψ of around 4%. Finally, the cross-sectional relationship between the manufacturing employment share and the share of refugees, i.e. column 2 in Table 9, is informative about the extent of comparative advantage of industrialists, i.e. ϕ^I_M. In particular, Figure 8 shows that

[18] In particular, the model implies that the local share of refugees π^R_{rt} relative to the economy-wide average $\overline{\pi}^R$ evolves according to

$$\pi^R_{rt} - \overline{\pi}^R = \left(\pi^R_{rt-1} - \overline{\pi}^R\right) \frac{(1-\psi)\, L_{rt-1}}{(1-\psi)\, L_{rt-1} + \psi \frac{(V_r w_{rt})^{\varepsilon}}{\sum_k (V_k w_{kt})^{\varepsilon}} L_{t-1}}.$$

a larger value of ϕ_M^I reduces the extent to which an inflow of refugees increases the manufacturing employment share.[19]

To identify the parameters of the exogenous productivity process for Q_{rt}, I use the panel regression of income per capita in 41, which is reported in column 5 of Table 9. The correlation parameter ϱ is an important determinant of the serial correlation of income per capita. To calibrate the dispersion parameter ϖ, I target the dispersion in the estimated residuals, i.e. $var\left(\hat{v}_{rt}\right)$, from 41.

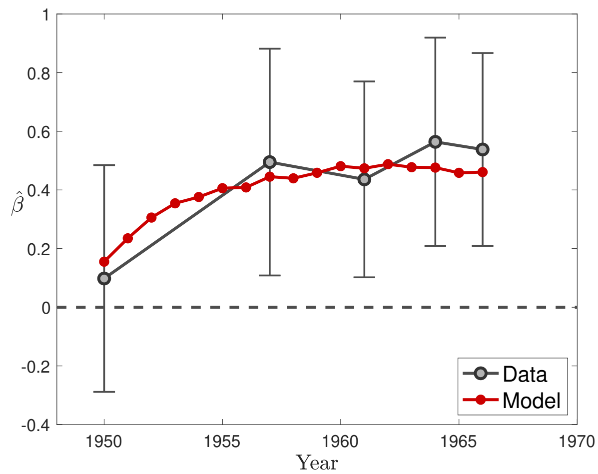
Last, I estimate the extent of skill heterogeneity χ from the differences in earnings between refugees and natives. As reported in the last column of Table 9 I rely on the micro data from the EVS and estimate a Mincer-type regression of log earnings on demographics and a dummy for whether or not the individual is a refugee. Empirically, refugees earn 10% less than native on average. Though the lens of the model, the only difference between refugees and natives is the extent of spatial sorting. Without skill heterogeneity, there would be no systematic earnings differences between natives and refugees. With skill heterogeneity, industrial workers among native sort towards regions, which are specialized in non-agricultural production and which offer higher wages. In contrast, refugees are not sorted upon their arrival but only sort gradually. The parameter χ is therefore an important determinant of the relative earnings between natives and refugees.

5.2 Model Fit

The estimated model captures the salient features of the spatial allocation of economics activity between 1950 and 1965 reasonably well. To see this, Figures 9 and 10 summarize different aspects of the model fit. In Figure 9 I report the effect of refugee inflows on income per capita at different horizons, i.e. the results reported in Table 5 and in columns 3 and 4 of Table 9. I depict both the regression coefficients from the data (together with the associated standard errors) and the corresponding estimates from the model. The model is calibrated to match the response in 1950 and 1961. Figure 9 shows that the model also captures the dynamics well. In particular, it replicates the fact that the effect on GDP per capita is increasing and that it plateaus in the early 60s (even though empirically the standard errors are large). Hence, the combination of parameters governing the extent of spatial mobility (ψ) and the variety gains (ρ and λ) can replicate the effect of refugee inflows on GDP per capita in the short- and long-run.

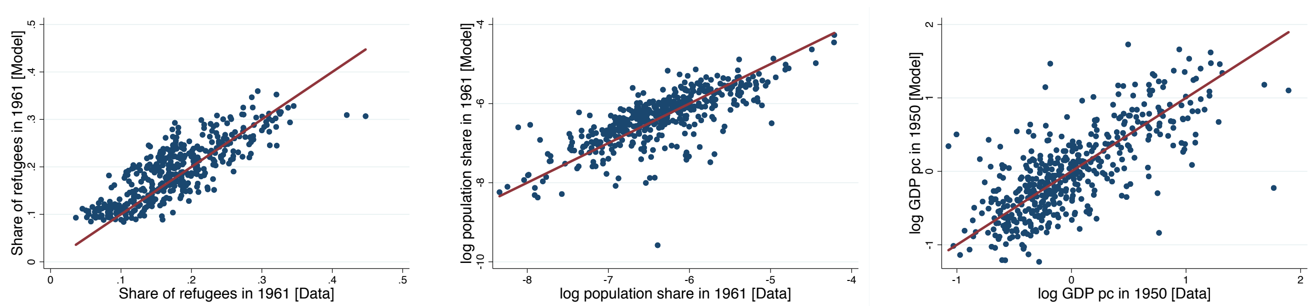
In Figure 10 I zoom into the spatial aspects of the allocation both in the model and the data. Recall that the model is calibrated to match the population distribution and income per capita in 1933 and that it by construction matches the share of refugees in 1950. In as far as the post-war allocations are concerned, the model is only calibrated to match the regressions reported in Table 9. Figure 10 nevertheless shows that the model replicates a variety of aspects of the spatial allocation. In particular, I report - both for the data and the model - the share of refugees in 1961, the population share and income per capita in1950. Note that the red lines are 45 degrees lines and not best-fit lines. The two panels on the left and in the center, show that the model approximates the endogenous population distribution in 1961 well: both the number of people (middle panel) and the share of refugees (left panel) in the model are a very good predictor for the actually allocation observed in the data. The discrepancies between the model and the data arise naturally due to the idiosyncratic productivity shocks of the productivity process, which imply that the allocations in the model are stochastic. In the right panel, I report the correlation of GDP pc in 1950 within states. As for the population distribution, the model provides a reasonable fit for the empirically observed distribution of GDP pc.

[19] The intuition for this result is the following: the higher ϕ_M^I, the more are native workers already spatially sorted. Given that refugees are allocated towards rural areas, the higher ϕ_M^I, the fewer industrialist native workers are present in the labor markets that experience a refugee inflow. This makes the aggregate labor supply less elastic and there is less crowding-in among natives.



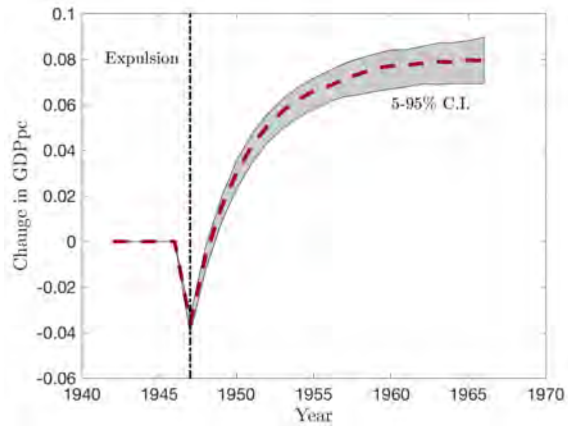
Notes: This figure reports the regression coefficient $\hat{\beta}_t$ from the specification $\ln y_{rt} = \beta_t \times \mu_{r1950} + x'_{rt}\gamma + u_{rt}$ where μ_{r1950} denotes the share of refugees in 1950, $\ln y_{rt}$ denotes income per capita at time t and x_{rt} contains controls (state FE, $\ln y_{r1933}$, log population density in 1939, wartime destruction, border FE and distance to the inner-german border). The black line corresponds to the data (including 90% confidence intervals) the red line corresponds to the model. The coefficients in 1950 and 1961 are explicitly targeted (see Table 8).

Figure 9: The Dynamic Effect of Refugee Inflows on Income per Capita



Notes: The figure reports the correlation between the allocations in the model and the data for the share of refugees in 1961 (left panel), the distribution of the population in 1961 (middle panel) and log GDP pc in 1950 after taking out a set of state fixed effects (right panel).

Figure 10: Fit of the Model



Note: The figure displays the aggregate treatment effect of the refugee inflow. Specifically, I calculate aggregate GDPpc for the model with refugee inflows relative to GDPpc for a counterfactual economy without the refugee inflow. The red lines reflects one particular sample path of the productivity process Q_{rt}. The shaded area displays a 90% confident interval from the bootstrap distribution of the aggregate treatment effect.

Figure 11: The Aggregate Impact of the Refugee Settlement

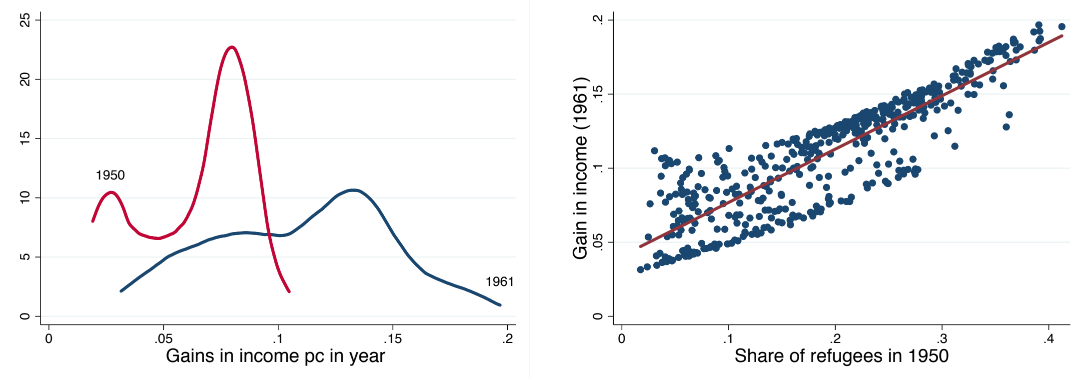
5.3 The Aggregate Effects of the Refugee-Settlements

How large was the aggregate impact of the refugee settlement on economic activity in Germany? While this object is not identified from the empirical estimates which rely on cross-sectional information, I can compute this "aggregate treatment effect" in the calibrated model by comparing the equilibrium with the refugee inflows with a counterfactual Germany where the refugees did not arrive. Importantly, I hold the sequences of productivity shocks constant when doing the comparison. The results are reported in Figure 11.

Figure 11 shows that the influx of refugees initially reduces GDPpc by about 3%. This is due to the fact that agricultural production is subject to decreasing returns. Due to the endogenous nature of technological progress, however, this initial drop is relatively short-lived and the population increase causes income per capita to increase. Given the estimated parameters, the aggregate effect peaks after roughly 15 years. Quantitatively, the refugee inflow increases GDPpc by about 5% at the 15 year horizon.

Because spatial productivity Q_{rt} is subject to shocks, the exact treatment effect depends on the particular sample path of local technologies. To get a sense on the importance of this source of randomness for estimated treatment effect, I also report the associated confidence intervals. I calculate these by a bootstrap procedure where I simulate the economy for M different histories of potential productivity histories $\{Q_{rt}\}_{rt}$ and then calculate the 5 and 95% quantile of the distribution of the estimated treatment effects. Figure 11 shows that the implied treatment effect is reasonably precisely estimated.

This aggregate effect masks a vast amount of spatial heterogeneity. This is for example seen in Figure 12. In the right panel, I depict the distribution of changes in income per capita in 1950 (red line) and 1961 (blue line). It is clearly apparent that some regions win more than others and that this dispersion is increasing between 1950 and 1961, reflecting the slow build of local technologies. In the right panel I show the link between local growth and local market size: the higher the inflow of refugees, the higher the increase in income per capita. 10 - 15 years after the expulsion, there is still a sizable gradient. Of course, in the very long-run, the population inflow is transitory and the initial inflow does not predict spatial growth. Given the estimated mobility frictions however, this process takes a long time to play out. Furthermore, the local build up of technology is an additional force limiting the extent of population outflows.



Note: The left panel depicts the distribution of average counterfactual income gains in 1950 and 1961. The right panel depicts the average counterfactual income gains in 1961 against the share of refugees in 1950. The average counterfactual income gains are calculated as the average (over 50 sample paths of productivity) relative income per capita in the calibrated equilibrium relative to an economy without the refugee inflow.

Figure 12: Spatial Heterogeneity of the Impact of the Refugee Settlement

6 Conclusion

In this paper I used a particular historical setting to empirically estimate wether local technology responds to changes in local factor supplies. I focused on the expulsion of the ethnic German population from their domiciles in Central and Eastern Europe and their subsequent resettlement in Western Germany. In the three years after the Second World War, roughly 8m people arrived in Western Germany. At the time, this amounted to an increase in population by about 20%. Furthermore, counties in Western Germany differed substantially in the extent to which they participated in the incoming refugee flows. Using both the policies of the US and UK Military Government in Post-War Germany and the pre-war geographic distance from counties in Western Germany to the expulsion regions to isolate the exogenous component in refugee flows, I study the long-run economic consequences of such labor supply shocks across 500 counties in Western Germany.

I find a positive relationship between refugee inflows and the size of the manufacturing sector and local income per capita. Together with the fact that refugees' earning were lower then the ones of natives, these cross-county results are consistent with models featuring an endogenous response of local technology to population size, but hard to reconcile with a neoclassical model, where productivity is exogenous.

I then propose a new spatial growth model, which can rationalize the above findings. In particular, the model implies that the effect of population inflows can increase over time if frictions of mobility are important and technological progress is endogenous. I calibrate the structural parameters to the regression evidence of the natural experiment using indirect inference. The model is able to quantitatively match the salient features of this historical episode and implies that the settlement of refugees increased aggregate income per capita by about 7% after 15 years. This is in stark contrast with the short-run effect, which lowered income per capita by 3 %.

References

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

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

_ , "When does labor scarcity encourage innovation?," *Journal of Political Economy*, 2010, *118* (6), 1037–1078.

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," *Unpublished manuscript*, 2018.

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.

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.

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.

Desmet, Klaus and Esteban Rossi-Hansberg, "Spatial Development," *American Economic Review*, 2014, *104* (4), 1211–43.

_ **, David Krisztian Nagy, and Esteban Rossi-Hansberg**, "The Geography of Delevopment: Evaluating Migration Restrictions and Coastal Flooding," 2015. Working Paper.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "Labor supply shocks, native wages, and the adjustment of local employment," *The Quarterly Journal of Economics*, 2016, p. qjw032.

Eckert, Fabian, "Growing apart: Tradable services and the fragmentation of the us economy," 2019.

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," Technical Report, National Bureau of Economic Research 2016.

Fajgelbaum, Pablo and Cecile Gaubert, "Optimal spatial policies, geography and sorting," Technical Report, National Bureau of Economic Research 2018.

— **and Stephen J. Redding**, "External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914," Working Paper 20217, National Bureau of Economic Research June 2014.

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.

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.

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

Krugman, Paul, "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 1980, pp. 950–959.

— , "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 1980, *70* (5), 950–959.

Kucheryavyy, Konstantin, Gary Lyn, and Andrés Rodríguez-Clare, "Grounded by Gravity: A Well-Behaved Trade Model with Industry-Level Economies of Scale," Technical Report, National Bureau of Economic Research 2016.

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

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.

Nunn, Nathan, Nancy Qian, Sandra Sequeira et al., "Migrants and the Making of America: The Short and Long Run Effects of Immigration during the Age of Mass Migration," Technical Report, CEPR Discussion Papers 2017.

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.

— , "Population of the US Zone of Germany (Summary)," Technical Report, Civil Administration Divison 1947.

— , "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.

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

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," *The American Economic Review*, 2008, *98* (5), 1766–1797.

Redding, Stephen J. 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.

Rivera-Batiz, Luis A. and Paul M. Romer, "Economic Integration and Endogeneous Growth," *Quarterly Journal of Economics*, 1991, pp. 531–55.

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

Statistisches Bundesamt, "Die Vertriebenen und Fluechtlinge in der Bundesreublik Deutschland 1946-1953," in "Statistik der Bundesrepublik Deutschland," Vol. 114 Kohlhammer Stuttgart 1955.

Walsh, Conor, "Entrepreneurship and Local Growth," 2019. Working Paper.

7 Appendix: Theoretical Results

This section derives the main results of the model presented in Section 4.

7.1 Equilibrium where the free entry condition (17) might be slack

In the main body of the text I focused on the case where the free entry condition (17) holds with equality. This might not necessarily be the case. In particular, if $w_{rMt}h_{rt}^{E} > \pi_{irt}$ the net entry rate will be negative, i.e. $N_{rt} = (1-\delta)N_{rt-1}$. Equation (17) implies that this is the case if

$$w_{rMt}f_E\left((1-\delta)N_{rt-1}\right)^{-\lambda} > \frac{1-\beta}{\rho}\frac{P_{rMt}Y_{rMt}}{N_{rt}}.$$

Rearranging terms and substituting $w_{rMt}H_{rPt} = \beta P_{rMt}Y_{rMt}$ yields

$$f_E\left((1-\delta)N_{rt-1}\right)^{-\lambda} > \frac{1-\beta}{\rho\beta}\frac{H_{rPt}}{N_{rt}} = \frac{1-\beta}{\rho\beta}\frac{H_{rPt}}{(1-\delta)N_{rt-1}}.$$

Hence, there is no gross entry if and only if

$$f_E\left((1-\delta)N_{rt-1}\right)^{1-\lambda} > \frac{1-\beta}{\rho\beta}H_{rPt}.$$

From (15) it is still the case that $H_{rXt} = \frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right)H_{rPt}$. The labor market clearing condition in (22) then implies that

$$H_{rMt} = H_{rPt} + \frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right)H_{rPt} = \left(1+\frac{(\rho-1)(1-\beta)}{\rho\beta}\right)H_{rPt} = \frac{\rho-(1-\beta)}{\rho\beta}H_{rPt}.$$

Hence, gross investment will be zero if

$$f_E\left((1-\delta)N_{rt-1}\right)^{1-\lambda} > \frac{1-\beta}{\rho-(1-\beta)}H_{rMt}. \tag{42}$$

Hence, the sectoral labor allocations are:

- if (42) is satisfied, we have

$$\begin{aligned} H_{rPt} &= \frac{\rho\beta}{\rho-(1-\beta)}H_{rMt} \\ N_{rt} &= (1-\delta)N_{t-1} \end{aligned}$$

- If (42) is violated, there is entry in equilibrium. The labor market clearing condition in (22) therefore requires that

$$H_{rPt} = \beta\left(H_{rMt} + f_E\left((1-\delta)N_{rt-1}\right)^{1-\lambda}\right).$$

  The number of plants is given in (18) as

$$N_{rt} = \frac{1-\beta}{\rho\beta}\frac{(1-\delta)^{\lambda}}{f_E}H_{rPt}N_{rt-1}^{\lambda}.$$

7.2 Deriving the linear equilibrium system

The SBGP is determined from

$$\frac{Q_{rt}}{Q_{jt}} = \left(\frac{L_r}{L_j}\right)^{\frac{1}{1+\nu}}. \tag{43}$$

$$\frac{L_r}{L_j} = \left(\frac{V_r w_{rt}}{V_j w_{jt}}\right)^{\varepsilon}. \tag{44}$$

and

$$\frac{w_{rt} L_r}{w_{jt} L_j} = \left(\frac{A_r}{A_j}\right)^{\sigma-1} \left(\frac{Q_r}{Q_j}\right)^{(\sigma-1)\zeta\vartheta} \left(\frac{w_{rt}}{w_{jt}}\right)^{1-\sigma}. \tag{45}$$

Let express everything relative to an arbitrarily selected region 1, i.e. $j = 1$. Furthermore, denote for any variable X

$$x_r \equiv \ln \frac{X_r}{X_1}.$$

Equations (43)-(45) then imply

$$\begin{aligned} q_r &= \frac{1}{1+\nu} l_r \\ l_r &= \varepsilon v_r + \varepsilon w_r \\ l_r &= (\sigma - 1)\, a_r + (\sigma - 1)\, \zeta\vartheta q_r - \sigma w_r \end{aligned}$$

Solving for the relative population size, l_r, I get that

$$\begin{aligned} l_r &= (\sigma - 1)\, a_r + \frac{(\sigma - 1)\, \zeta\vartheta}{1+\nu} l_r - \frac{\sigma}{\varepsilon} l_r + \sigma v_r \\ &= (\sigma - 1)\, a_r + \sigma v_r + \frac{(\sigma - 1)\, \zeta\vartheta}{1+\nu} l_r - \frac{\sigma}{\varepsilon} l_r \\ &= \frac{1}{1 + \frac{\sigma}{\varepsilon} - \frac{(\sigma-1)\zeta\vartheta}{1+\nu}} \left[(\sigma - 1)\, a_r + \sigma v_r\right]. \end{aligned}$$

For the system to be well behaved I assume that

$$1 + \frac{\sigma}{\varepsilon} - \frac{(\sigma - 1)\, \zeta\vartheta}{1+\nu} = 1 + \frac{\sigma}{\varepsilon} - \left(\frac{\sigma - 1}{\rho - 1}\right) \frac{1-\beta}{1+\nu} > 0. \tag{46}$$

Note that this is a restriction - as $\nu \to -1$, this equation will not hold. Also note that the exogenous case of a standard Armington model is nested as the case $\beta = 1$. Using (46) we get that

$$l_r = \frac{(\sigma - 1)}{1 + \frac{\sigma}{\varepsilon} - \left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}} \times a_r + \frac{\sigma}{1 + \frac{\sigma}{\varepsilon} - \left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}} \times v_r.$$

Then we can solve for relative wages as

$$\begin{aligned} w_r &= \frac{1}{\varepsilon} l_r - v_r \\ &= \frac{1}{\varepsilon}\left(\frac{(\sigma-1)}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \frac{\sigma}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times v_r\right) - v_r \\ &= \frac{\frac{1}{\varepsilon}(\sigma-1)}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \frac{\frac{1}{\varepsilon}\sigma}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times v_r - v_r \\ &= \frac{\frac{1}{\varepsilon}(\sigma-1)}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \left(\frac{\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}-1}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\right)\times v_r. \end{aligned}$$

Finally, note that total factor productivity in region r is given by

$$TFP_{rt} = A_{rt} Q_{rt}^{\frac{1-\beta}{\rho-1}}.$$

Hence,

$$\begin{aligned} tfp_r &= a_r + \frac{1-\beta}{\rho-1} q_r \\ &= a_r + \frac{1-\beta}{\rho-1}\frac{1}{1+\nu} l_r \\ &= a_r + \frac{1-\beta}{\rho-1}\frac{1}{1+\nu}\left(\frac{(\sigma-1)}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \frac{\sigma}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times v_r\right) \\ &= a_r + \left(\frac{\frac{\sigma-1}{\rho-1}\frac{1-\beta}{1+\nu}}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \frac{\frac{1-\beta}{\rho-1}\frac{\sigma}{1+\nu}}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times v_r\right) \\ &= \frac{1+\frac{\sigma}{\varepsilon}}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times a_r + \frac{\frac{1-\beta}{\rho-1}\frac{\sigma}{1+\nu}}{1+\frac{\sigma}{\varepsilon}-\left(\frac{\sigma-1}{\rho-1}\right)\frac{1-\beta}{1+\nu}}\times v_r. \end{aligned}$$

7.3 Sectoral Labor Supply and Earnings

In this section we derive some convenient properties of the selection model. Recall that the distribution of individual skills is given by $F_s(z) = e^{-\phi_s z^{-\theta}}$, where ϕ_j parametrizes the average level of productivity of individuals in sector s and θ the dispersion of skills. The following result will turn out to be useful

Lemma 3. *Let* $\{x_i\}_{i=1}^n$ *be distributed iid according to*

$$F_{x_i}(x) = e^{-A_i x^{-\theta}}.$$

Then

$$E\left[x_i | x_i = \max_i [x_i]\right] = \Gamma\left(1-\frac{1}{\theta}\right)\left(\sum_{i=1}^n A_i\right)^{1/\theta}. \tag{47}$$

Note that this object does not depend on i.

Proof. Suppose that $i=1$ and let us derive the conditional distribution of x_1, conditional that x_1 is the highest

$\{x_j\}_j$. The joint distribution of $\{x_j\}_j$ is given by

$$F\left(x_1, x_2, ..., x_3\right) = \prod_{j=1}^{n} F\left(x_i\right) \tag{48}$$

because of independence. Hence, we get that

$$\begin{aligned}
P\left(x_1 \leq m | x_1 = \max_j \left[x_j\right]\right) &= \frac{1}{P\left(x_1 = \max_j \left[x_j\right]\right)} \int_0^m \prod_{j=2}^{n} P\left(x_j < x\right) dF_{x_1}\left(x\right) \\
&= \frac{1}{P\left(x_1 = \max_j \left[x_j\right]\right)} \int_0^m \prod_{j=2}^{n} e^{-A_j x^{-\theta}} A_1 \theta x^{-\theta-1} e^{-A_1 x^{-\theta}} dx \\
&= \frac{A_1}{P\left(x_1 = \max_j \left[x_j\right]\right)} \int_0^m \theta x^{-\theta-1} \prod_{j=1}^{n} e^{-A_j x^{-\theta}} dx \\
&= \frac{A_1}{P\left(x_1 = \max_j \left[x_j\right]\right)} \int_0^m \theta x^{-\theta-1} e^{-\left(\sum_j A_j\right) x^{-\theta}} dx.
\end{aligned}$$

Now let us derive $P\left(x_1 = \max_j \left[x_j\right]\right)$. We get that

$$P\left(x_1 = \max_j \left[x_j\right]\right) = P\left(x_j \leq x_1 \text{ for all } j \geq 2\right) = \int_{x_1} \frac{\partial F}{\partial x_1}\left(x_1, x_1, x_1, ..\right) dx_1.$$

Using (48), we get that

$$\begin{aligned}
\frac{\partial F\left(x_1, x_2, ..., x_3\right)}{\partial x_1} &= \frac{\partial F\left(x_1\right)}{\partial x_1} \left(\prod_{j=2}^{n} F\left(x_i\right)\right) = e^{-A_1 x^{-\theta}} A_1 \theta x_1^{-\theta-1} \left(\prod_{j=2}^{n} e^{-A_i x_j^{-\theta}}\right) \\
&= A_1 \theta x_1^{-\theta-1} \left(\prod_{j=1}^{n} e^{-A_i x_i^{-\theta}}\right) \\
&= A_1 \theta x_1^{-\theta-1} e^{-\sum_i^n A_i x_i^{-\theta}}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\int_{x_1} \frac{\partial F}{\partial x_1}\left(x_1, x_1, x_1, ..\right) dx_1 &= \int_{x_1} A_1 \theta x_1^{-\theta-1} e^{-\left(\sum_i^n A_i\right) x_1^{-\theta}} dx_1 \\
&= \frac{A_1}{\sum_i^n A_i} \int_{x_1} \frac{\theta}{\left(\sum_i^n A_i\right)^{1/\theta}} \left(\frac{x_1}{\left(\sum_i^n A_i\right)^{1/\theta}}\right)^{-\theta-1} e^{-\left(\frac{x_1}{\left(\sum_i^n A_i\right)^{1/\theta}}\right)^{-\theta}} dx_1 \\
&= \frac{A_1}{\sum_i^n A_i}.
\end{aligned}$$

Substituting this above yields

$$\begin{aligned}
P\left(x_1 \leq m | x_1 = \max_j \left[x_j\right]\right) &= \left(\sum_i^n A_i\right) \int_0^m \theta x^{-\theta-1} e^{-\left(\sum_j A_j\right) x^{-\theta}} dx \\
&= \int_0^m \frac{\theta}{\kappa} \left(\frac{x}{\kappa}\right)^{-\theta-1} e^{-\left(\frac{x}{\kappa}\right)^{-\theta}} dx,
\end{aligned}$$

where $\kappa = \left(\sum_j A_j\right)^{1/\theta}$. This is a Frechet distribution with shape θ and scale κ, i.e.

$$F_{x_1|x_1=\max_j[x_j]}(m) = e^{-\left(\frac{x}{\left(\sum_j A_j\right)^{1/\theta}}\right)^{-\theta}} = e^{-\left(\sum_j A_j\right)x^{-\theta}}.$$

This implies (47) □

Lemma 3 is useful because it allows us to calculate sectoral earnings and the sectoral supply of efficiency units. Consider first sectoral earnings. Let $y_{rs}^i = w_{rs} z_s^i$ be the earnings of individual i in region r working in sector s. The distribution of earnings is

$$P\left(y_{sr}^i < y\right) = P\left(z_s^i < \frac{y}{w_{rs}}\right) = e^{-\phi_s\left(\frac{y}{w_{rs}}\right)^{-\theta}} = e^{-\phi_s w_{rs}^{\theta} y^{-\theta}}.$$

Hence, Lemma 3 implies that

$$E\left[y_{rs}^i | y_{rs}^i = \max_s \left\{y_{rs}^i\right\}\right] = \Gamma\left(1 - \frac{1}{\theta}\right)\left(\sum_s \phi_s w_{rs}^{\theta}\right)^{1/\theta}.$$

Similarly, the average labor supply in sector s is given by

$$\begin{aligned} E\left[z_s^i | y_{rs}^i = \max_s \left\{y_{rs}^i\right\}\right] &= E\left[z_s^i | w_{rs} z_s^i = \max_s \left\{w_{rs} z_s^i\right\}\right] \\ &= E\left[z_s^i | z_s^i = \max_k \left\{\frac{w_{rk}}{w_{rs}} z_k^i\right\}\right] \\ &= \Gamma\left(1 - \frac{1}{\theta}\right)\left(\phi_s + \sum_{k \neq s} \phi_k \left(\frac{w_{rk}}{w_{rs}}\right)^{\theta}\right)^{1/\theta} \\ &= \Gamma\left(1 - \frac{1}{\zeta}\right)\frac{1}{w_{rs}}\left(\sum_s \phi_s w_{rs}^{\theta}\right)^{1/\theta}. \end{aligned}$$

Also, the share of people working in sector s is given by

$$\pi_{rs} = P\left(y_{rs}^i = \max_k \left\{y_{rk}^i\right\}\right) = \frac{\phi_s w_{rs}^{\theta}}{\sum_s \phi_s w_{rs}^{\theta}} = \phi_s\left(\frac{w_{rs}}{\left(\sum_s \phi_s w_{rs}^{\theta}\right)^{1/\theta}}\right)^{\theta}.$$

It is useful to define the endogenous scalar of *average earnings in region r as*

$$\overline{w}_r = \left(\sum_s \phi_s w_{rs}^{\theta}\right)^{1/\theta}.$$

Then we can write the sectoral employment share π_{rs} as

$$\pi_{rs} = \phi_s\left(\frac{w_{rs}}{\overline{w}_r}\right)^{\theta},$$

and the aggregate amount of sectoral earnings $w_{rs}H_{rs}$ as

$$\begin{aligned} w_{rs}H_{rs} &= L_r P\left(y^i_{rs} = \max_k \left\{y^i_{rk}\right\}\right) E\left[y^i_{rs}|y^i_{rs} = \max_s \left\{y^i_{rs}\right\}\right] \\ &= L_r \Gamma\left(1-\frac{1}{\theta}\right)\left(\pi_{rs}\overline{w}_r\right). \end{aligned}$$

Note that we can write

$$\pi_{rs}\overline{w}_r = \phi_s w^{\theta}_{rs}\left(\overline{w}_r\right)^{1-\theta}.$$

Hence,

$$w_{rs}H^h_{rs} = L_r\Gamma\left(1-\frac{1}{\theta}\right) w^{\theta}_{rs}\left(\phi_s\left(\overline{w}_r\right)^{1-\theta}\right).$$

Note that the aggregate level of aggregate human capital of individuals working in sector s in region r is given by

$$\begin{aligned} H_{rs} &= L_r\Gamma\left(1-\frac{1}{\theta}\right) w^{\theta-1}_{rs}\left(\phi_s\left(\overline{w}_r\right)^{1-\theta}\right) = L_r\Gamma\left(1-\frac{1}{\theta}\right)\left(\phi_s\left(\frac{w_{rs}}{\overline{w}_r}\right)^{\theta-1}\right) \\ &= L_r\Gamma\left(1-\frac{1}{\theta}\right)\left(\phi_s\left(\frac{\pi_{rs}}{\phi_s}\right)^{\frac{\theta-1}{\theta}}\right) = L_r\Gamma\left(1-\frac{1}{\theta}\right)\left(\left(\phi_s\right)^{\frac{1}{\theta}}\pi_{rs}{}^{\frac{\theta-1}{\theta}}\right). \end{aligned}$$

This also implies that *aggregate* earnings in region r, Y_r, are given by

$$Y_r = \left[\sum_s w_{rs}H_{rs}\right] = L_r\Gamma\left(1-\frac{1}{\theta}\right)\left(\overline{w}_r\right).$$

7.4 Static Allocations

Consider the manufacturing sector in region r. Let P be the price for the aggregate good and w the wage rate. The profits of the final good firm are therefore given by

$$\Pi = \max_{[x_i],H,T} PQX^{1-\beta}H^{\beta} - wH - \int_0^N p_i x_i di,$$

where X is given in (7). As usual we can solve this problem sequentially.

Optimal spending on $[x_i]$ Consider the problem

$$\max_{[x_i]}\left(\int_0^N x_{it}^{\frac{\rho-1}{\rho}} di\right)^{\frac{\rho}{\rho-1}} \quad \text{s.t.} \quad \int_0^N p_i x_i di \leq M.$$

Standard arguments imply that prices, sales and profits of inputs producers are given

$$p_i = \frac{\rho}{\rho-1}w \qquad \pi_i = \frac{1}{\rho}P_X X\frac{1}{N}$$

where the price index for the aggregate X, P_X, is given by

$$P_X = \frac{\rho}{\rho-1}w \times N^{-\frac{1}{\rho-1}}. \tag{49}$$

Given (49), the optimal quantity of X and H solve the problem

$$\Pi = \max_{H,T} \; PQX^{1-\beta}H^{\beta} - wH - P_X X.$$

Hence,

$$\begin{aligned} P_X X &= (1-\beta)\, PY \\ wH &= \beta PY \end{aligned}$$

The aggregate price is given by

$$P = \frac{1}{Q}\left(\frac{P_X}{1-\beta}\right)^{1-\beta}\left(\frac{w}{\beta}\right)^{\beta} = \frac{1}{Q}\left(\frac{1}{1-\beta}\right)^{1-\beta}\left(\frac{1}{\beta}\right)^{\beta}\left(\frac{\rho}{\rho-1}\right)^{1-\beta} wN^{-\frac{1-\beta}{\rho-1}}.$$

Similarly,

$$Y = \frac{wH}{\beta P} = N^{\frac{1-\beta}{\rho-1}} Q \frac{H}{\left(\frac{\beta}{1-\beta}\frac{\rho}{\rho-1}\right)^{1-\beta}}$$

Relative Employment Relative earnings for workers in the intermediate goods sector are given

$$\begin{aligned} wH_X &= w\int_i^N h_i di = \frac{\rho-1}{\rho} P_X X \\ &= \frac{\rho-1}{\rho}(1-\beta)\, PY \\ &= \frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right) wH_P. \end{aligned}$$

Hence,

$$H_X = \frac{\rho-1}{\rho}\frac{1-\beta}{\beta} H_P.$$

7.5 Proof of Proposition 2

To see that equations (23)-(28) fully characterize the equilibrium, let $[N_{rt-1}]_r$ be given. Let $\left[w_{rAt}, w_{rMt}, L^I_{rt}, L^F_{rt}\right]_r$ be given. From (25) and (26) we can directly calculate $[H_{rAt}, H_{rPt}]_r$. From (28) this implies $[N_{rt}]_r$. Given $[H_{rAt}, H_{rPt}, N_{rt}]_r$, (23) and (24) are $2 \times R$ equations in $2 \times R$ unknowns, which can be used to find $[w_{rAt}, w_{rMt}]_r$. Note that (23) and (24) are homogeneous of degree zero. To determine the level of $[w_{rAt}, w_{rMt}]_r$, we therefore need to impose a numeraire. Finally, (27) are $2 \times R$ equations in $2 \times R$ unkowns, which can be used to find $\left[L^I_{rt}, L^F_{rt}\right]_r$.

7.6 Persistence of Population Shocks

Equation (30) implies that the change in productivity N_{rt} in response to a change in the manufacturing workforce $\{d\ln H_{rPj}\}_{j=t_0}^{t_o+t}$ is given by

$$d\ln N_{r\tau} = \sum_{j=t_0}^{t_0+\tau} \lambda^{-(j-(t_0+\tau))} d\ln H_{rPj}.$$

Assume that the initial shock at t_0 dies out slowly, i.e.

$$d \ln H_{rPj} = d \ln H_{rPt_0} \times p^{j-t_0}$$

where $p < 1$.Hence if $p = 0$ then $d \ln H_{rPj} = 0$ for $j > t_0$. This implies that

$$\begin{aligned}
\sum_{j=t_0}^{t_0+\tau} \lambda^{-(j-(t_0+\tau))} d \ln H_{rPj} &= \sum_{j=t_0}^{t_0+\tau} \lambda^{-(j-(t_0+\tau))} d \ln H_{rPt_0} \times p^{j-t_0} \\
&= d \ln H_{rPt_0} \left(\sum_{j=t_0}^{t_0+\tau} \lambda^{-(j-(t_0+\tau))} p^{j-t_0} \right) \\
&= d \ln H_{rPt_0} \left(\lambda^{\tau} \sum_{j=t_0}^{t_0+\tau} \left(\frac{\lambda}{p} \right)^{-(j-t_0)} \right) \\
&= d \ln H_{rPt_0} \lambda^{\tau} \frac{1 - \left(\frac{p}{\lambda}\right)^{\tau+1}}{1 - \frac{p}{\lambda}} \\
&= d \ln H_{rPt_0} \frac{\lambda^{\tau+1} - p^{\tau+1}}{\lambda - p}.
\end{aligned}$$

Hence,

$$\Psi_{\tau}(p, \lambda) \equiv \frac{d \ln N_{r\tau}}{d \ln H_{rPt_0}} = \frac{\lambda^{\tau+1} - p^{\tau+1}}{\lambda - p}.$$

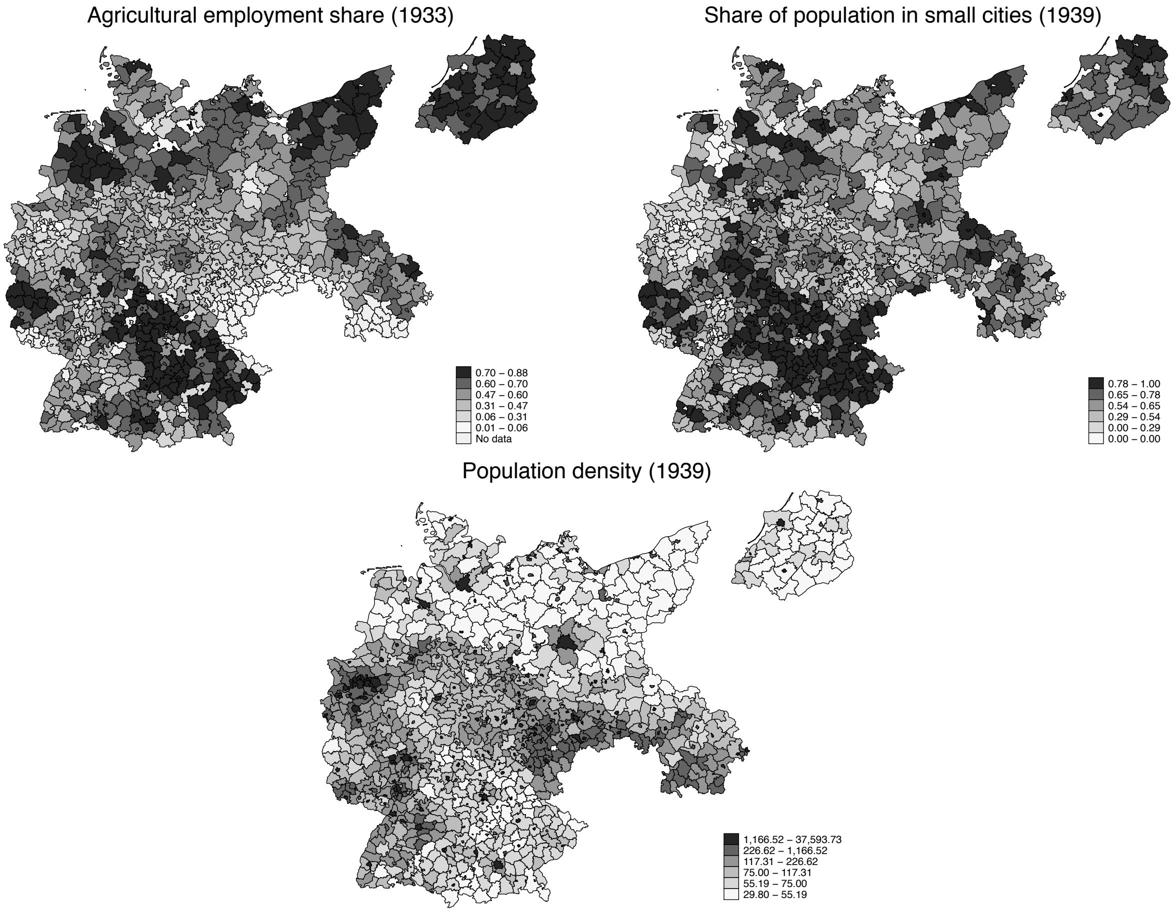
If $p = 0$ we that $\Psi_{\tau}(0, \lambda) = \lambda^{\tau} \to 0$. If $p = 1$, then

$$\Psi_{\tau}(1, \lambda) = \frac{1 - \lambda^{\tau+1}}{1 - \lambda} \to \frac{1}{1 - \lambda}.$$

Note also that for $p < 1$ we always have

$$\lim_{\tau \to \infty} \Psi_{\tau}(p, \lambda) = 0$$

Note also that $\varUpsilon_{\tau}$ is increasing in λ and p.



Notes: The map on the left shows the agricultural employment share in 1933. The map on the right shows the share of the county population that lives in cities with less than 2000 inhabitants. The map at the bottom shows the population density in 1939.

Figure 13: Economic Geography in the Pre-War Period

8 Appendix: Empirical Results

8.1 Additional Empirical Results

In this section I report additional empirical results.

Germans in Eastern and Middle Europe before 1939

In Figure 13 I display three aspects of the spatial distribution of economic activity in the pre-war period: the agricultural employment share in 1933, a measure of the rural population in 1939 (as measured by the share of the county population that lives in cities with less than 2000 inhabitants) and the population density in 1939. The maps in Figure 13 show that the German Territories of Eastern Europe were more agricultural and rural than Western Germany prior to the war. Finally, in Table 10 I compare the population in West Germany and the Eastern Territories according to their occupational and educational characteristics. The distribution of formal skills was very similar. The only slight difference is the higher popularity of vocational schools (*Berufsschule*) in pre-war Western Germany, a fact that is due to the bigger importance of the manufacturing industry. Similarly, the occupational distribution is also very similar in the pre-war period.

	West Germany	Eastern Territories
	Educational Attainment	
Elementary School	66.3	65.9
High School	8.3	11
Vocational School	18.4	15.5
College	6.8	7.6
	Occupational composition	
Self-employed (Agriculture)	10.3	12.3
Skilled Employee	7.7	8.4
Unskilled Employee	7.9	8.5
Skilled Worker	3.6	2.9
Unskilled Worker	23.8	21.8

Notes: This table reports the educational, sectoral and occupational distribution in West Germany and the Eastern Territories of the German Reich in 1939.

Table 10: Economic Characteristics in 1939

	Share of refugees in 1950					
State	Mean	P10	P25	P50	P75	P90
Baden_Wuerttemberg	0.143	0.058	0.082	0.118	0.222	0.245
Bavaria	0.229	0.158	0.194	0.235	0.265	0.294
Hesse	0.182	0.099	0.152	0.191	0.220	0.244
NRW	0.113	0.057	0.067	0.114	0.143	0.177
Niedersachsen	0.275	0.161	0.213	0.290	0.345	0.360
RhPf	0.052	0.032	0.040	0.050	0.063	0.071
SH	0.336	0.259	0.312	0.351	0.386	0.391
Western Germany	0.189	0.060	0.107	0.194	0.255	0.312

Notes: The table report moments of the distribution of refugee shares within the different states in Western Germany.

Table 11: Distribution of Refugees Across States in Western Germany

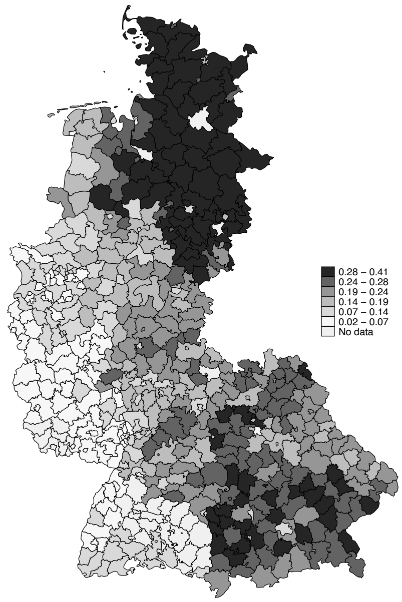
Allocation of refugees in Western Germany

In Figure 14 I display the share of refugees (relative to the county population) for all counties in Western Germany in 1950. As also seen in Figures 3 and 4, for many counties the initial shock is very large and there is a clear East-West trajectory. There are also two important "centers" of refugee destinations in the north (in the states Schleswig-Holstein and Lower Saxony) and in the South (in Bavaria). Again, this has geographical reasons. Many expellees from the Easter Territories arrived in Western Germany via a northern route along the coast of the Baltic Sea and hence arrived in Western Germany in the north. Similarly, the expellees from the southern parts of Eastern Europe (most importantly the Sudeten) arrived in Bavaria and therefore settled there.

These differences across states are also seen in Table 11, which reports the distribution of refugees across states. Because the entire analysis is conducted conditional on state fixed effects, the variation across states is not used for identification.

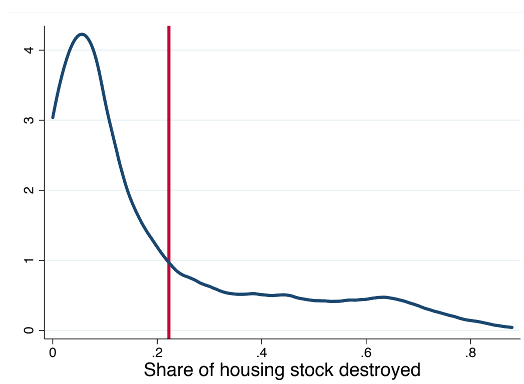
Wartime destruction

In Figure 15 I depict the cross-sectional distribution of war-time destruction, i.e. the share of the housing stock, which was damaged in the war. It is clearly seen that there are many counties, where more than 60% of their housing stock was damaged during the war.



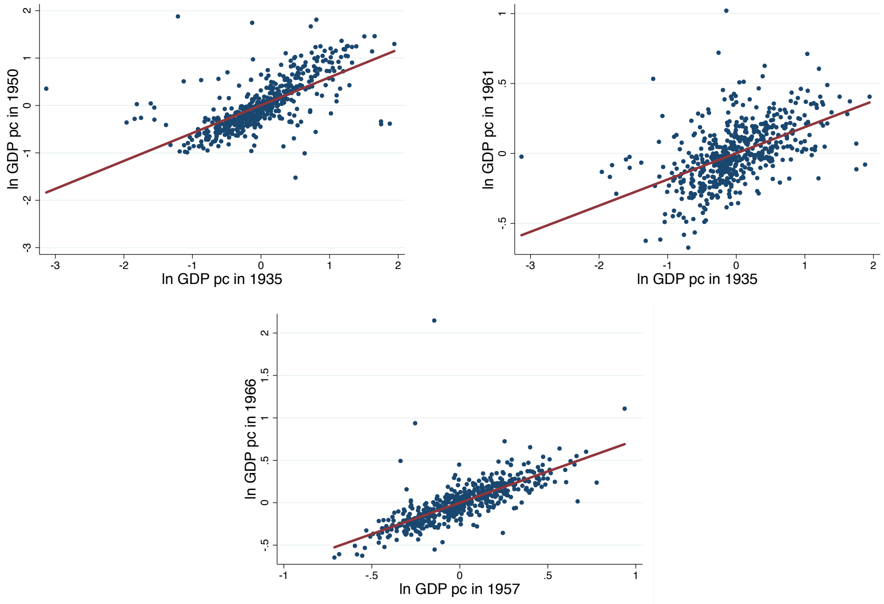
Notes: This figure depicts the share of refugees (relative to the entire population) for each county of Western Germany in 1950. Counties are harmonized at the level of 1975. Source: Statistisches Bundesamt (1955)

Figure 14: The Allocation of Refugees in Wester Germany: 1950



Notes: The figure shows the distribution of war-time destruction across all Western German counties. War time destruction is measured as the share of the housing stock that was destroyed during the war. The data is drawn from the 1950 housing census.

Figure 15: The Distribution of wartime destruction



Notes:

Figure 16: The correlation of the different measures of GDP

	(1) ln pop dens. 1939	(2) Pop share in small cities 1939	(3) Manufacturing share in 1939	(4) Manufacturing share in
Share of refugees in 1950	-10.678***	2.491***	-0.504***	-0.559***
	(0.845)	(0.205)	(0.085)	(0.098)
State FE	Yes	Yes	Yes	Yes
Observations	536	536	535	524
R^2	0.345	0.326	0.277	0.264

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

Table 12: Corollaries of the Initial Allocation of Refugees

Measures of GDP

As explained in Section 3.1, I use data on regional value added taxes to measure income per capita in 1935 and 1950. For the years 1957-1966 I have information on income per capita at the county-level. In Figure 16 I show the within-state correlation of the different measures in 1935 and 1950 (Panel A), 1935 and 1961 (Panel B) and 1957 and 1966 (Panel C). Reassuringly there is a strong positive correlation.

Initial Allocation of refugees

8.2 Robustness of Results of Section 3

In this section I provide additional results for the empirical results reported in Section 3.

	Manufacturing share in 1950								
Share of refugees in 1950	0.315***	0.276***	0.315***	0.292***	0.265***	0.309***	0.274***		0.243***
	(0.047)	(0.045)	(0.077)	(0.053)	(0.046)	(0.047)	(0.049)		(0.059)
Share of refugees in 1946								0.218***	
								(0.068)	
Manufac. share 1939	✓	✓	✓	✓	✓	✓	✓	✓	✓
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls 1933					✓				
Labor Supply		✓							
Construction		✓							
City Dummy						✓			
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Clustered Std errors			✓						
Pop weights				✓					
Excluding Bavaria							✓		
Observations	535	535	535	535	523	488	344	403	403
R^2	0.896	0.908	0.896	0.914	0.912	0.899	0.930	0.884	0.885

Table 13: Refugees and Manufacturing Employment in 1950: Robustness

Manufacturing Employment and Manufacturing Plants

Table 4 showed that refugee inflows were positive related to subsequent manufacturing employment. In this Section I report additional robustness results. These are contained in Table 13 for 1950 and in Table 14 for 1961. The tables show that the results reported in Table 2 are robust to explicitly controlling for regional variation in labor supply as proxied by the employment share and the share of male inhabitants (column 2), to clustering the standard errors at the state level (column 3), to weighing counties by their population in 1939 (column 4), by controlling for population density and sectoral employment shares in 1933 (column 5), by explicitly incorporating fixed effects at the city level (column 6) and by including the state of Bavaria, where counties are decisively smaller as in other states (column 7). In column 8 I replicate the results based on the share of refugees in 1946. Finally, in column 9 I use the same counties as in specification 8 and show that the coefficient is very similar.

In Table 7 I showed a positive relationship between the inflow of the refugees and the growth of manufacturing plants. In Tables 15 and 16 I show that these results are broadly robust to a variety of controls and specifications. When I restrict the sample by either excluding the state of Bavaria or by focusing only the set of counties where the refugee share in 1946 is observed, the coefficients are no longer significant.

GDPpc

Tables 17 and 18 report a variety of robustness checks for the results reported in Table 5 in the main text. In Table 17 I focus on GDPpc in 1950 as the dependent variable, Table 18 covers the case of 1957 - 1966. For brevity I only report the pooled specification - the results for the individual regressions for the years 1957, 1961, 1964 and 1966 are available upon request. Tables 17 and 18 confirm the results reported in Table 5 qualitatively and quantitatively. In particular, I report specifications with clustered standard errors (at the state level), there the individual observation are weighted with the 1939 county population, where I drop the state of Bavaria, where I control for a fixed effect for large city states and when I use the share of refugees in 1946 as the dependent variable. Except for the case of using the 1946 refugee share for GDP pc in the long-run, all results mirror the ones reported in Table 5.

	Manufacturing share in 1961								
Share of refugees in 1950	0.249***	0.258***	0.249***	0.177***	0.196***	0.276***	0.255***		0.209***
	(0.058)	(0.058)	(0.049)	(0.058)	(0.061)	(0.060)	(0.062)		(0.075)
Share of refugees in 1946								0.163**	
								(0.081)	
Manufac. share 1939	✓	✓	✓	✓	✓	✓	✓	✓	✓
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓	✓	✓
Controls 1933					✓				
Labor Supply		✓							
Construction		✓							
City Dummy						✓			
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Clustered Std errors			✓						
Pop weights				✓					
Excluding Bavaria							✓		
Observations	535	535	535	535	523	488	344	403	403
R^2	0.896	0.908	0.896	0.914	0.912	0.899	0.930	0.884	0.885

Table 14: Refugees and Manufacturing Employment in 1961: Robustness

	ln num of manufac. plants 1950						
Share of refugees in 1950	0.407**	0.407	0.510**	0.660***	-0.076		0.294
	(0.203)	(0.235)	(0.243)	(0.196)	(0.259)		(0.236)
Share of refugees in 1946						-0.257	
						(0.283)	
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓
ln plants in 1939	✓	✓	✓	✓	✓	✓	✓
Manufac. share in 1933				✓			
Manufac. share in 1939	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓
Controls 1933				✓			
State FE	✓	✓	✓	✓	✓	✓	✓
Clustered Std errors		✓					
Pop weights			✓				
Excluding Bavaria					✓		
Observations	519	519	519	505	340	390	390
R^2	0.864	0.864	0.927	0.888	0.847	0.876	0.876

Table 15: Refugees and Manufacturing Plants in 1950: Robustness

	ln num of manufac. plants 1956						
Share of refugees in 1950	0.873**	0.873*	0.198	1.007**	0.196		0.780
	(0.434)	(0.448)	(0.360)	(0.436)	(0.490)		(0.548)
Share of refugees in 1946						-0.072	
						(0.540)	
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓
ln plants in 1939	✓	✓	✓	✓	✓	✓	✓
Manufac. share in 1933				✓			
Manufac. share in 1939	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓
Controls 1933				✓			
State FE	✓	✓	✓	✓	✓	✓	✓
Clustered Std errors		✓					
Pop weights			✓				
Excluding Bavaria					✓		
Observations	519	519	519	505	340	390	390
R^2	0.770	0.770	0.892	0.785	0.737	0.769	0.771

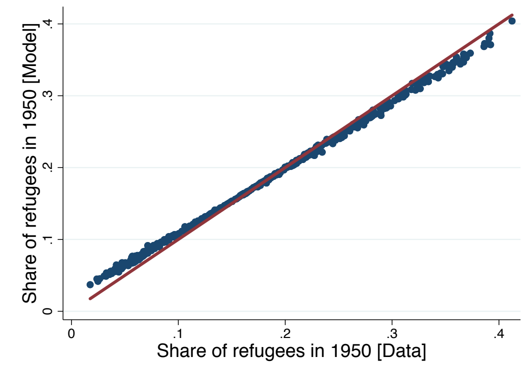
Table 16: Refugees and Manufacturing Plants in 1956: Robustness

	ln GDPpc in 1950							
Share of refugees in 1950	0.178	0.178	0.260	0.075	0.119		0.237	-0.344
	(0.322)	(0.436)	(0.418)	(0.436)	(0.311)		(0.371)	(0.240)
Share of refugees in 1946						0.209		
						(0.377)		
ln GDPpc in 1935	✓	✓	✓	✓	✓	✓	✓	✓
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓	✓
Industry Controls	✓	✓	✓	✓	✓	✓	✓	✓
City Dummy					✓			
Clustered Std error		✓						
Population weighted			✓					
Excluding Bavaria				✓				
Dropping outliers								✓
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	523	523	523	337	476	396	396	492
R^2	0.702	0.702	0.743	0.638	0.742	0.723	0.723	0.858

Table 17: Refugees and GDPpc in 1950: Robustness

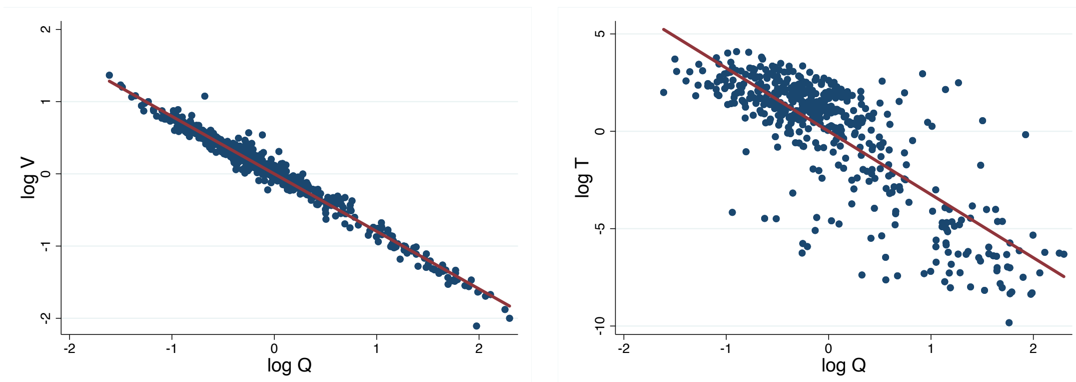
	log GDP per worker in 1957-1966						
Share of refugees in 1950	0.617***	0.617***	0.387*	0.415***	0.655***		0.712***
	(0.100)	(0.150)	(0.207)	(0.132)	(0.110)		(0.133)
Share of refugees in 1946						0.159	
						(0.125)	
ln GDPpc in 1935	✓	✓	✓	✓	✓	✓	✓
Year	✓	✓	✓	✓	✓	✓	✓
ln pop dens 1939	✓	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓	✓
Industry Controls	✓	✓	✓	✓	✓	✓	✓
City Dummy					✓		
Clustered Std error		✓					
Population weighted			✓				
Excluding Bavaria				✓			
State FE	✓	✓	✓	✓	✓	✓	✓
Observations	2076	2076	2076	1332	1888	1568	1568
R^2	0.750	0.750	0.647	0.731	0.751	0.721	0.727

Table 18: Refugees and GDPpc in 1957-1966: Robustness



Notes: The figure shows a scatter plot of the share of refugees in 1950 in the data (x axis) against the model (y axis).

Figure 17: The Share of Refugees in 1950: Model vs Data



Notes: The left (right) panel shows the correlation between innate productivity $\ln Q$ and amenities $\ln V$ (land endowments $\ln T$) along the spatial BGP in 1935.

Figure 18: Correlation of Spatial Fundamentals

9 Appendix: Quantitative Results

9.1 Details of the calibration

As explained in the main text, I model the inflow of refugees as a one-time-shock in 1947 and allocate them across counties according to the empirically observed share of refugees in 1950. This implies that the model-implied share of refugees does not perfectly coincide with the one in the data. However, as seen in Figure 17 where I plot the refugee share in the data against the refugee share in the model, the discrepancy is very small. In Figure 18 I show the correlation of spatial fundamentals (conditional on state FE).