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

Michael Peters*

December 2019

Abstract

Virtually all theories of economic growth predict a positive relationship between population size and productivity. In this paper I study a particular historical episode to provide direct evidence for the empirical relevance of such scale effects. In the aftermath of the Second World War about 8m ethnic Germans were expelled from their domiciles in Eastern Europe and transferred to West Germany. This inflow increased the German population by 20%. Using variation across counties I show that the settlement of refugees had a large and persistent effect on the size of the local population, that it increased manufacturing employment and that it raised income per capita, particularly in the long-run. I then show that these findings are quantitatively consistent with an idea-based model of spatial growth. The theory highlights that a shock to the local population can have persistent effects on local productivity, which are larger in the long-run. When calibrated to the cross-sectional estimates, the model implies that the refugee settlement increased aggregate income per capita by about 25% after 25 years.

*eMail: m.peters@yale.edu. I am very grateful to Julieta Caunedo, Penny Goldberg, Tarek Hassan, Chad Jones, Pete Klenow, Sam Kortum, 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 the Atlanta FED, ASU, Chicago, Columbia, Cornell, Harvard, LSE, MIT, NYU, Philadelphia FED, 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. Fil Babalievsky provided stellar research assistance.

1 Introduction

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

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

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

To estimate the cross-sectional relationship between refugee inflow and local economic development, I constructed a novel panel dataset for 500 West German counties between 1930 and 1970 from original historical sources. Two features of the refugee settlement allow me to use it as a shifter of local labor supply. First, the refugees were not free to settle in the location of their choice but the population transports were organized by the Military Governments of the US and the UK, the governing bodies of West Germany at the time. Second, the dominant consideration to allocate the inpouring refugees to particular regions was the availability of housing rather than future economic prospects. With millions of refugees being transferred, the population of West Germany in 1950 exceeded its pre-war level by about 20%. At the same time, the Allied bombing campaign had reduced the housing stock by almost 25% on average and in many cities by more than 75%. Hence, refugees were assigned to rural, low population density localities where housing was relatively abundant. These aspects of the historical setting allow me to tease out the exogenous component of the initial refugee allocation both by directly controlling for the determinants of the allocation rule and by using an instrumental variable strategy, which exploits the distance to the pre-war population centers in Eastern Europe.

My results imply a positive relationship between population size and local productivity. First, I show that the initial allocation of refugees was very persistent. By 1961, i.e. 10 - 15 years after the initial settlement, counties that received more refugees were still substantially larger and the share of refugees was still higher. Second, I establish a positive relationship between the allocation of refugees and manufacturing employment growth in the 1950s and 60s. Third, I document that the inflow of refugees raised local productivity. Importantly, such productivity gains accrued slowly over time: while the effect of refugee inflows on income per capita in 1950 is statistically indistinguishable from zero, it is positive and large in the late 50s and early 60s.

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

The model highlights an important distinction between the short-run and the long-run elasticity of productivity with respect to population size. The short-run elasticity describes the relationship between productivity and the local population holding current productivity constant. This elasticity depends on the elasticity of substitution across varieties, which parametrizes the strength of variety gains, and is isomorphic to agglomeration externalities commonly used in quantitative models of economic geography. In contrast, 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. The crucial parameter for the size of the long-run elasticity is the - what I call - inter-temporal knowledge elasticity, which determines how quickly the costs of creating new ideas decline in the existing stock of ideas. If this elasticity is positive, the long-run elasticity exceeds the short-run elasticity as the dynamic accumulation of local productivity acts as an amplifying force. If this elasticity exceeds unity, population shocks have permanent effects on local productivity and that the spatial distribution of economic activity is not stationary.

I then structurally estimate the model using the empirical variation from the natural experiment, i.e. by targeting the cross-county regression coefficients estimated in the first part of the paper. The main moments of interest are the relationship between refugee inflows and income per capita at different time horizons, the spatial persistence of the allocation of the refugee population, and the response of local manufacturing employment to the refugee settlement. It is the discrepancy between the short- and long-run scale elasticity that allows the model to reconcile the small effects of refugee inflows on income per capita in 1950 with the large effects in the 1960s.

My empirical estimates imply that moving frictions were substantial and that the dynamic amplification of the initial shock was powerful. The long-run scale elasticity is about 8 times as large as the short-run scale elasticity. At the same time, the estimated knowledge elasticity is small enough so that I can comfortably reject the case of non-stationarity. While this implies that a population shock will not affect the spatial distribution of economic activity in the very long-run, my estimates show that this long-run can indeed be very long: the effect of the refugee settlement on local income per capita is increasing during the 1950s and 1960s, peaks in the mid 70s, and then slowly reverts back to zero.

Finally I use the model to quantify the aggregate consequences of the refugee settlement. The combination of decreasing returns to scale in the agricultural sector and increasing returns to scale in the manufacturing sector imply that the effect is a priori ambiguous. It is also not identified from the cross-sectional estimates because of general equilibrium interactions, in particular the extent to which non-treated regions benefit indirectly through trade and migration linkages. The quantitative results reflect the importance of the dynamic amplification as the inflow of refugees reduced income per capita by about 5% in the short-run but increased it by about 25% after 25 years. Importantly, these effects are much larger than a naïve extrapolation of the cross-sectional estimates that does not take general equilibrium effects into account.

Related Literature The paper is related to a large literature in economic growth, which highlights the importance of market size effects (see, for example, the survey articles by Jones (2005) or Akcigit (2017)). Of particular relevance is the semi-endogenous growth model by Jones (1995), where the inter-temporal knowledge elasticity also takes center stage. While Jones (1995) highlights the importance of this elasticity to distinguish models of endogenous and semi-endogenous growth in the time series, my model stresses that this distinction has a clear spatial counterpart: the 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. Recent papers that focus on nexus between population and productivity growth are Jones (2019) and Peters and Walsh (2019).

The paper also contributes to a recent literature on dynamic models of trade and economic geography - see e.g. Desmet et al. (2018), Desmet and Rossi-Hansberg (2014), Nagy (2017) and Walsh (2019). These contributions are theoretically extremely elegant in that they deliver static policy functions even though all agents are forward looking. The necessary assumptions, however, imply that the local population is not a state variable of the equilibrium system. This makes these frameworks not well suited to study the consequences of an exogenous shock to the population. My model allows me to do so at the cost of assuming that agents are myopic. Expectedly, this simplifies the analysis considerably and allows me to treat the population as a slow-moving state-variable. This dynamic interaction between spatial mobility and local productivity is also studied in Allen and Donaldson (2018), albeit in a more reduced form way. A dynamic model of trade and migration is also analyzed in Caliendo et al. (2019), who however assume that regional productivity is exogenous.

There is also a close connection to a large and growing literature on models of economic geography, which often posit the existence of exogenous agglomeration economies - see for example Allen and Arkolakis (2014); Ahlfeldt et al. (2015); Fajgelbaum and Redding (2014); Ramondo et al. (2016); Eckert and Peters (2016); Faber and Gaubert (2019) or the recent survey by Redding and Rossi-Hansberg (2017). These reduced-form specifications imply that scale elasticities are stable and time-invariant. My theoretical and empirical results highlight that such scale elasticities are likely to be very different in the long and the short-run. This finding is reminiscent of the literature on directed technological change, which also stresses the difference between short- and long-run elasticities (Acemoglu, 2002, 2007).

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

Finally, various papers use the German context as a source of historical experiments.[1] Burchardi and Hassan (2013) use the fall of the Berlin Wall to measure the importance of social ties, Ahlfeldt et al. (2015) exploit the partition of Berlin to estimate agglomeration forces within city-blocks in Berlin and Redding and Sturm (2008) use the division of Germany

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

as a shift in market access. The post-war population expulsions, which are the focus of this paper, have also been analyzed in Braun and Mahmoud (2014) and Braun and Kvasnicka (2014). These contributions do not focus on the effect on local productivity.

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

2 The Historical Setting

The Presence of Ethnic Germans in Eastern Europe before 1939

The presence of ethnic Germans in Eastern Europe is by no means a novel phenomenon but dates back to the Middle Ages.[2] On the eve of the Second World War about 17m Germans inhabited regions to the east of what is Germany today. As seen in Table 1, roughly 10m people resided in areas, which are part of todays' Poland and Russia but used to be part of the German Reich. This area, known as the East German Territories, encompassed for example the regions of East Prussia and Silesia. In addition, there were sizable German minorities in other countries of Eastern Europe, most notably the so-called Sudetenland. This region in the north of Czechoslovakia has a long tradition of German settlements and was annexed by the Nazi Government in 1938. But many other countries like Poland, Hungary and Romania were also home to substantial German minorities.

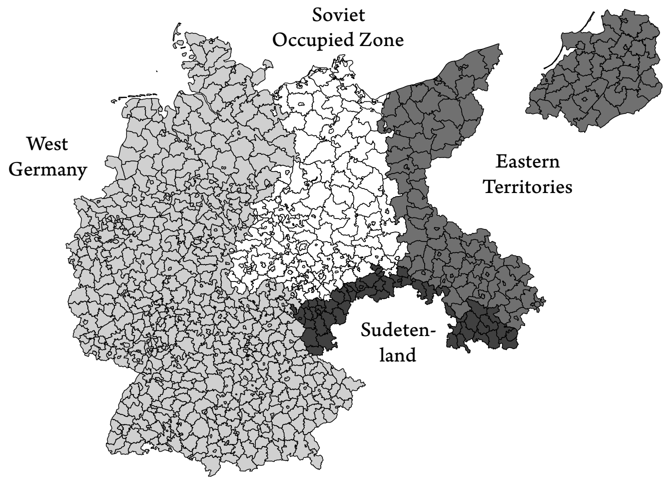
East German Territories	Czechoslovakia	Hungary	Romania	Poland	Others	Total
9.6m	3.5m	0.6m	0.8m	1m	1.4m	16.9m

Notes: The table shows the ethnic German population in different regions in Eastern Europe in 1939. The category "Others" comprises Danzig, the Baltic States and Yugoslavia.

Table 1: The German Population in Eastern Europe in 1939

The geography of the German Reich in 1939 is shown in Figure 1. I also display the boundaries of the individual counties, which is the 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 and 60s in these regions. In the far east, shown

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



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

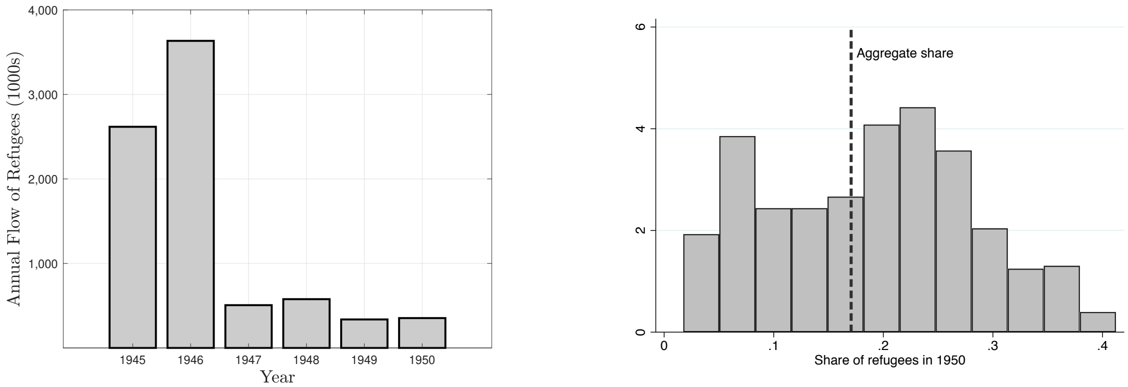
Figure 1: The German Reich in 1939

in medium dark, are the East German Territories. In the south-east, shown in dark, is the aforementioned Sudetenland in the north of Czechoslovakia. Finally, the light shaded area in the middle will become the Soviet Occupied Zone (in 1945) and then turn into the German Democratic Republic (in 1949). This area will not be part of the analysis of this paper.

The Expulsions and the Potsdam Conference in 1945

The Second World War brings an end to the presence of the German population in Eastern Europe as the entire German population either fled or were expelled in the aftermath of the war. This population transfer, where roughly 12 million ethnic Germans were forced to leave their domiciles, is one of the largest in world history. The expulsion can be broadly divided into three phases. The first wave of refugees arrived in West Germany during the last months of the war when Soviet forces made their appearance at the eastern German border. After the German defeat in May 1945, the so-called wild expulsions started. These took place in the spring and summer of 1945 mainly in Poland and Czechoslovakia, where both the army and privately organized militias systematically expelled the remaining German population. It is only at the Potsdam Conference in the Summer of 1945, that the Military Governments of the US, UK and Russia tried to put an end to these unorganized expulsions and legalized them ex-post.[3] In the official protocol of the conference they noted that *"the Three Governments,*

[3]The Potsdam Conference took place from 17 July to 2 August 1945 and was the main conference where the governments of Russia, the US and the UK discussed the restructuring of Germany and Europe after



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

Figure 2: The Settlement of Refugees in West Germany

having considered the question in all its aspects, recognize that the transfer to Germany of German populations, or elements thereof, remaining in Poland, Czechoslovakia and Hungary, will have to be undertaken. They agree that any transfers that take place should be effected in an orderly and humane manner." Within the following two years, the majority of the ethnic German population was transferred from Eastern Europe to West Germany and the Soviet Occupied Zone.

In the left panel of Figure 2 I depict the flow of refugees that arrived in West Germany. By the end of 1946, almost 6m refugees had arrived in West Germany. Between 1947 and 1950 another half million refugees arrived per year. By 1950, the inflow of refugees had increased the population in West Germany by about 8m individuals.[4] To put this population inflow into perspective, Table 2 reports a decomposition of the population dynamics in West Germany between 1939 and 1950. From the initial population of about 40m in 1939, West Germany suffered military and civilian losses of about 2.5m during the Second World War. At the same time, the country saw the arrival of 8m refugees and an additional 1.5m people fleeing the Soviet Occupied Zone. Hence, despite the casualties during the war, the population of West Germany increased by almost 8m people, i.e. by 20%, between 1939 and 1950.[5]

WW2. In addition to the expulsion of the German population, they also decided (among other things) on the redrawing of Germany's eastern border, the trials of the German war criminals, the division of Germany and Austria into different occupation zones and the payments of war reparations.

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

[5]In Section A-2.1 in the Appendix I show that refugees and natives were very similar demographically. In particular, both the age distribution and the share of males is almost identical in both groups.

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

Notes: The table reports aggregate population trends in West Germany between 1939 and 1950. "Inflows from SOZ" are individuals who fled the Soviet Occupied Zone.

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

The Initial Settlement in West Germany

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

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

To appreciate this unequal spatial distribution it is important to remember that an orderly settlement was an almost impossible task in war-torn Germany. A particular concern was the availability of housing as the rising population came hand in hand with a sharply diminished housing stock, which was heavily destroyed during the Allied Bombing Campaign. Werner Nellner, one of the leading post-war economic historians, described 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" (Nellner, 1959, p. 73). This uncoordinated settlement was already considered an enormous problem at the time. As early as in 1946, P.M. Raup, Acting Chief of the Food and Agricultural Division of the Office of the Military Government of the US (OMGUS) complained that "both the planning and the execution of the support measures for German expellees was conducted entirely under welfare perspectives. The people in charge at the Military Government are social service officials. Similarly on the side of the German civil government, the department in charge is the social service agency. Entire communities are moved so that the population of some counties is increased by 25-30% and the agency in charge was founded to support the elderly, disabled people and the poor. ... The whole problem has not been handled as one of settlements of entire communities but as an emergency problem supporting the poor." (Grosser and Schraut, 2001, p. 85).

These descriptions of the refugee settlement are also apparent in the data. In Table 3 I report the results of a set of bivariate regressions of the share of refugees in 1950 on different pre-war characteristics and a set of state fixed effects and report the coefficients on the respective characteristics. In column 1, I show that the share of refugees is strongly negatively correlated with the population-weighted distance to the expulsion region (the "expulsion distance" ED_c), which I calculate as

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

where $d_{c,j}$ is the geographical distance between county j and r, ER denotes the set of expulsion regions, i.e. the areas, which the german population had to leave and pop_r^{1939} is the size of their population in 1939. Hence, counties that were closer to the population centers of ethnic Germans in the pre-war period, experienced larger refugee inflows. This is exactly what one would expect if indeed the Military Governments of the US and UK experienced an institutional overload in distributing the refugees, which kept pouring in at the eastern border. In columns 2 and 3 I focus on the availability of housing. The share of refugees was much larger in regions with a low population density in the pre-war period as the abundance of land made it easier to expand the supply of housing. Furthermore, refugees were also much more likely to settle in counties, which suffered less from destruction during the war.[6]

These correlations imply that the inflow of refugees was systematically correlated with regional pre-war characteristics. In particular, the fact that refugees were settled in regions with low population density and little destruction during the war, implies that less developed,

[6]About 23% of the aggregate housing stock was damaged during the Allied bombing campaign. Moreover, there is considerable heterogeneity as a large share of counties saw more than 60% of their housing stock damaged during the war (see Section A-2.1 in the Appendix).

	Expulsion Distance	ln pop dens 1939	War time Dest.	Manufac. share 1939	Manufac. share 1933	Ag. share 1933	Rural share 1933	GDP pc 1935
β	-0.076***	-0.023***	-0.190***	-0.132***	-0.109***	0.087***	0.080***	-0.017***
	(0.011)	(0.002)	(0.011)	(0.022)	(0.020)	(0.011)	(0.008)	(0.005)
N	536	536	536	535	523	523	536	523
R^2	0.669	0.724	0.752	0.656	0.662	0.691	0.705	0.651

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. Each column reports the coefficient β of a regression $\mu_r = \delta_s + \beta x_r + u_r$ where μ_r is the share of refugees in 1950, x_r are the different regional characteristics in the respective columns and δ_s is a set of state fixed effects.

Table 3: Spatial Correlates of Refugee Inflows

remote areas saw their population increase the most. This is seen in the remaining columns of Table 3. A county's share of refugees is negative correlated with the share of manufacturing employment (both in 1933 and 1939) and positively correlated with its agricultural employment share. Moreover, counties with a larger share of refugees are more likely to be rural (as measured by the share of the population living in small cities) and have lower GDP per capita in 1935. My empirical strategy will take these systematic correlations into account.

Even though the size of the refugee settlement was systematically correlated with local characteristics, the absence of a specific allocation plan implied that there was very little spatial sorting of particular refugees into particular localities. If refugees had been spatially sorted by the government authorities, one would expect that the composition of the settled refugees would vary systematically with the pre-war industrial make-up. In particular, one would expect that refugees from the manufacturing-intensive Sudetenland would be sent to locations where the pre-war manufacturing share would be higher. This is not the case. As I show in Section A-2.1 in the Appendix, neither the manufacturing share nor GDP per capita predicts the composition of the refugee population. Importantly, this is not the case for individuals fleeing the Soviet Occupied Zone. These individuals are a natural control group as they were not part of the organized refugee treks but were free to settle. Empirically, these migrants do in fact settle systematically in richer and more manufacturing intensive locations.

The Spatial Persistence of the Initial Settlement

The long-run consequences of the initial settlement obviously depend crucially on the persistence of the shock. If the settled refugees left their originally assigned areas relatively quickly or simply replaced the native population, the likely consequences would be small. Table 4 shows that this was not the case but that the initial allocation of refugees was highly persistent and constituted an important source of regional population growth. In columns 1 - 3 I show that there is a strong correlation between the share of refugees in 1946, 1950, 1955

	Share of refugees in ...			Population growth ...	
	1950	1955	1961	1939-1950	1939 - 1961
Share of refugees in 1946	0.909***				
	(0.014)				
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)
N	404	536	484	536	488
R^2	0.952	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. All specifications control for state fixed effects.

Table 4: Refugees: Persistence and Population Growth

and 1961. In columns 4 and 5 I show that this persistence in the refugee settlement is also strongly correlated with subsequent population growth. Quantitatively, an increase in the share of refugees by 10 percentage points lead to an increase in population size by 1961 by 13%.

3 Refugees, Industrialization and Spatial Growth

In Section 2 I documented that localities differed vastly in the exposure to the inflow of refugees and that this inflow raised the local population in the long-run. In this section I turn to the economic consequences. My main analysis focuses on the effect on regional industrialization and income per capita. These cross-sectional estimates also form the backbone of my structural analysis as I estimate the structural parameters of my theory with indirect inference to fit these regression results.

3.1 Data

I use a variety of novel historical datasets, many of which were digitized for this project. The majority of the analysis exploits spatial variation across 500 counties in West Germany and links refugee inflows in the late 1940s to economic outcomes in the 1950s and 1960s. To do so, I constructed a panel dataset for West Germany at the county-level spanning the time-period from 1933 to 1970. The dataset was constructed by digitizing and spatially harmonizing a host of historical publications.

The basis of my dataset is comprised of the population censuses for the years 1933, 1939, 1950, 1961 and 1970, which are published individually for each of the nine states. For each of these years, the publications report a variety of outcomes at the county-level like the level of population, sectoral employment shares, occupational employment shares, sex ratios and various other characteristics.

I augmented this dataset with six additional pieces of information. The first concerns the regional allocation of refugees, which I digitized from a special statistical publication published in 1953. Secondly, in the 60s and 70s, the 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. Unfortunately, I was not able to find county-level GDP measures for the pre-war and immediate post-war period. To be able to control for pre-war economic activity and to estimate the effect of refugee inflows on income per capita in the short-run, I digitized county-level information from tax records, which report value added taxes for each county in 1935 and 1950. Because tax rates did not vary across space in Germany, I take these measures as being proportional to local GDP.

Fourth, I digitized the county-level results for four waves of the manufacturing census in 1933, 1939, 1950 and 1956. They report the number of plants at the county-level and hence allow me to directly measure plant entry, which is an important theoretical mechanism of my model. Fifth, I provide new measures of the extent of war time destruction and regional housing supply, which I digitized from the historical housing census conducted in 1950. This census contains information on the extent of war damages for each county.[7] Finally, I digitized the historical migration census from 1955, which reports inflows and outflows for each of the 500 counties. This information is extremely useful to estimate the structure of mobility costs in the quantitative model.

I complement my analysis with two micro datasets. The first is the Mikrozensus Zusatzerhebung 1971 (MZU 71), a special appendix to the census conducted in 1971 to measure social mobility. It includes identifiers on individuals' refugee status and retrospective information about individuals' employment characteristics in 1939, 1950, 1960 and 1971. The MZU 71 has roughly 200,000 observations, 40,000 of which are refugees. Hence, this survey is representative in that refugees account for 20% of the population. The second is the Einkommens- und Verbrauchsstichprobe 1962/63 (EVS 62), which is a micro dataset conducted in 1962 to measure household expenditure and hence similar to the Consumer Expenditure Survey in the US. It has about 32,000 observations and I exploit the information on refugees' relative earnings to estimate the distribution of human capital in the quantitative model.

[7] Crucially, 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 wartime 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.

	Sectoral employment shares					
	Manufacturing			Agriculture	Services	Manufacturing
	1939	1950		1950		1961
Share of refugees (1950)	-0.077	0.235***	0.315***	-0.264***	-0.042	0.249***
	(0.089)	(0.083)	(0.047)	(0.072)	(0.069)	(0.058)
Manufac. share (1939)			0.875***	-0.534***	-0.341***	0.885***
			(0.036)	(0.044)	(0.040)	(0.040)
N	535	536	535	535	535	535
R^2	0.461	0.486	0.896	0.863	0.702	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 (col 1), in 1950 (cols 2, 3), the agricultural employment share in 1950 (col 5), the service employment share in 1950 (col 6) and the manufacturing employment share in 1961 (cols 7). 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. The specifications in cols 3 - 7 also control for the log of the distance to the inner german border and a fixed effect for whether a county is a county at the inner german border.

Table 5: Refugees and Sectoral Employment Shares

3.2 Refugees, Industrialization and Local Productivity

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

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

where π_{rt}^{M} denotes the manufacturing employment share at time t and μ_{r1950} is the share of refugees in 1950. Furthermore I control for a set of state fixed effects (δ_s), population density in 1939 (l_{r1939}) and the extent of wartime destruction (wd_r), which are the important determinants of the housing supply (and hence refugee flows), 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. I estimate this specification both via OLS and with an instrumental variable strategy (see Table 7 below).

Consider first the OLS results reported in Table 5. In column 1 I consider a "Placebo"-like specification and show that - reassuringly - there is no relationship between the share of refugees in 1950 and the manufacturing share in 1939.[8] 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 column

[8]In Section A-2.1 in the Appendix I also report a set of regressions, which suggest counties with high exposure to the refugee settlement were not on differential trends. In particular, I show that the share of refugees is uncorrelated with pre-war population growth between 1933 and 1939 and if anything negatively correlated with the change in the manufacturing employment share between 1933 and 1939.

	ln GDP per capita					
	1950	1957-1966	1957	1961	1964	1966
Share of refugees (1950)	-0.062	0.455***	0.461**	0.607***	0.645***	0.603***
	(0.302)	(0.100)	(0.199)	(0.185)	(0.218)	(0.198)
N	523	2076	518	518	518	518
R^2	0.722	0.730	0.661	0.574	0.436	0.427

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The dependent variable is log GDP per capita in 1950 (cols 1) and log GDP per capita in 1957, 61, 64 and 66 in cols 2-6. In column 2 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 log GDP per capita in 1935, 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.

Table 6: Refugees and GDP per capita

3 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 4 and 5 I show that the sectoral reallocation takes place exclusively between the manufacturing and the agricultural sector. In contrast, the employment share of the 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. The last column shows that the expansion of the manufacturing sector is long-lived as the inflow of refugees is strongly related to manufacturing specialization in 1961. This is consistent with the results in Table 4, which showed that the inflow of refugees had a persistent positive effect on spatial population growth.

I now estimate the effect of the initial refugee settlement on regional income per capita. My main empirical specification takes a similar form to the specification in (2):

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

where y_{ct} denotes GDP per capita in region r at year t. Hence, β captures the effect of the inflow of refugees on the growth rate of income per capita since 1935. All remaining variables are defined as above. Again, I estimate (3) using both OLS and an IV strategy.

Table 6 contains the OLS results. The first two columns show that income per capita growth between 1935 and 1950 is essentially unrelated to the inflow of refugees. Note that if agricultural production is subject to decreasing returns (as suggested by the results in Table 2), this implies that the manufacturing sector is subject to increasing returns, which for the local economy as a whole cancel out. In the remaining columns I focus on the long-run effects. In column 2 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. Hence, the long-run effect is positive, suggesting a form of dynamic agglomeration. In the remaining columns I estimate the long-run effect separately for each year. The results are quantitatively consistent with the pooled regression. Interestingly the point estimates are increasing between 1957 and 1964. However, these differences are too small to detect statistically.

Instrumental Variable Estimates For the structural estimation of my theory I use the OLS estimates reported in Tables 5 and 6 as identified moments and estimate the structural parameters via indirect inference by targeting the regression coefficients. As complementary evidence that these results reflect the causal effect of changes in the local refugee population I now present an instrumental variable strategy to estimate specifications (2) and (3). The identification strategy is based on the systematic geographic variation between the share of refugees and the population-weighted distance to the expulsion regions.

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 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, 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, pushing against the main findings of this paper. Third, in Section A-2.2 of the Appendix I also offer an additional instrumental variable strategy, which is less subject to these concerns but also less precisely estimated.[9]

The results are contained in Table 7. The first column contains the first stage relationship and confirms that there is still a strong negative relationship between the expulsion distance and the share of refugees even if the controls, which are present in the second stage, are

[9]This strategy exploits the 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 houses, 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. The results of this estimation strategy are qualitatively similar but less precisely estimated. The estimated relationship with the manufacturing share (both in 1950 and 1961) is somewhat larger but not statistically different from the baseline results. The relationship with GDP per capita is very noisy. While the point estimates are positive, neither in 1950 nor in the later periods 1957-1966 is the relationship significant.

	First stage	Man. empl. share		ln GDP pc	
	Share of refugees in 1950	1950	1961	1950	1957 - 1966
Expulsion Distance	-0.118***				
	(0.033)				
μ_{r1950}		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)
N	523	523	523	523	2076
R^2	0.784	0.909	0.803	0.702	0.749

Note: Robust standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The distance to the expulsion population is calculated according to (1) and used as an instrument for the share of refugees. Column 1 contains the first stage relationship, columns 2 - 5 contain the second stage relationship. The dependent variable is the manufacturing employment share in 1950 (column 2) and 1961 (column 3) and log GDP per capita in 1950 (column 4). In column 5 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, a fixed effect for whether a county is a border county and the manufacturing employment share in 1933 and 1939.

Table 7: Instrumental Variable Estimates

included. The remaining columns contain the IV results for manufacturing employment and income per capita. Columns 2 and 3 show the results for manufacturing employment and correspond to columns 4 and 6 in the OLS specification of Table 5. Columns 4 and 5 focus on income per capita (the corresponding OLS specifications are shown in columns 1 and 2 of Table 6). All these estimates are very similar to the corresponding OLS estimates.

Robustness In Section A-2.2 in the Appendix I report a battery of robustness checks for the results reported in Tables 5, 6 and 7. In particular, (i) I control for spatial variation in labor supply (as proxied by the aggregate employment share and the share of males at the county level) or local demand for reconstruction (as proxied by the share of the housing stock built after 1945), (ii) I report the results when counties are weighted by their population size and when standard errors are clustered at the state level, (iii) I control for both population density and manufacturing employment shares in 1933 (in addition to 1939), (iv) I use the refugee share as of 1946 (instead of 1950) as the dependent variable and (v) I show that the results are not driven by particular cities or particular states. For the manufacturing employment share I find that the results are virtually identical to the baseline results reported in Table 5. For the case of GDP per capita, the only difference is that the effects of the refugee share in 1946 on GDP per capita in the 1960s is positive and but not statistically significant. All other specifications are statistically indistinguishable to the baseline results reported in Table 6. Similarly, the IV estimates reported in Table 7 are insensitive to these considerations.

	Average	Quantiles				
		10%	25%	50%	75%	90%
Manufacturing	1.36	0.93	1.10	1.31	1.54	1.81
Agriculture	0.37	0.18	0.21	0.29	0.40	0.76

Note: The table reports the distribution of the sectoral employment share of refugees relative to the one of natives in each county for the state of Bavaria, i.e. $s_{rst}^{Ref}/s_{rst}^{Nat}$ where s_{rst}^{Ref} (s_{rst}^{Nat}) is the employment share of refugees (natives) in sector s in region r.

Table 8: The Manufacturing Bias of Refugees' Labor Supply

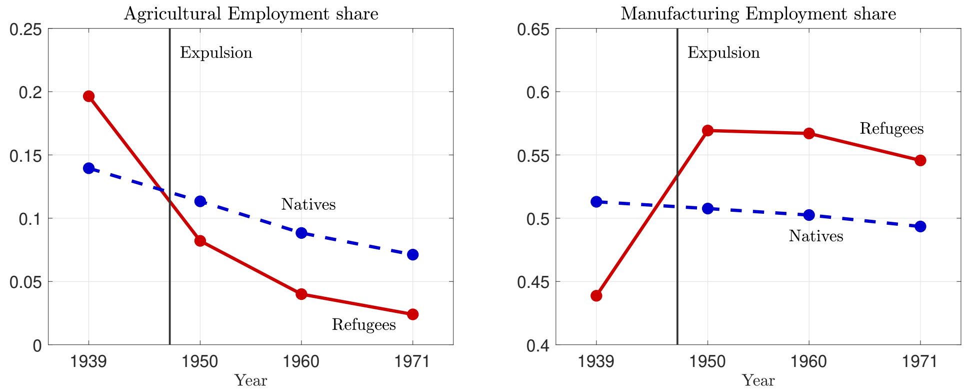
3.3 Spatial Sorting: Refugees as a Source of Manufacturing Labor

One important reason for the expansion of the local manufacturing sector was that the incoming refugees often ended up as manufacturing workers. In Table 8 I report the distribution of refugees' sectoral employment shares relative to natives *within* counties, i.e. $s_{rs}^{Ref}/s_{rs}^{Nat}$ where s_{rs}^{Ref} (s_{rs}^{Nat}) is the employment share of refugees (natives) in sector s in county r.[10] A value of unity indicates that refugees and natives have the same sectoral employment shares. The table shows a clear pattern of comparative advantage within local labor markets: facing the same wages, refugees are on average 36% more likely to work in manufacturing. In contrast, the average agricultural employment share among refugees is only 37% as large as the one of natives. The remaining columns show that these patterns hold throughout the entire distribution of counties. In less than 20% of counties are refugees less likely to work in manufacturing and there is no instance for refugees to be more likely to work in the agricultural sector.

This sectoral sorting is also apparent in the longitudinal microdata of the MZU 71. This unique supplement to the 1971 population census asked every respondent where he or she lived in 1939 and in which occupation and sector he or she worked in 1939, 1950, 1960 and 1971. By analyzing the time-series of these retrospective questions, I can measure snapshots of the employment life-cycle for both refugees and natives for a 40 year window. In Figure 3 I depict the sectoral life-cycle profile for the cohort of individuals born between 1915 and 1919. This cohort is 20-25 years old in 1939 and in their late twenties or early thirties at the time of the expulsion around 1947. In 1971, this cohort is 50-55 years old, i.e. still in the labor force. The two panels show the agricultural employment share (left panel) and the manufacturing employment share (right panel). The vertical line indicates the time of the expulsion. The process of reallocation is vividly apparent. Among refugees, 20% of the twenty year olds in 1939 used to work in the agricultural sector.[11] After the expulsion and their resettlement to West

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

[11] 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



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 vertical line.

Figure 3: The Life-Cycle of the 1915-1919 Cohort

Germany, only 8% still did so. In contrast, the share of manufacturing employment, within the same cohort of individuals, increases from 44% to 57% after the settlement. The pattern for natives is strikingly different as the time period of the expulsion is hardly noticeable. The secular decline in agricultural and manufacturing employment for both natives and refugees in the post-war period reflects the process of structural change towards the service sector.

One way to rationalize these findings is as a consequence of spatial sorting. Because refugees were predominantly allocated to rural areas, the average native working in a rural area might have had a comparative advantage in the agricultural sector relative to a randomly selected refugee, if native workers were spatially sorted based on their skills, Intuitively, the share of engineers within an arriving refugee trek might have been higher that in the rural native population. This mechanism will feature prominently in the theory and in Section 5 I provide direct evidence for the importance of spatial sorting.

4 Theory: A Model of Spatial Growth

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

cohorts.

4.1 Environment

I consider an economy with R regions (counties in the data). Individidual 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 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 $s = A, M$ are in turn CES aggregates from a set of differentiated, regional varieties. In particular, each region produces a differentiated variety, which is tradable across space (subject to an iceberg trade cost τ_{rj}) and aggregated according to $Y_{st} = \left(\sum_{r=1}^{R} Y_{rst}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$, where Y_{rst} is the amount of sector s output sourced from region r and $\sigma > 1$ denotes the elasticity of substitution across varieties, which for simplicity I assume to be the same between agricultural and manufacturing goods. Letting P_{rst} denote the price of goods from r in r, consumers in region j pay $\tau_{rj} P_{rst}$ for region r goods. The price index of sector s goods in region j is therefore given by $\overline{P}_{jst} = \left(\sum_r \left(\tau_{rj} P_{rst}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$.

Production The agricultural good is produced using labor, measured in efficiency units, and land according to

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

where T_r denotes the fixed total supply of land in region r and represents a source of comparative advantage for the production of agricultural goods. H_{Art} denotes the total amount of labor employed for agricultural production and Q_{rt} is productivity in region r at time t. Evidently, agricultural production is subject to decreasing returns to scale. The returns to the fixed factor land accrue to a set of immobile land-owners which only consume.

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

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

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

$$X_{rt} \quad = \quad \left(\int_0^{N_{rt}} x_{it}^{(\rho-1)/\rho} di \right)^{\rho/(\rho-1)},$$

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

The regional productivity term Q_{rt} evolves according to the persistent process

$$\ln Q_{rt} = (1 - \varrho) \ln Q_r + \varrho \ln Q_{rt-1} + \varpi u_{rt}, \tag{5}$$

where Q_r is a fixed, region-specific level of innate productivity, ϱ governs the regional autocorrelation and u_{rt} is a productivity shock, which is distributed iid with a unit variance. 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 measure of input varieties N_{rt} is determined endogenously and provides the link between local productivity and labor supply. At the end of each period (after production has taken place) an exogenous fraction δ of firms exits. Firm entry takes place in the beginning of the period. The labor requirement to start a new firm in region r at time t, h_{rt}^E, is given by

$$h_{rt}^E = f_E N_{rt-1}^{-\lambda}, \tag{6}$$

where $\lambda \leq 1$. The parameter f_E parametrizes the size of entry costs and λ governs the extent of dynamic spillovers as in Jones (1995). As I 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 effects. I 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 N_{rt-1} affects the costs of creating new knowledge.

The entry technology in (6) nests three important benchmark models as special cases and the structural estimation will allow me to distinguish between them. If $\lambda = 0$ and $\delta = 1$, the model is the static model of Krugman (1980): firms only live for a single period and the cost of entry do not depend on the number of varieties already available. If $\lambda > 1$ there are dynamic spillovers as entry costs are declining in the measure of existing varieties. The case of $\lambda = 1$ is the specification of Romer (1990), where the costs of creating new varieties are inversely proportional to the level of knowledge $N_{t-1,r}$. I will show below that this specification of the model leads to fully endogenous growth. The intermediate case of $0 < \lambda < 1$ is the semi-endogenous growth model of Jones (1995), where growth in the long-run is fully determined by population growth. As I discuss in detail below, these different parameterizations have strikingly different implications for the dynamic effects of refugee inflows on income per capita.

Sectoral Labor Supply I model the sectoral supply of human capital using the usual Roy-type machinery. Individuals are characterized by a two dimensional efficiency vector $z_{it} = (z_{iAt}, z_{iMt})$, where z_{ijt} denotes the number of efficiency units individual i can supply to sector j 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 parameterizes 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. 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 parameterize the respective distribution of skills in sector j. Standard arguments show that the share of individuals of type $\nu \in \{F, I\}$ working in sector j in region r is given by

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

$\overline{w}_{rt}^{\nu}$ also denotes the average earnings of type ν in region r at time t. Allowing for latent types is important in that it gives a distinct role for the composition of the local population to determine aggregate labor supply. In particular, if region r is inhabited by L_{rt}^{ν} workers of type ν, the total supply of human capital in sector j is given by

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

where $\Gamma_{\theta} = \Gamma(1-\theta^{-1})$ is the gamma function, $L_{rt} = L_{rt}^I + L_{rt}^F$ is the total population in county r and $\omega_{rt}^{\nu} = L_{rt}^{\nu}/L_{rt}$ is the share of ν-types. Hence, aggregate sectoral labor supply depends on the total population size and the degree of spatial sporting encapsulated in ω_{rt}^{ν}. Because population mobility is endogenous, the extent of spatial sorting is also endogenous.

To connect this model to the historical evidence, denote the number of natives and refugees at time t by L_t^N and L_t^R. For parsimony, I assume that the population of refugees and natives are exactly identical in terms of their skills and I denote the number of natives (refugees) of type ν 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 if native workers are endogenously spatially sorted but refugees are sent to rural locations.

Spatial Mobility Individuals are mobile across space, but subject to a friction. While individuals know their type before making their moving decisions, they do not observe their particular skill realization z_{it}. The utility for individual i who currently lives in region j to move to region r at time t is then given by

$$\mathcal{U}_{jrt}^{i} = V_r \overline{u}_{rt}^{i} \eta_{jr} \xi_{rt}^{i}.$$

Here, V_r denotes an exogenous amenity in region r, $\overline{u}_{rt}^{i}$ is the expected utility individual i achieves in region r, η_{jr} parametrizes the moving costs from j to r and ξ_{rt}^{i} 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 $\varepsilon > 1$, i.e. $F(\xi) = e^{-\xi^{-\varepsilon}}$. Given the utility function in (4), expected utility (or real income) of individual i of type ν in region r is given by $\overline{u}_{rt}^{\nu} \propto \overline{w}_{rt}^{\nu} / \left(\overline{P}_{rAt}^{\alpha} \overline{P}_{rMt}^{1-\alpha} \right)$, where $\overline{w}_{rt}^{\nu}$ is the expected income type ν can earn in region r and $\overline{P}_{rst}$ is the price index in sector s in region r.

Motivated by the fact that the spatial allocation of refugees was highly persistent, I allow for (in addition to the moving costs encapsulated in η_{jr}) a second mobility friction a la Calvo: at each point in time individuals have the option to move with probability $\psi > 0$ (see also Bilal (2019)). The law of motion of the spatial population distribution then takes the simple form

$$L_{rt}^{\nu} = (1-\psi) L_{rt-1}^{\nu} + \psi \sum_{j=1}^{R} L_{jt-1}^{\nu} m_{jrt}^{\nu},$$

where L_{rt}^{ν} denotes the number of $\nu-$types in region r at time t and m_{jrt}^{ν} denotes their probability to move from j to r in period t

$$m_{jrt}^{\nu} = \frac{\left(V_r \eta_{jr} \overline{u}_{rt}^{\nu}\right)^{\varepsilon}}{\sum_k \left(V_k \eta_{jk} \overline{u}_{rt}^{\nu}\right)^{\varepsilon}}.$$

The combination of $\psi < 1$ and $\eta_{jr} \neq \eta_{kr}$ for $j \neq k$ parsimoniously captures the intensive and extensive margin of costly migration. In my quantitative application I assume a particular functional form $\eta_{jr} \propto d_{jr}^{-\kappa}$, where d_{jr} is the geographic distance between j and r and κ is a parameter, which I estimate. The higher κ, the more local are migration flows. If $\kappa = 0$, moving costs are constant across space and the migration probabilities m_{jr} will be equalized across regions of origin j.[12]

4.2 Equilibrium

The timing of events is as follows. At the beginning of the period, the set of state variables in region r is given by $\left(Q_{rt-1}, N_{rt-1}, L^{\nu}_{rt-1}\right)$. Then the exogenous productivity shock Q_{rt} is realized. Given $\{Q_{rt}, N_{rt-1}\}_r$, individuals make their mobility decision, potential entrants decide on their entry activities and production and consumption take place. Individuals' mobility decisions and firms' entry decisions determine the future set of state variables $\{L^{\nu}_{rt}, N_{rt}\}_r$.

Static Allocations To solve for the static equilibrium allocations, consider first the manufacturing sector.[13] 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 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 shared among the N_{rt} firms. This implies that profits of firm i in region r are given by

$$\pi_{ir} = \frac{1}{\rho}\frac{(1-\beta)\, P_{rM}Y_{rM}}{N_r}.$$

Similarly, workers in the production of the manufacturing good receive a share β of total revenue, i.e. $w_{rM}H_{rP} = \beta P_{rM}Y_{rM}$. One can also show that total labor demand of intermediate firms is given by

$$H_{rX} = \int_0^{N_{rt}} x_i di = \frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right) H_{rP},$$

i.e. total intermediate employment in region r, H_{rX}, is always proportional to total production employment H_{rP}. Aggregate manufacturing output Y_{rMt} is therefore given by

$$Y_{rMt} = \left(\frac{\rho-1}{\rho}\frac{1-\beta}{\beta}\right)^{1-\beta} Q_{rt} N_{rt}^{\frac{1-\beta}{\rho-1}} H_{rPt}. \tag{7}$$

Equation (7) shows the usual variety gains: a larger mass of varieties N_{rt} increases productivity and these productivity gains are higher the lower the elasticity of substitution ρ. I there also

[12] The law of motion of the population highlights that the past population distribution $\left\{L^{v}_{jt-1}\right\}_j$ is a dynamic state variable. This is in contrast to the theory developed in Desmet et al. (2018), which allows for fully forward looking agents but imposes sufficient assumptions so that the spatial mobility problem ends up being static in nature. This framework therefore does not naturally lend itself to study an exogenous change in the population distribution.

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

sometimes refer to N_{rt} as "productivity".

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 the labor-land ratio H_{rA}/T_r because of decreasing returns.

Entry, Market Size and Dynamic Agglomeration As for workers, I assume that entering firms act myopically, only considering static profits as part of their entry decision.[14] Free entry therefore requires that[15]

$$\pi_r = w_{rMt} h_{rt}^E = w_{rMt} f_E N_{rt-1}^{-\lambda}. \tag{8}$$

Using the expressions for profits π_{ir} and for the manufacturing wage w_{rMt}, the free entry condition implies a simple expression for the evolution of regional varieties

$$N_{rt} = \frac{1-\beta}{\rho\beta} \frac{1}{f_E} \underbrace{H_{rPt}}_{\text{Market size}} \times \underbrace{\left(N_{rt-1}\right)^{\lambda}}_{\text{Dynamic agglomeration}}. \tag{9}$$

Equation (9) 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 scale effect: a larger workforce H_{rPt} triggers the entry of new varieties because it goes hand in hand with larger profits. Note that H_{rPt} emerges as a sufficient statistic that summarizes all equilibrium effect on sectoral wages and aggregate demand, which are determined as part of the trade and spatial equilibrium. The second term captures the dynamic agglomeration force. As long as $\lambda > 0$, the equilibrium features persistence whereby the existing number of varieties positively predicts the future number of varieties. As I will discuss in detail below, this extent of persistence has important implications for both the time-series and the spatial predictions of the theory.

Static Agglomeration and the Short-Run Scale Elasticity Even though regional manufacturing output is a linear function of the number of production workers (see (7)), the number of active firms N_{rt} is itself a function of market size. Combining (7) and (9) therefore

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

[15] 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 (8) holds with inequality, i.e. $\pi_{irt} < w_{rMt} f_E \left((1-\delta) N_{rt-1}\right)^{-\lambda}$. As I show in Section A-1.2 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 (8) holds with equality in the main text. In the quantitative application I of course allow for the general case where (8) might be slack.

yields the equilibrium production function

$$Y_{rMt} = \varsigma Q_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{1+\vartheta} \quad \text{where} \quad \vartheta = \frac{1-\beta}{\rho-1}, \tag{10}$$

and ς is an inconsequential constant. I refer to ϑ as the "short-run" scale elasticity, because, holding a location's state variables (Q_{rt}, N_{rt-1}) fixed, manufacturing productivity is increasing in the size of the local workforce if $\vartheta > 0$. The modifier "short-run" is important. As I show below, the long-run elasticity between productivity and population size is qualitatively different and depends directly on the knowledge elasticity λ. Note also that the production structure in (10) is isomorphic to a production function with exogenous agglomeration forces common in many models of economic geography (Redding and Rossi-Hansberg, 2017).

Dynamic Equilibrium Given the environment above, I now characterize the dynamic equilibrium of this economy.

Definition 1. *A 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 $\left\{L_{rt}^{N\nu}, L_{rt}^{R\nu}\right\}_{\nu rt}$, employment allocations $\{H_{rAt}, H_{rPt}, H_{rXt}, H_{rEt}\}_{rt}$, regional manufacturing varieties $\{N_{rt}\}_{rt}$ and quantities of tradable goods $\{Y_{rAt}, Y_{rMt}\}_{rt}$ such that (i) firms and consumers behave optimally, (ii) labor and good markets clear at each point in time, (iii) the evolution of the local population is consistent with individuals' optimal mobility decisions and (iv) the evolution of the number of varieties is consistent with free entry.*

To represent the equilibrium in a parsimonious way, note that aggregate labor demand in the manufacturing sector is given by[16]

$$H_{rMt} = H_{rPt} + H_{rXt} + H_{rEt} = \frac{1}{\beta} H_{rPt} - (1-\delta) f_E N_{rt-1}^{1-\lambda}. \tag{11}$$

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}. Note also that the composition of refugees and natives within a locality does not matter conditional on the composition of types.

Proposition 2. *Let $\left\{L_{r0}^I, L_{r0}^F, N_{r0}, Q_{r0}\right\}_{r=1}^R$ be given. The dynamic equilibrium is fully char-*

[16]Note that the number of workers who are engaged in entry activities in region r is given by $H_{rEt} = f_E N_{rt-1}^{-\lambda} \left(N_{rt} - (1-\delta) N_{rt-1}\right) = \frac{1-\beta}{\rho\beta} H_{rPt} - (1-\delta) f_E N_{rt-1}^{1-\lambda}$.

acterized by the spatial goods market clearing equation

$$\frac{w_{rMt}H_{rPt}}{\beta} = \sum_{j=1}^{R}\left(\frac{\tau_{rj}P_{rMt}}{\overline{P}_{jMt}}\right)^{1-\sigma}(1-\alpha)\left(\frac{w_{jAt}H_{jAt}}{\gamma}+\frac{w_{jMt}H_{jPt}}{\beta}\right) \quad (12)$$

$$\frac{w_{rAt}H_{rAt}}{\gamma} = \sum_{j=1}^{R}\left(\frac{\tau_{rj}P_{rAt}}{\overline{P}_{jAt}}\right)^{1-\sigma}\alpha\left(\frac{w_{jAt}H_{jAt}}{\gamma}+\frac{w_{jMt}H_{jPt}}{\beta}\right), \quad (13)$$

where $P_{rMt} = \frac{1}{Q_{rt}N_{rt-1}^{\lambda\vartheta}H_{rPt}^{\vartheta}}w_{rMt}$, $P_{rAt} = \frac{1}{Q_{rt}}w_{rAt}\left(\frac{H_{rAt}}{T_r}\right)^{1-\gamma}$ *and* $\overline{P}_{jst} = \left(\sum_r (\tau_{rj}P_{rst})^{1-\sigma}\right)^{1/(1-\sigma)}$, *the labor market clearing conditions*

$$H_{rAt} = \Gamma_\theta \sum_{\nu=I,F} L_{rt}^{\nu}\left(\phi_A^{\nu}\right)\left(\frac{w_{rAt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1} \quad (14)$$

$$\frac{1}{\beta}H_{rPt} - (1-\delta)f_E N_{rt-1}^{1-\lambda} = \Gamma_\theta \sum_{\nu=I,F} L_{rt}^{\nu}\left(\phi_M^{\nu}\right)\left(\frac{w_{rMt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1}, \quad (15)$$

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

$$L_{rt}^{\nu} = (1-\psi)L_{r,t-1}^{\nu} + \psi\sum_{j=1}^{R}L_{jt-1}^{\nu}\frac{\left(V_r\eta_{jr}\overline{w}_{rt}^{\nu}\right)^{\varepsilon}}{\sum_k\left(V_k\eta_{jk}\overline{w}_{kt}^{\nu}\right)^{\varepsilon}} \quad \text{for } \nu = I,F, \quad (16)$$

$$N_{rt} = \frac{1-\beta}{\rho\beta}\frac{1}{f_E}H_{rPt}\times N_{rt-1}^{\lambda}, \quad (17)$$

and the law of motion for Q_{rt} *in (5).*

Proof. See Section A-1.3 in the Appendix. □

The first four equations are the goods and labor market clearing conditions.[17] Given the number of pre-determined varieties N_{rt-1} and the mass of individuals L_{rt}^{ν}, these equations are static.

The crucial equations are the laws of motion for the dynamic state variables $\{L_{rt}, N_{rt}\}$ as they govern how the joint distribution of population and regional productivity evolves. Whether the equilibrium allocations show persistence depends crucially on the extent of spatial mobility (governed by ψ and η_{jk}) and the knowledge production function (parametrized by the knowledge elasticity λ and the depreciation of the capital stock δ). With free mobility, i.e. $\psi = 1$ and $\eta_{jk} = 1$, the distribution of people across space ceases to be a state variable.

[17] As for the owners of land, I assume that firm profits accrue to a set of spatially immobile entrepreneurs who simply consume these returns.

If in addition there are no dynamic spillovers, i.e. $\lambda = 0$, and varieties depreciate fully, i.e. $\delta = 1$, the model is a static model of economic geography with agglomeration forces as in Allen and Arkolakis (2014) or Ahlfeldt et al. (2015). This parameterizations of the model is inconsistent with the empirical finding that the population shock was persistent and that the effect on income per capita was increasing over time.

4.3 Population Inflows and Dynamic Agglomeration

To see why population shocks have dynamic consequences, note that (9) implies that

$$\ln N_{rt} = \alpha_0 + \underbrace{\lambda \ln N_{rt-1}}_{\text{Mean reversion}} + \underbrace{\ln H_{rPt}}_{\text{Labor supply shocks}},$$

where $\alpha_0 = \ln\left(\frac{1-\beta}{\rho\beta}\frac{1}{f_E}\right)$. Hence, as long as $\lambda > 0$, local productivity N_{rt} endogenously follows a persistent AR(1) process where the "shocks" are endogenous as they depend directly on the extent to which factors are mobile across space. This implies that for any period t_0 and time period $\tau \geq t_0$

$$\ln N_{r\tau} \quad = \quad \Lambda\left(t, t_0\right) + \lambda^{\tau-(t_0-1)} \ln N_{rt_0-1} + \sum_{j=t_0}^{\tau} \lambda^{\tau-j} \ln H_{rPj}, \tag{18}$$

where $\Lambda\left(t, t_0\right) = \alpha_0 \sum_{j=t_0}^{\tau} \lambda^{j-t_0}$. Hence, 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 the inter-temporal knowledge elasticity λ. Intuitively: local productivity at any point in time encapsulates the entire history of local scale, as such scale led to plant entry, which made the creation of future varieties easier. If there is no mean reversion, i.e. $\lambda = 1$, past population size does not have to be discounted at all: population shocks have persistent productivity effects and the spatial productivity distribution is not stationary.

This expression is useful to understand the dynamic productivity effects of an exogenous change in labor supply at time t_0. For concreteness, suppose there is a positive shock at t_0, which subsides at rate $p < 1$. Hence, for period $t_0 + d$ we have $d \ln H_{rPd+t_0} = d \ln H_{rPt_0} \times p^d$. If $p = 0$, the manufacturing workforce H_{rPt} returns to its pre-shock level immediately. If $p = 1$, the shock is persistent. As I show in Section A-1.5 in the Appendix, the elasticity of productivity at time $d + t_0$ with respect to the shock at time t_0 is described by the impulse response function $\Psi_d\left(p, \lambda\right)$

$$\frac{d \ln N_{rd+t_0}}{d \ln H_{rPt_0}} \equiv \Psi_d\left(p, \lambda\right) = \frac{\lambda^{d+1} - p^{d+1}}{\lambda - p}.$$

If the shock is purely transitory, i.e. $p = 0$, we get that, assuming $\lambda < 1$, $\Psi_d(0, \lambda) = \lambda^d \xrightarrow{d \to \infty} 0$. Hence, a transitory population inflow shows persistence but the effect is *declining* over time. On the other extreme, if the shock was permanent, i.e. $p = 1$, we have $\Psi_d(1, \lambda) = \frac{1-\lambda^{d+1}}{1-\lambda} \xrightarrow{d \to \infty} \frac{1}{1-\lambda}$, i.e. the effect is *increasing* over time as the endogenous creation of varieties acts as an amplifying force. If $\lambda = 1$, the productivity effect keeps accumulating and $\Psi_d(1, 1) \to \infty$. 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 and the persistence can be sizable. This captures the qualitative pattern that I documented empirically: the short-run effect of the inflow of refugees on GDP per capita in 1950 was small, the long-run effect on GDP per capita the 1960s was large.

4.4 The Long-Run Scale Elasticity

To study the long-run impacts of the refugee inflow, consider the behavior of the economy along a non-stochastic spatial balanced growth path (SBGP), which I define as an allocation where the population distribution is stationary and regional wages grow at a common rate. Along a SBGP innate productivity Q_{rt} is equal to its long-run level Q_r and 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 varieties grow at a common rate in all locations:

$$g_N = \frac{N_{rt}}{N_{rt-1}} = \frac{1-\beta}{\rho\beta} \frac{1}{f_E} H_{rPt} N_{rt-1}^{\lambda-1}. \tag{19}$$

Equation (19) has obvious similarities to the growth equation analyzed in Jones (1995). For g_N to indeed be constant across space, $H_{rPt} N_{rt-1}^{\lambda-1}$ has to be equalized across locations. Consider first the case of $\lambda < 1$. This implies that the level of productivity along a SBGP is given by

$$N_{rt} = \left(\frac{1}{g_N} \frac{1-\beta}{\rho\beta} \frac{1}{f_E} \right)^{\frac{1}{1-\lambda}} H_{rPt}^{\frac{1}{1-\lambda}}. \tag{20}$$

Hence, the equilibrium growth rate is given by $g_N = \frac{n}{1-\lambda}$, where n is the growth rate of the manufacturing workforce H_{rPt}, which could either be driven by population growth or by the accumulation of human capital. Hence, if $\lambda < 1$ this is a model of semi-endogenous growth as in Jones (1995), where the growth rate is endogenous but in equilibrium determined directly from exogenous parameters.

Equation (20) has two important implications for the long-run distribution of economic activity across space. First, it is apparent that the spatial distribution of productivity reflects local scale effects: regions where H_{rPt} is large will have high productivity. Hence, regional scale has level effects and the spatial distribution of productivity is jointly determined with individuals' sorting behavior across sectors and space. Crucially, however, this long-run relationship between productivity and population size is fundamentally different from the short-run relationship. In (10) I showed that productivity in the manufacturing sector is given by

$$Y_{rMt}/H_{rPt} \propto Q_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{\vartheta},$$

i.e. ϑ is exactly the elasticity between local productivity and local scale, holding N_{rt-1} constant. Equation (20) shows that this elasticity along a SBGP is quite different. In particular, substituting for the BGP value of N_{rt}, we get that

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

i.e. the long-run elasticity between population size and productivity is given by $\vartheta/(1-\lambda)$. Note that this long-run elasticity always exceeds the short-run elasticity as long as there are dynamic spillovers in the entry technology, i.e. $\lambda > 0$.

The second important implication is that the long-run distribution of economic activity is stationary and fully determined from regional fundamentals. This implies that the effects of population shocks will necessarily vanish in the long-run. Hence, population shocks might show persistence but will not affect the stationary distribution. To see this, consider the following special case of the model.[18] Suppose there are no trade and migration costs and 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, i.e. $\phi_M^R = \phi_A^I = 0$. This implies that $N_{rt} \propto \left(L_{rt}^I\right)^{\frac{1}{1-\lambda}}$ because total human capital in manufacturing is proportional to the number of industrial workers, i.e. $H_{rPt} \propto L_{rt}^I$. The distribution of workers along a SBGP is described by the spatial mobility condition $L_{rt}^I \propto (V_r w_{rMt})^{\varepsilon}$. Finally, good market clearing requires that $w_{rMt} L_{rt}^I \propto w_{rMt}^{1-\sigma} \left(Q_r N_r^{\vartheta}\right)^{1-\sigma}$.

These three equations fully summarize the long-run distribution of varieties N_{rt}, wages w_{rMt} and individuals L_{rt}^I as a function of the underlying fundamentals. In particular, upon defining[19]

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

it can be shown that

[18] See Section A-1.4 in the Appendix for details of the derivation.

[19] The restriction that $\zeta > 0$ is a stability condition similar to the restrictions derived in Allen and Arkolakis (2014) and Allen and Donaldson (2018).

$$\ln L_{rt}^{I} = \kappa_L + \zeta \times \ln Q_r + \frac{\sigma}{\sigma - 1}\zeta \times \ln V_r \tag{21}$$

$$\ln w_{rMt} = \kappa_w + \frac{\zeta}{\varepsilon} \ln Q_r + \zeta \left(\frac{\vartheta}{1-\lambda} - \frac{1}{\sigma - 1} \right) \times \ln V_r, \tag{22}$$

where and κ_L and κ_w summarize terms, which are constant across space. Equations (21) and (22) show how the long-run elasticity $\vartheta / (1-\lambda)$ amplifies existing differences in amenities V_r and productivity Q_r. Because ζ is increasing in $\vartheta / (1-\lambda)$, the endogenous nature of variety creation acts as an amplifying force. Interestingly, wages are *increasing* in local amenities if the long-run scale elasticity is large relative to the dispersion force of product differentiation governed by $1/(\sigma - 1)$. 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, making "nicer" places also high-paying locations.

The case of $\lambda = 1$ is qualitatively different. First of all, it is apparent from (19) that - generically - there does not exists a SBGP where all regions grow at the same rate as this would require the amount of human capital to be equalized across space. Hence, if $\lambda = 1$, the economy does not permit a stationary distribution of economy activity. The linear relationship between growth and the level of population is of course exactly the case of "strong scale effects", which is at the heart of most models of endogenous growth. Equation (19) therefore highlights the spatial analog 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.[20]

5 The Aggregate Effects of the Refugee Settlement

The model developed in the last section is a parsimonious framework to quantify the short- and long-run effects of population inflows in the context of a spatial model of growth. In this section I estimate the structural parameters of the theory by fitting the empirical results established in Section 3. This exercise has two main purposes. First, I show that the empirical results presented in Section 3 are quantitatively consistent with the theory. Secondly, the

[20] This result is due to the assumption that there are no other forms of spatial knowledge diffusion. If the local cost of variety creation depended on the number of varieties in other locations, it would be possible to have a stationary long-run population distribution and full endogenous growth. See Krugman (1979), Eaton and Kortum (1999) and Desmet et al. (2018).

model allows me quantify the effect of the refugee-settlement on aggregate income, which is not identified from the empirical findings alone.

5.1 Refugees and Variety Growth: Direct Evidence

The theory highlights the endogenous link between population size and productivity through the entry of new plants. Interestingly this mechanism 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)

The theory suggests to consider the relationship

$$\ln N_{rt} = \alpha_0 + \beta \ln L_{r1950} + \eta \ln N_{r1939} + x_r' \xi + u_{rt},$$

where $\ln L_{r1950}$ denotes the log of population density in 1950 and N_{rt} denote the number of plants at time t, which I can measure from the digitized historical manufacturing census files for 1933, 1939, 1950 and 1956. This specification is exactly the accumulation equation (9) if the total population L_{rt} was proportional to the number of production workers H_{rPt}. To estimate this specification empirically (see Section A-2.4 in the Appendix for details), I instrument the size of the local population L_{rt} with either the share of refugees in 1950 or the geographical expulsion distance Instrumenting for the size of the population is required if counties differ for example in their entry costs f_E as such variation would both induce a higher growth rate of the number of plants and population inflows.

For the short-run response in 1950, I find that $\beta = 0.68$, for the long-run response in 1956 I find $\beta = 1.1$. Hence, the long-run elasticity is much larger than the short-run elasticity, suggesting that $\lambda > 0$. Equation (18) also implies the coefficient η to be equal to $\lambda^{t-(1939-1)}$. I indeed estimate that η is below one and smaller in the long- compared to the short-run, which implies that $\lambda < 1$. This is consistent with the structural estimates below even though these estimates do not use the information on the dynamics of plant entry.

5.2 Estimation and Identification Strategy

The model is fully parametrized by 17 structural parameters and a tupel of fundamentals $[Q_r, V_r, T_r]$ per region. I calibrate 5 parameters externally and estimate the remaining 12 parameters within the context of this paper

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

My empirical strategy to identify the vector of structural parameters Ω is as follows. I estimate the knowledge elasticity λ and the elasticity of substitution ρ, the "Calvo"-mobility friction ψ, the parameters governing the productivity process (ϱ, ϖ) and two parameters of the skill distribution $\left(\chi, \phi_M^I\right)$ through indirect inference, i.e. by running the regressions of Section 3 on the data generated by the model. The remaining parameters $\left(\varepsilon, \kappa, \phi_A^I, \alpha, \xi\right)$ are calibrated to match five additional moments.[21] I identify the fundamentals $[Q_r, V_r, T_r]$ by calibrating the model to the cross-regional data on GDP per capita, sectoral employment shares and population size in 1933, which I assume to correspond to a steady-state of the system. Finally, the five parameters I set externally are the trade elasticity σ, the labor share in the two sectors γ and β, the dispersion of skills θ, which determines the elasticity of substitution across sectors, and the exogenous rate of firm exit δ.[22] For my baseline analysis I assume that $\sigma = 5$, $\gamma = 0.5$, $\beta = 0.6$, $\theta = 1.5$ and $\delta = 0.1$.

To replicate the historical experiment of the refugee settlement in the model, I simulate the dynamic evolution of the economy starting in 1933 and "shock" the economy with the inflow of refugees in the post-war period. Because the majority of refugees arrived around the year 1947, I assume that all refugees arrived in 1947 and I allocate them according to the empirically observed share of refugees in 1950.[23] Hence, the model - by construction - replicates the correlation between the share of refugees and population density, GDP per capita and sectoral employment shares in 1933 and is consistent with the identification assumption of the my OLS strategy that the share of refugees is uncorrelated with the history of regional productivity shocks u_{rt} (but correlated with spatial fundamentals Q_r, T_r or V_r). I then simulate the evolution of the economy until 1966 to calculate the regression coefficients and moments on the model-generated data. To account for the sampling variation induced by the stochastic productivity process, I replicate the analysis 30 times and calculate the average of all moments and regression coefficients.

For convenience I report all regression results I target via indirect inference in Table 9. The table summarizes the results reported earlier in Tables 4, 5 and 6 and also indicates the precise set of control variables. I always run the exact same specification in the model-generated data

[21] Here, ξ denotes the elasticity of trade costs with respect to distance, i.e. $\tau_{jr} = d_{jr}^{\xi}$.

[22] The fixed cost of entry f_E can be normalized by an appropriate choice of units for N_{rt}

[23] 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 ψ is small.

Dep. variable	$\ln y_{1950}$	$\ln y_{1961}$	s_{M1950}	μ_{1950}	$\ln pop_{1961}$	$\ln y_t$	ln earnings
Target	ρ	λ	ϕ_A^I	ψ	ϱ	ϱ	χ
Share of refugees	-0.062	0.465**	0.264***	0.601***	1.112***		
	(0.302)	(0.189)	(0.072)	(0.035)	(0.196)		
$\ln y_{r,t-1}$						0.555***	
						(0.041)	
Refugee							-0.098***
							(0.008)
N	523	519	535	484	488	1650	32584
R^2	0.722	0.508	0.863	0.733	0.912	0.894	0.316

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. For the definition of variables, see the notes for Tables 4, 5 and 6. The regressions control for log GPD pc in 1935 (cols 1 and 2), log population density 1939, wartime destruction and geographical controls (cols 1,2,3 and 5), the manufacturing employment share in 1939 (col 3), individual demographics (col 7), the log of population in 1939 (col 5), state fixed effects (cols 1-7) and county effects (col 6).

Table 9: Indirect Inference: Regression Evidence as Identified Moments

(with the exception of the control for the extent of wartime destruction, which does not have a model-counterpart). While all parameters are calibrated jointly, I now discuss the moments which mainly inform the different parameters in turn. See also Section A-2.5 in the Appendix for more details.

Growth & Scale: ρ and λ I estimate the two scale parameters ρ and λ by indirect inference using the short and the long-run relationship between refugee inflows and GDP per capita reported in columns 1 and 2 of Table 9. 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. If ρ is large, the effects of locally decreasing returns through the agricultural sector and elastic demand dominate and the effect of refugee inflows on GDP per capita is negative. The relationship between refugee inflows and long-run GDP per capita is increasing in λ, as λ amplifies the strength of the dynamic agglomeration force.

Spatial Mobility: ε, ψ and κ The Calvo parameter ψ governs the extensive margin of population mobility. The higher ψ, the faster will the distribution of refugees mimic the ones of natives, i.e. regions with an initial "excess" of refugees revert quickly to the mean. I therefore identify ψ by targeting the regression coefficient between the share of refugees in 1961 and in 1950 reported in column 4 of Table 9. My strategy to match the dispersion of idiosyncratic regional tastes ε is as follows. The higher ε, the more individuals agree about the attractiveness of a location and a region with *net* outflows will only see very few people moving into the particular location. Empirically, however, gross migration flows are much larger than net flows as "reverse" flows occur frequently. Even among regions that experience net outflows, the average share of inflows relative to outflows is 0.7, i.e. for every person that leaves a shrinking region, on average 0.7 new migrants appear. The ratio of inflows to outflows in region with net population losses is monotone in ε. The parameter κ governs the extent to which moving costs increase in distance: the higher κ, the more local are migration flows. The historical migration survey reports - for each county - the share of out-migrants that remains in their state. Because this share of within-state migration is monotonously increasing in κ, I calibrate this parameter to match the empirically observed value of 0.67.

Skill distribution: ϕ_M^I, ϕ_A^I and χ To estimate the average human capital of industrialists, ϕ_M^I and ϕ_A^I, I use two moments. First, I target the relationship between refugee inflows and the expansion of the manufacturing sector reported in column 3 of Table 9. The lower ϕ_A^I, the more will industrialist natives sort towards urban areas. The inflow of refugees will therefore trigger a larger increase in the local manufacturing sector. Secondly, I target the average earnings premium in manufacturing relative to agriculture, i.e. the "agricultural productivity gap" (Gollin et al., 2014). Holding ϕ_A^I fixed, ϕ_M^I increases relative human capital of industrialists and hence the measured agricultural gap. To estimate the extent of skill heterogeneity χ, I exploit differences in earnings between refugees and natives at the micro-level by estimating a Mincer-type regression of log earnings on demographics and a dummy for whether or not the individual is a refugee. Empirically, refugees earn about 9% less than natives on average (see column 7 in Table 9). Through the lens of the model, the only difference between refugees and natives is the extent of spatial sorting as industrial workers among natives sort towards manufacturing-intensive regions which offer high wages. The parameter χ is therefore an important determinant of the relative earnings between natives and refugees.

Productivity process: ϱ and ϖ To estimate the parameters of the productivity process of Q_{rt}, I use three moments. First of all, I aim to match the stochastic process of regional income per capita for the years 1957, 1961, 1964 and 1966 by running the panel regression $\ln y_{rt} = \delta_r + \beta \ln y_{rt-1} + v_{rt}$ in the model-generated data. Here δ_r is a county fixed effect. I target the coefficient β (column 6 of Table 9) and the dispersion in the estimated residuals, i.e. $sd(\hat{v}_{rt})$. In addition I also target the relationship between refugee inflows and long-run population growth between 1939 and 1961 (column 5 in Table 9).

Spatial fundamentals $[Q_r, V_r, T_r]$ and the agricultural expenditure share α To estimate the spatial fundamentals, I assume that the economy is in a steady-state in 1933 and calibrate $[Q_r, T_r, V_r]$ by matching the population distribution $\{L_{r1933}\}_r$, the sectoral employment shares $\left\{s^M_{r1933}\right\}_r$ and income per capita $\{y_{r1933}\}_r$.[24] Formally, for a given set of structural parameters, there is a one-to-one mapping between the three fundamentals and the three moments for each region. Note also that the steady-state implies a particular endogenous distribution of the number of varieties N_{r1933} and the extent of spatial sorting, i.e. the allocation of industrial types ω^I_{r1933} across counties. It can be shown that the observed aggregate employment share in agriculture identifies the level of α.[25]

Trade costs ζ I estimate the elasticity of trade costs with respect to distance ζ from the gravity relationship of within-country trade flows. Because I do not have access to historical trade flow data from Germany, I target the moment reported in Monte et al. (2018), who use data on shipments within the US and estimate a distance elasticity of -1.29.

5.3 Estimation Results and Model Fit

In Table 10 I report the estimated structural parameter (column 3) and the fit of the model (columns 5 and 6). The model is able to replicate the targeted moments quite well. In particular, it matches the fact that the short-run effect of the refugee-inflow on GDP per capita is negative (row 1) but the long-run effect is positive (row 2). It also matches the spatial persistence of refugee flows (row 5), the long-run effect of the refugee settlement on population growth (row 9) and the expansion of the local manufacturing sector (row 3). Regarding the process of spatial mobility, the model accounts for the fact that two thirds of outflows remain in the same state (row 7) and that bi-directional flows occur frequently (row 6).

In terms of the implied structural parameters, I estimate the inter-temporal knowledge elasticity λ to be 0.88 and the elasticity of substitution ρ to be 3.5. Hence, the short-run scale elasticity in the non-agricultural sector is equal to $\vartheta = \frac{1-\beta}{\rho-1} = 0.16$ and the long-run scale elasticity is $\frac{1}{1-\lambda} \approx 8$ times as large and equal equal to 1.3. To match the persistence of the refugee settlement, the model implies a moving hazard of $\psi = 0.079$.

Non-targeted Moments The model also matches a variety of non-targeted moments. For brevity, I only summarize this analysis here and refer to Section A-2.6 in the Appendix for details.

[24] In principle, these parameters could be identified even without the assumption of the economy to be in a steady-state in the pre-war period. This would however require at least two periods where income per capita was observed. I only have access to the data on GDP per capita at the county level for a single period prior to the war.

[25] Intuitively, if there was only a single region and no skill heterogeneity, the employment share would be equal to α and independent of sectoral productivities. A similar reasoning still holds true in the full model with skill heterogeneity, multiple regions and trade and migration costs.

	Panel A: Estimated structural parameters				
	Structural Parameter		Moment		
				Data	Model
ρ	Elasticity of substitution	3.502	Table 9, Col. 1	-0.062	-0.063
λ	Intertemp. knowledge elasticity	0.880	Table 9, Col. 2	0.465	0.489
ϕ_A^I	HC of industrialists in agric.	0.763	Table 9, Col. 3	0.264	0.302
ϕ_M^I	HC of industrialists in manuf.	14.741	Agricultural prod. gap	1.500	1.479
ψ	Frequency of mobility shocks	0.079	Table 9, Col. 4	0.601	0.564
ε	Spatial labor supply elasticity	2.580	Inflows/outflows in declining regions	0.73	0.834
κ	Elasticity of moving costs	-1.081	Share of outflows within own state	0.67	0.663
ϱ	Correlation of productivity	0.982	Table 9, Col. 5	1.112	1.270
			Table 9, Col. 6	0.555	0.577
ϖ	Dispersion of prod. shocks	0.038	Std. dev. of residuals, Table 9, Col. 6	0.045	0.033
χ	Share of industrial workers	0.242	Table 9, Col. 7	-0.098	-0.04
ζ	Elasticity of trade costs	0.322	Gravity relationship	1.29	1.29
α	Spending share on agriculture	0.239	Agricult. empl. share in 1933	36.7%	36.7%

Notes: The table reports the structural parameters (column 3) and the targeted moments both as observed empirically (column 5) and as implied by the model (column 6). All moments are targeted simultaneously by minimizing the function $\sum_{m=1}^{N_M} \omega_m \left(W_m^{Data} - W_m^{Model}\right)^2$,where $m = 1, ..., N_M$ denotes the respective moments, W_m^{Model} (W_m^{Data}) denotes the moment value in the model (data) and ω_m is a vector of weights of weighting the different moments. I weigh all moments equally. The agricultural productivity gap is from Gollin et al. (2014). The gravity relationship is taken from Monte et al. (2018). The exogenously set parameters are $\sigma = 5$, $\gamma = 0.5$, $\beta = 0.6$, $\theta = 1.5$ and $\delta = 0.1$. The time-invariant regional fundamentals $[V_r, Q_r, T_r]_r$ are calibrated to perfectly match the population distribution, the agricultural employment share and income per capita in 1933, which I treat as the allocation in a steady-state.

Table 10: Structural parameters

I first show that the model matches the cross-sectional relationship between refugee inflows and income per capita at different time horizons. While the model is calibrated to match the response in 1950 and 1961, it is statistically indistinguishable from the data for the whole time period 1950-1964. I also confront the models' spatial implications and show that the model is a good predictor for several non-targeted spatial moments: the size of the population in 1961, the share of refugees in 1955, the manufacturing employment share in 1961, the employment share of refugees in manufacturing relative to the one of natives within counties in 1950 and the two moments which I use to estimate the spatial labor supply elasticity ε and the distance elasticity of moving costs κ, i.e. the share of a region's outflows, which are within-state and the ratio of inflows relative to outflows in 1955.

Finally, I revisit a key feature of the historical experiment, namely the refugees' "manufacturing bias", i.e. the fact that the employment share in manufacturing among refugees exceeded the one for natives within a local labor market. According to the theory, this is consequence of spatial sorting, where in many rural locations the share of industrialists among refugees exceeded that for natives. Specifically, the theory implies that

$$\underbrace{\pi_{rMt}^{Ref} - \pi_{rMt}^{Nat}}_{\text{Bias}} = \underbrace{\left(\omega_{rt}^{I,Ref} - \omega_{rt}^{I,Nat}\right)}_{\text{Sorting}} \times \underbrace{\left(\pi_{rMt}^{I} - \pi_{rMt}^{F}\right)}_{\text{Comp. Advantage}}. \tag{23}$$

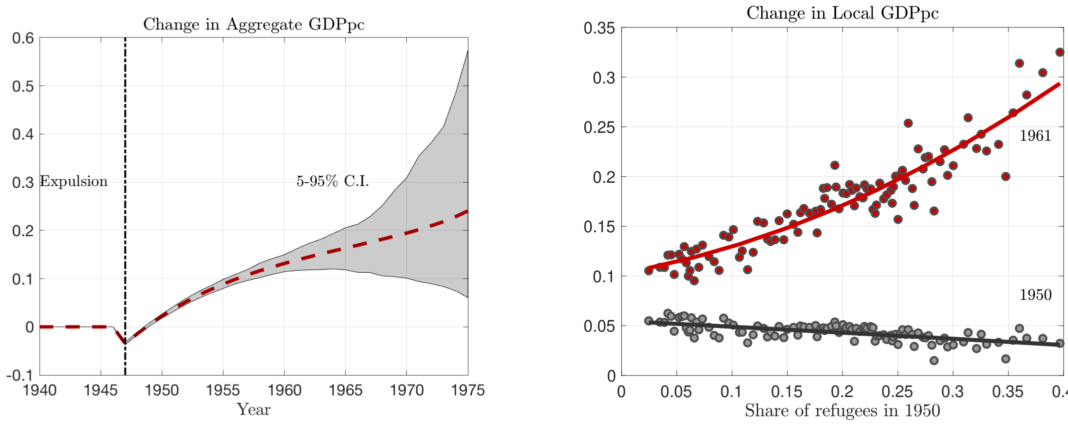
Given that $\pi_{rMt}^{I} > \pi_{rMt}^{F}$ as industrialists have a comparative advantage in manufacturing, the relative manufacturing bias of refugees reported in Table 8 (i.e. $\pi_{rMt}^{Ref} > \pi_{rMt}^{Nat}$) suggests that the share of industrialists among refugees exceeded the one for natives in many locations (i.e. $\omega_{rt}^{I,Ref} > \omega_{rt}^{I,Nat}$). In Section A-2.6 in the Appendix I show that the calibrated model is perfectly consistent with this implication. I first document that the share of skilled natives $\omega_{r}^{I,Nat}$ is strongly positively correlated with the pre-war manufacturing employment share, because skilled natives sort towards rich, manufacturing intensive locations in equilibrium. Given that refugees were predominantly sent to rural location, this implies that most refugees resided in locations where $\omega_{r}^{I,Nat}$ was low. (23) therefore predicts that the pre-war manufacturing share should be negatively correlated with the manufacturing bias of refugees in 1950. This is exactly what I find in the data. Finally, this also implies that the increase in the regional manufacturing employment share should be particularly large in rural locations where the incoming refugees were the rare provider of manufacturing human capital. I also find this to be the case. When I allow the relationship between the refugee settlement and manufacturing employment to vary by the pre-war manufacturing share, I find that the effect is much larger in locations that were predominantly rural prior to the war.

5.4 The Aggregate Effects of the Refugee-Settlements

How large was the aggregate impact of the refugee settlement on economic activity in West Germany? While this object is not identified from the cross-sectional regression due to the usual "missing-intercept" problem (see for example Adao et al. (2018), Chodorow-Reich (2019) and Wolf (2019)), it can be computed in the calibrated model by comparing the equilibrium with refugee inflows with a counterfactual West Germany where the refugees did not arrive.

In the left panel of Figure 4 I plot the time path of the percentage change in aggregate income per capita due to the refugee settlement. More precisely, for a given sequence of regional productivity shocks, I compute the effect of the refugee settlement on aggregate income. Calculating this experiment for a whole set of regional productivity shocks allows me to estimate the distribution of this aggregate impact and I plot both the average effect in red and a 90% confidence interval in grey.

The graph shows that the influx of refugees initially *reduces* GDP per capita by about 5%. This is mostly due to the fact that agricultural production is subject to decreasing returns.



Note: The figure displays the aggregate treatment effect of the refugee inflow. Specifically, I calculate aggregate GDP per capita for the model with refugee inflows relative to GDP per capita 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 4: The Aggregate and Spatial Impacts of the Refugee Settlement

Due to the endogenous nature of technological progress this initial drop is short-lived and the population increase causes income per capita to increase. Given the estimated parameters, the average effect rises to about 15% in 1961 and reaches 24% by 1975.

To put these numbers into perspective, suppose there was only a single region and there were no exogenous productivity shocks. As I show in Section A-1.6 of the Appendix, the elasticity of long-run income per capita with respect to population size is given by

$$\frac{d\ln y_r}{d\ln L_t} = \underbrace{(1-\alpha)\frac{\vartheta}{1-\lambda}}_{\text{Increasing returns in manufacturing}} - \underbrace{\alpha\gamma}_{\text{Decreasing returns in agriculture}}.$$

Hence, the aggregate scale elasticity is an $\alpha-$weighted average between the long-run scale elasticity in manufacturing $\vartheta/(1-\lambda)$, which is positive, and the returns to scale in agriculture $-\gamma$, which are negative. Using the estimated parameters in Table 10, this expressions suggests an elasticity of 0.89 in the long-run, i.e. the inflow of refugees should have increased income per capita by about 22% in the long-run.

Note that the cross-sectional estimates reported in Table 6 provide a misleading answer for the aggregate impact of the refugee settlement shown in Figure 4. In particular, they underestimate the aggregate effect. Not only is the cross-sectional estimate between refugee inflows and GDP per capita in 1950 negative (even though the aggregate effect is positive) but the implied long-run estimates are also downward biased. The point estimate of 0.465 in Table 9 for example suggests that a 20% increase in the share of refugees increases GDP per capita by $0.2 \times 0.465 = 9\%$ in 1961. This is not only smaller than the average aggregate impact of about 15% but in fact outside the confidence intervals.

The reason is of course that non-treated regions also benefitted from the refugee inflow in general equilibrium. This is shown in the right panel of Figure 4, where I depict the correlation between the counterfactual percentage change in income per capita and the share of refugees, both in 1950 (dark dots) and 1961 (red dots).[26] In 1950 there is a negative correlation, in line with the negative cross-sectional estimate. Crucially, however, the entire locus is positive on average, i.e. non-treated regions benefit from the increase in the demand for their products from other counties. It is this general equilibrium effect, which is differenced out in the empirical cross-sectional estimates. While this makes these estimates not directly informative to quantify the aggregate impact of the refugee settlement, they are still useful moments to estimate the structural parameters.

If we fast-forward by a decade and look at the impact on income per capita in 1961 (red dots), we see a very different picture. First of all, the relationship between the local refugee settlement and income per capita is now strongly positive, reflecting the slow accumulation of technology whereby past market size positively affects local productivity. Empirically, this slope reflects the positive cross-sectional relationship between refugee inflows and income per capita in the long-run. Secondly, the entire locus is shifted upwards because regions that were initially non-treated benefit both from refugees' migration response and share in the aggregate gains through trade linkages. Of course, in the very long-run, the correlation between income per capita and refugee inflows in 1950 will be zero as the long-run distribution is stationary and fully determined from fundamentals. This also implies that eventually all regions benefit from the inflow proportionally.

Interestingly, the effect on manufacturing employment is qualitatively different. Both in the data (see Table 5) and in the model (see Section A-2.7 in the Appendix) there is a positive cross-sectional relationship in 1950 and 1961. Even though the manufacturing sector is the sole source of endogenous productivity gains, an increase in the employment share does not necessarily go hand-in-hand with an increase in income per capita if local technology accumulates slowly. In the short-run, the manufacturing sector absorbs the rising population for a given level of technology. This reduces productivity in the manufacturing sector due to selection.[27] In the long-run, a larger manufacturing labor force triggers entry and hence an

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

[27] Suppose for simplicity that there is no type heterogeneity and that H_{rPt} is proportional to H_{rMt} (which is the case along the BGP). Letting π_{rMt} denote the employment share in the manufacturing sector, total sectoral output per manufacturing worker is given by $\frac{Y_{rMt}}{L_{rt}\pi_{rMt}} \propto Q_{rt}N_{rt-1}^{\lambda\vartheta}\phi_M^{\frac{1+\vartheta}{\theta}}L_{rt}^{\vartheta}\pi_{rMt}^{\frac{\vartheta(\theta-1)-1}{\theta}}$. While a bigger population L_{rt} increases output per workers, the effect of a larger employment share π_{rMt} depends on the relative strength between agglomeration forces and the deterioration of average human capital through selection. The latter

upward shift in the labor demand schedule through the dynamic increase in productivity.

6 Conclusion

The positive relationship between population size and productivity is at the heart of virtually all theories of economic growth. In this paper I analyzed a particular historical setting to provide direct evidence for the empirical relevance of such scale effects. I focused on the expulsion of the ethnic German population and their subsequent resettlement in West Germany in the aftermath of the Second World War. As part of the post-war restructuring of continental Europe, the Military Governments of the US, the UK and Russia decided to expel the German population from their domiciles in Eastern Europe and transferred them to West Germany and the Soviet Occupied Zone. Between 1945 and 1948 almost 8m people arrived in West Germany. At the time, this amounted to an increase in the population by about 20%.

Because regions in West Germany differed substantially in the extent to which they were exposed to the refugee settlement, I am able to use the cross-sectional variation in refugee inflows to estimate the relationship between changes in population size and income per capita in both the short- and the long-run. My finding that refugee inflows increased income per capita is consistent with models that feature an endogenous response of local technology to population size, but hard to reconcile with a neoclassical framework, where productivity is exogenous.

I then propose a parsimonious idea-based model of spatial growth that can rationalize this finding both qualitatively and quantitatively. The model highlights the interplay between frictional mobility and local growth. In particular, it can deliver a persistent and protracted increase in local productivity if spatial mobility is subject to frictions and there are dynamic spillovers, whereby the existing number of varieties reduce the cost of future variety creation. The theory also stresses the distinction between short- and long-run scale elasticities, as the dynamic accumulation of local productivity acts as an amplifying force.

I estimate the structural parameters of my theory via indirect inference, using the cross-sectional regression results of the natural experiment as identified moments. I find that the dynamic amplification is sizable and that the long-run scale elasticity is 8 times as large as the short-run elasticity. At the aggregate level, the settlement of refugees increased income per capita by about 25% after 25 years. This is in stark contrast with the short-run effect, which lowered income per capita by 5%. Importantly, a naïve extrapolation of the cross-sectional relationship between refugee inflows and local income per capita would have substantially

negative effect dominates if $\vartheta < \frac{1}{\theta-1}$. According to my estimates, this inequality is satisfied as $\vartheta = \frac{1-\beta}{\rho-1} = 0.2$ and $\theta = 2$.

underestimated the aggregate effects as non-treated regions benefitted from the refugee inflow through migration and trade linkages.

A natural question is of course whether these results from this particular historical episode are quantitatively portable to predict the consequences of immigration episodes today. While I expect the basic mechanism to apply more generally, there are at least three aspects of this historical setting that seem particularly context-specific. First and foremost, the German economy just emerged from the Second World War. Hence, there is reason to suspect that new capital might have been particularly mobile across space. Secondly, the German refugees were allocated to rural labor markets and not to urban centers. This is of course very different from most voluntary migration episodes both in the modern era and in the past. Finally, the 1950s and 1960s were characterized by a secular rise in the manufacturing sector. To the extent that the agglomeration mechanisms highlighted in this paper are more potent in the manufacturing sector than in the service industries, the productivity effects of changes in population size might be smaller today.

References

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

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

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito, "Spatial Linkages, Global Shocks, and Labor Markets: Theory and Evidence," 2018.

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," 2018. Working Paper.

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

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.

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

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

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

Card, David, "The Impact of the Marial Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, 1990.

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

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

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

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

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment," *The Quarterly Journal of Economics*, 2016.

Eaton, Jonathan and Samuel Kortum, "International technology diffusion: Theory and measurement," *International Economic Review*, 1999, *40* (3), 537–570.

Eckert, Fabian and Michael Peters, "Spatial Structural Change," 2016. Working Paper.

Faber, Benjamin and Cecile Gaubert, "Tourism and economic development: evidence from Mexico's coastline," *American Economic Review*, 2019, *109* (6), 2245–93.

Fajgelbaum, Pablo 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.

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

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

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

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

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

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

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

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

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

Krugman, Paul, "A Model of Innovation, Technology Transfer, and the World Distribution of Income," *Journal of political Economy*, 1979, *87* (2), 253–66.

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

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

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

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

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

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

Peters, Michael and Conor Walsh, "Declining Dynamism, Increasing Markups and Missing Growth: The Role of The Labor Force," 2019. Working Paper.

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

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

Redding, Stephen J and Daniel M Sturm, "The costs of remoteness: Evidence from German division and reunification," *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.

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

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

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

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