Refugees and Local Agglomeration - Evidence from Germany's Post-War Population Expulsions

Michael Peters*

March 2017 [Preliminary]

Abstract

Can increases in the size of the local population raise productivity and spur economic development? This paper uses a particular historical episode to study this question empirically. After the Second World War, between 1945 and 1948, about 12m Ethnic Germans were expelled from their domiciles in Middle and Eastern Europe and transferred to Western Germany. At the time, this inflow amounted to almost 20% of the Western German population. Moreover, there are vast cross-sectional differences in the extent to which refugees were allocated to individual counties. In this paper I use this cross-sectional variation to study the effects of the inflow of refugees on Germany's regional economic development between 1950 and 1970. I find that refugee-inflows are positively correlated with income per capita, overall manufacturing employment and the entry of new plants. At the same time, refugees' earnings were substantially lower. Using a simple general equilibrium trade model I show that these patterns are consistent with theories of local agglomeration and endogenous technological change but hard to rationalize in a neoclassical framework with exogenous technology.

*eMail: m.peters@yale.edu. I am very grateful to Abhijit Banerjee, Penny Goldberg, Tarek Hassan, Virgiliu Midrigan, Gianluca Violante, Joachim Voth and especially Alwyn Young for their comments. I also thank seminar participants at Chicago Booth, Fordham, Harvard, MIT, NYU, Princeton, UBC, Wharton and Yale and the staff at the GESIS Institute in Mannheim, especially Bernhard Schimpl-Neimanns, for their hospitality and help in accessing the historical microdata of the 1971 census.

1 Introduction

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

In contrast to this large body of theoretical work there is relatively little direct empirical work on the importance of such mechanisms. This is due to a fundamental empirical challenge. As many of the theories above stress the importance of general equilibrium effects or aggregate consequences of agglomeration, which are external to the individual firm, one needs variation in factor supplies, which is both large and at the same time uncorrelated with other determinants of technological change. In this paper, I analyze a particular historical episode, that generates spatial variation in factor supplies and arguably satisfies both requirements.

At the end of the Second World War, during the Potsdam Conference, the Governments of the US, the UK and Russia decided to expel about 12m Ethnic Germans from their domiciles in Middle and Eastern Europe and transfer them to both Western Germany and the Soviet Occupied Zone. The ensuing expulsion was implemented between 1945 and 1948 and represents one of the largest population movements in world history. By 1950, about 8m people had been transferred to Western Germany. Given the population at the time, this amounted to an increase in the total population of Germany by about 20%. Moreover, there is substantial heterogeneity in the extent to which the inpouring refugees settled in different region. While some counties see their population almost double, other counties were far less affected.

In this paper I exploit this cross-sectional variation in refugee inflows across the 500 counties in Western Germany to study the link between population inflows, endogenous productivity responses and regional economic development. To guide the empirical analysis, I develop a simple two-sector general equilibrium trade model. Workers face a Roy-type occupational choice problem so that sectoral labor supplies depend on both relative wages and workers' relative skills. The population is comprised of refugees and natives, which differ in their skill endowments. While the agricultural sector uses a fixed factor ("land") and is subject to decreasing returns, productivity in the manufacturing sector is endogenous because the number of active firms responds to the size of the manufacturing workforce. I refer to the strength of this endogenous productivity response as the strength of agglomeration. Regions differ in their innate manufacturing productivity, their land supply, the human capital of their population and - crucially - in the number of refugees in the population. The model makes tight predictions on the spatial co-movement between refugees, income per capita, manufacturing employment, the number of active firms and sectoral employment patterns of refugees and natives. In particular, the qualitative properties of the model depend crucially on the strength of agglomeration.

I study the empirically relevant case where refugees have a comparative advantage in the manufacturing sector, but a lower level of human capital (and hence lower earnings) than natives. In that context I show that if agglomeration effects are sufficiently strong, a higher share of refugees in the population (i) increases manufacturing employment, (ii) induces new firms to enter, (iii) causes an increase in the local manufacturing shares of native workers (i.e. a "crowding-in" of native employment) and - most importantly - (iv) increases income per capita.

The reason why an inflow of refugees might *increase* income per capita despite the *decline* in average regional human capital is the interaction between firms' entry decisions and workers' occupational choice. As firms' do not capture the entire social surplus of their innovation (i.e. entry) expenses, equilibrium wages in the manufacturing sector are too *low* from a social point of view. A decline in agricultural human capital will therefore shift labor supply to the manufacturing sector and can hence be beneficial. If, in contrast, productivity was exogenous, a higher share of refugees will - given their lower average skills reflected in their lower earnings - always *reduce* income per capita. Moreover, refugees would tend to crowd-out native workers from the manufacturing sector.

Two features of the historical episode make the empirical variation extremely suitable to test these implications of the theory. The first concerns the determinants of the spatial distribution of refugees. With millions of refugees being transferred to the country, the Western German population in 1950 actually exceeded its pre-war level by about 13%. At the same time, the Allied bombing campaign had reduced the housing stock by almost 25% on average and in many cities by more than 90%. Hence, the dominant consideration for the Military Governments of the US and the UK to allocate the inpouring refugees across counties was the available housing supply. Moreover, the Military Governments implemented tight mobility controls, which essentially ruled out refugees' spatial adjustment before 1950.

These aspects of the historical context suggest that (i) refugees were mostly settled in rural areas with a more abundant housing supply, that (ii) conditional on these determinants of housing supply, the allocation of refugees was unrelated to other regional fundamentals stressed in the theory (like manufacturing productivity or the supply of human capital) and that (iii) the physical distance to the pre-war population centers of the expulsion regions remained a powerful determinant of the allocation of refugees until 1950. I use these insights to construct two complementary empirical strategies to tease out the exogenous component of the initial refugee allocation. I first use information on pre-war population density and a - to the best of my knowledge - novel data set on the extent of war-time destruction for all 500 German counties to control for the political allocation rule in a OLS strategy. I verify that, conditional on this rule, the refugee allocation is indeed uncorrelated with a host of measures of pre-war economic development, which should reflect regional productivity or human capital. I also consider an instrumental variable strategy, where I exploit the distance to the pre-war population centers in Eastern and Central Europe.

The second aspect of this empirical setting is equally important. In order to identify an endogenous productivity response of the local manufacturing industry, it is necessary to exploit a source of variation, which actually affects equilibrium labor supply in the manufacturing sector. Moreover, such variation should not *increase* the supply of average human capital, as this would raise GDP per capita even in the absence of a productivity response. The historical context of the expulsion of the ethnic Germans satisfies both these requirements. First of all, I show that refugees were much more likely to work in the manufacturing sector than natives, that is refugees had a comparative advantage in manufacturing. Secondly, I present direct micro data evidence from the early 60s that refugees had indeed lower earnings than natives. This is consistent with refugees being - on average - less skilled than natives.

To perform my analysis, I use a variety of datasets. Most importantly, I exploit novel historical data on the regional development of Germany between 1933 and 1970. In contrast to many other countries there are no surviving records of the historical micro census data with sufficient regional breath to calculate outcomes at the level of the roughly 500 Western German counties. However, the local statistical offices did publish summary statistics of the respective census at the county-level at the time, which I was able to digitize. For the years 1933, 1939, 1950, 1961 and 1970 I digitized the population census to measure sectoral and occupational employment shares, sex-ratios, population density and other characteristics. I then augmented this dataset with information on the allocation of refugees, on the extent of war destruction from the county-level results of the housing census, on regional GDP in the 1950s and 60s and measures of plant entry from the 1933, 1939 and 1956 waves of the German census of manufacturers.

Using this data, I show that local productivity did increase in response to the inflow of refugees. Most importantly, I show a positive relationship between refugee inflows in 1950 and local GDP per capita in 1961. In the theory, this positive relationship is driven by an expansion of the manufacturing sector and an increase in firm entry. I confirm that both these predictions are borne out in the data. In particular, refugee-receiving counties have higher manufacturing employment shares in the 1950 and 1960s and experienced higher rates of plant entry between 1939 and 1956. Additionally, I find no evidence of refugees crowding out native employment - if anything, the relationship between regional refugee inflows and native manufacturing employment is positive. Together with refugees' lower earnings, these patterns are consistent with an endogenous local productivity but hard to rationalize in a neoclassical framework with constant technologies.

Finally, I also present direct micro evidence on why refugees' labor supply was biased towards the manufacturing sector, i.e. on the source of their comparative advantage. More specifically, I exploit a special supplement to the census conducted in 1971 that aimed to measure the extent of social and economic mobility of the German population. The data contains retrospective information about employment characteristics in 1939, 1950, 1960 and 1971 for about 200.000 individuals and explicitly identifies refugees. Using this data I can measure snapshots of refugees' and natives' employment life-cycle pre- and post-expulsion. For refugees, I find a drastic reallocation from self-employed, agricultural work into unskilled occupations in the manufacturing sector after the expulsion. No such changes are observed for the native population. My preferred interpretation of the evidence is that refugees did not necessarily have superior skills in manufacturing but that they faced barriers to work in agriculture. As the agricultural sector in Germany was widely dominated by small, family farms, such barriers took plausibly the form of frictions in the agricultural land market. Using data on the distribution of farm size and occupational employment patterns within the agricultural sector corroborates this interpretation.

Related Literature On the theoretical side, the paper is related to a large literature in economic growth, which argues that innovation incentives' respond to changes in factor supplies. While this is true for many models of growth (e.g. the basic Romer (1990) model), this reasoning is at the heart of the literature on directed technological change and the bias of innovation (see e.g. Acemoglu (2002, 2007, 2010)), the relationship between economic integration and growth (Rivera-Batiz and Romer, 1991)) or the interaction between market size and specialization (Krugman, 1980a). Empirically, Hanlon (2015) also uses historical data to test for the prevalence of directed technological change. He uses the blockade of US-UK trade during the US Civil War and the resulting drop in the aggregate supply of US cotton to study firms' incentives in the UK to adoption technologies, which are biased towards other varieties of cotton. In contrast to this paper, Hanlon (2015) does not focus on the implications on income per capita.

The paper is also related to the recent literature on model of economic geography. Of particular relevance are the papers by Desmet et al. (2015), Desmet and Rossi-Hansberg (2014) and Nagy (2016), all of which present growth models with a realistic geography, where local innovation incentives (and hence productivity) do respond to local factor supplies. These models are therefore consistent with the empirical findings of this paper. At a more reduced-form level, my findings are also consistent with a large static literature on economic geography, which posits the existence of exogenous agglomeration economies - see for example Fajgelbaum and Redding (2014); Ahlfeldt et al. (2015); Allen and Arkolakis (2014) or the recent survey by Redding and Rossi-Hansberg (2016).

Finally, there is a large literature, which uses the German context as a source of historical experiments. Of particular relevance is Burchardi and Hassan (2013), who use a related source of variation. They use the settlement of refugees coming from the Soviet Occupied Zone and the interaction with the fall of the Berlin Wall in 1989 to measure the importance of social ties. The current paper is different. First of all, I look at outcomes in the 1950s and 1960s.[1] Secondly, I particularly focus on the evolution of local productivity and GDP per capita as a function of

[1] I therefore also rely on a different identification strategy. Burchardi and Hassan (2013) use the distribution of wartime destruction as an instrument for the settlement of refugees leaving the Soviet Occupied Zone during the 1950s. As the extent of wartime destruction is likely to directly affect manufacturing output in 1950, I do not use it for the allocation or refugees. However, when I use it as an

local labor supply. Finally, I focus on a different group of refugees. Burchardi and Hassan (2013) focus on refugees from Eastern Europe who were *first* sent to the Soviet Occupied Zone and *then* left for Western Germany. I in contrast *only* focus on the refugees from the East, who were directly sent to Western Germany, allocated according to the available housing stock and subject to migration restrictions.

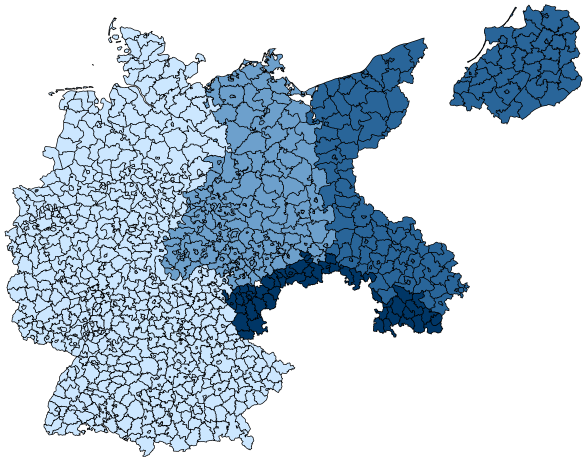
Other papers using German history as a source of variation include Ahlfeldt et al. (2015), who exploit the partition of Berlin as a shock to the distribution of economic activity, Redding and Sturm (2008), who use the Division of Germany as a shifter in market access and Fuchs-Schündeln and Schündeln (2005), who exploit the distribution of occupational patterns at the time of the German reunification to generate variation in income risk to test for the importance of pre-cautionary savings. On a methodological note, my paper is related to a small but growing literature, which uses natural experiments in macroeconomics - see Fuchs-Schündeln and Hassan (2015) for a recent survey.

Given the historical setting and the empirical strategy, there is also a close connection to the literature on immigration. In a recent paper Burstein et al. (2017) study the effects of immigration on native employment patterns within occupations. They also stress the role of tradability to determine wether natives are crowded-in or crowded-out through immigration. They, however, do not focus on possibility of immigration affecting local productivity. In a classic study, Card (1990) used the unexpected shock of the Miami-Boatlift to study the effect of Cuban immigrants on the labor market in Miami. This paper and many other papers in that literature (see e.g. Peri (2016); Dustmann et al. (2016)) are mainly concerned with the short-run impact of immigrants on wages and employment prospects of natives. Not only do I focus entirely on the longer-run outcomes, but I am also mostly interested in comparing average outcomes (like employment shares and income per capita) across regions, instead of relative wages within regions. In a recent paper, Akcigit et al. (2017) also relate the location choice of US immigrants in the 19th century to measures of innovation. They stress a different mechanism in that they focus on the innovation potential of the inflowing immigrants themselves. In my context, refugees were not the main source of new ideas.[2] Instead they encouraged firm entry through an increase in market size. See also Nunn et al. (2017), who study the long-run effects of immigration in the US and Hornung (2014), who uses data on textile plants to analyze the productivity effects of the Huguenot re-settlement for the 18th century. Finally, this historical setting has also been analyzed in Braun and Mahmoud (2014) and Braun and Kvasnicka (2014). In contrast to my paper, these contributions do not focus on the effect of refugee inflows on local productivity and also do not use the spatial allocations across counties.

The remainder of the paper is structured as follows. In the next section I describe the historical setting and the political environment leading to the population expulsions. In Section 3 I present the theory to link refugee flows and the endogenous productivity response and to derive the empirical implications. Section 4 contains the empirical analysis. I describe in detail the determinants of the initial allocation and my empirical strategy. I then analyze the relationship between the inflow of refugees, income per capital, manufacturing employment and the entry of new plants. In Section 5 I use the historical micro data on life-cycle employment patterns of refugees and natives to present further evidence on the mechanism, in particular on prevailing frictions in the agricultural market. Section 6 concludes.

instrument for outcomes in the 1960s, I find very similar results then when using my identification strategy.

[2] While there are of course individual instances of refugees brining their entrepreneurial capital to Western Germany, I present direct evidence that this effect is unlikely to be quantitatively important in my context.



Notes: The figure shows the German Reich in the boundaries of 1939. The light shaded part in the west is the area of to-be Wester Germany. The darker shaded part in the middle is the area of the to-be GDR. The medium-blue 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 annexxed by Germany in 1938. During the Potsdam Conference in 1945, Germany lost the Eastern Territories of the German Reich and the Sudentenland.

Figure 1: The German Reich in 1939

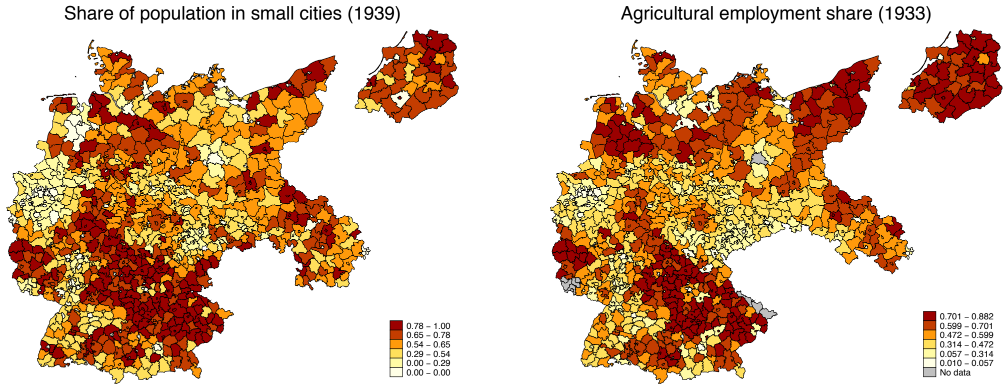
2 The Historical Setting

Germans in Eastern and Middle Europe before 1939

The presence of Germans in Middle and Eastern Europe is by no means a novel phenomenon. In fact, the settlement of ethnic Germans in Eastern Europe dates back to the Middle Ages. At the beginning of the Second World War in the summer of 1939, there are two groups to distinguish. On the one hand, there are large parts of todays' Poland and Russia, which used to be part of the German Reich. This encompasses for example the regions of East Prussia and Silesia. On the other hand, there were vast German minorities in other countries of Eastern Europe, most importantly the so-called Sudetenland in Czechoslovakia. This region in the north of Czechoslovakia has a long tradition of German settlements and was annexed by the Nazi Government in 1938.

To see that more clearly, consider the map shown in Figure 1. The map shows the territory of the German Reich on the eve of the Second World War, in the summer of 1939. To get a sense of the economic geography, I also display the individual counties, which is the source of cross-sectional variation I will be using for this paper. In the West, shown with a light shade, are the territories, which are going to become West Germany in 1949. These regions form the main part of the analysis in this paper, as I will be measuring post-war outcomes in the 50s and 60s in these regions. In 1939, roughly 38m people live in these areas. In the far East, shown in medium blue, are the "Eastern Territories of the German Reich". This is the part of the German Reich, which will no longer be part of Germany after 1945. These regions were home to roughly 10m people in 1939. In the south-east, shown in dark blue, is the aforementioned Sudetenland in the north of Czechoslovakia. According to the German Census in 1939, roughly 3m Germans were living there in 1939. Finally, in the middle is the area of the the German Reich, which will become the Soviet Occupied Zone (in 1945) and then turn into the German Democratic Republic (in 1949). This area will not be part of the analysis in this paper. Not shown on the map, there are are additional smaller German minorities living in other countries in Eastern Europe, in particular Poland, Hungary and Romania - see Table 12 in the Appendix.

To get a sense of the economic geography in 1939, consider Figure 2. On the map on the left I depict a measure of urbanization in 1939, namely the share of the county population living in cities with less than 2000 inhabitants. The



Notes: The left map shows a measure of urbanization in 1939, namely the share of the county population living in cities with less than 2000 inhabitants. The map on the right displays the agricultural employment share in 1933.

Figure 2: Economic Development Pre-War: Urbanization and Agricultural Employment

map on the right depicts the distribution of regional agricultural employment share in 1933. It is clearly seen that there are systematic regional differences between West Germany and the Eastern Territories. While the Eastern Territories are often rural and hence agricultural intensive, one can also see the industrialized, densely population part in the Ruhr-region in central- and western Germany. This is seen more clearly in Table 1, which compares the population in West Germany and the Eastern Territories according to various economic characteristics in 1939. Panel A contains the educational characteristics of the two populations. It is seen that the distribution of formal skills was very similar. The only slight difference is the higher popularity of vocational schools (*Berufsschule*) in pre-war Western Germany, a fact that is due to the bigger importance of the manufacturing industry (see below). This is no longer true when we consider the sectoral structure of employment, given in Panel B. While the reliance on services and the public sector is very similar, there is a large disparity in the employment shares of manufacturing and agriculture. In particular, the agricultural sector is roughly 60% bigger in the Eastern Territories, with a commensurate smaller manufacturing sector. This is very consistent with the regions in the East having higher agricultural employment shares in 1939. Finally, Panel C depicts the distribution across occupational classes, which again is very similar.

The Potsdam Conference in 1945

The Second World War marks a drastic change in the geography of Europe and Germany in particular. Germany was not only divided into the four Allied Occupation Zones, but also lost a substantial part of its landmass in Eastern Europe as means of war reparations. Specifically, the Sudetenland was returned to the Czech Republic and the Eastern Territories were allocated to both Poland and Russia respectively. Int that process, the governments of Russia, the UK and US also decided to expel the German population from these territories. The official protocol of the conference reads:

> "The Three Governments, having considered the question in all its aspects, recognize that the transfer to Germany of German populations, or elements thereof, remaining in Poland, Czechoslovakia and Hungary, will have to be undertaken. They agree that any transfers that take place should be effected in an orderly and humane manner."

The subsequent population transfer is one of the largest transfers in world history. Between 1946 and 1950, roughly 12 million ethnic Germans were expelled and 8 million people were allocated to Western Germany (Reichling, 1958,

	West Germany	Eastern Territories
	Educational Attainment	
Elementary School	66.3	65.9
High School	8.3	11
Vocational School	18.4	15.5
College	6.8	7.6
	Sectoral composition	
Agriculture	14.4	22.2
Manufacturing	52.6	43.1
Services	18.3	17
Public Sector	14.5	17.6
	Occupational composition	
Self-employed (Agriculture)	10.3	12.3
Skilled Employee	7.7	8.4
Unskilled Employee	7.9	8.5
Skilled Worker	3.6	2.9
Unskilled Worker	23.8	21.8

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

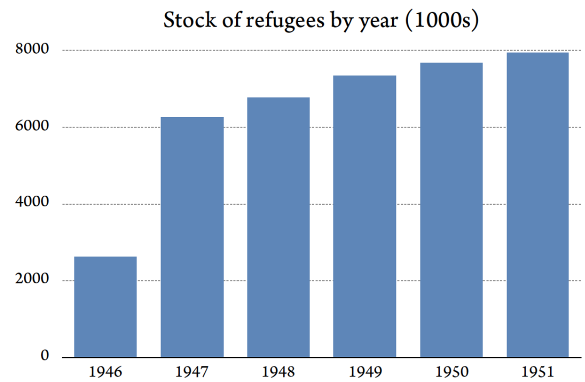
Table 1: Economic Characteristics in 1939

p. 17).

The expulsion can be broadly divided into three phases. The first wave of refugees arrived in Western Germany during the last months of the war. Soviet forces made their appearance at the eastern German border in the summer of 1944. Trying to reach Berlin, soviet soldiers were advancing through the German Eastern Territories at great speed causing the German population to flee westwards. As the Nazi government considered the evacuation of German territories a defeatist act and executed a strict "no retreat" policy to use the civil population as a shield slowing down the Russian army, most inhabitants evacuated their homes fully unprepared. Because there were hardly any official evacuation plans as trains and ships were often reserved for the German soldiers, most refugees fled their homes by joining refugee treks, which suffered enormous casualties during the flight (de Zayas, 1993). After the German defeat in May 1945, the so-called wild expulsions started. These where mainly taking place in the spring and summer of 1945 before the Potsdam Agreement was signed in August 1945, most importantly in Poland and Czechoslovakia, where a substantial German minority resided. Under the backing of the respective governments, both the army and privately organized militias started to systematically expel the German population. It is only the Potsdam agreement which tried put an end to these unorganized expulsions and legalized them ex-post. Within the following two years, the majority of the German population was transferred from Middle- and Eastern Europe to Western Germany and the Soviet Occupied Zone. The timing of the arrival in Western German is depicted in Figure 3 below. It is clearly seen that the vast majority of the population transfer takes place in the two years immediately following the war. By 1948 almost 7m expellees were already present in Western Germany.[3] This amounted to roughly 20% of the population living in Western Germany at the time. Despite the casualties during the war, the population of Western Germany had therefore increased from 42m people in 1939 to 1950 (Steinberg, 1991).

In this paper I will exploit the cross-sectional variation in refugee flows across counties in Western Germany.

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



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

Figure 3: Expellees' arrival in Western Germany

In Figure 4 I show the allocation of refugees across Western German counties in 1950. Two observations stand out. First of all, for many localities, the initial shock is very large. There are many counties, where the share of refugees in the county population exceeds 30%. Second of all, the variation across counties is also sizable. We see a clear East-West trajectory. This is not surprising as the flow of refugees arrived from the East. There are also two important "centers" of refugee destinations in the north (in the states Schleswig-Holstein and Lower Saxony) and in the South (in Bavaria). Again, this has geographical reasons. Many expellees from the Easter Territories arrived in Western Germany via a northern route along the coast of the Baltic Sea and hence arrived in Western Germany in the north. Similarly, the expellees from the southern parts of Eastern Europe (most importantly the Sudeten) arrived in Bavaria and therefore settled there.

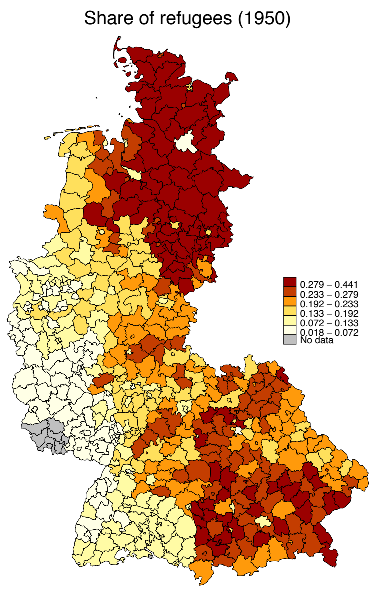
A crucial part of the empirical analysis will naturally rely on the determinants of the cross-sectional variation shown in Figure 4, in particular the extent to which the allocation of refugees was correlated systematically with differences across counties. As I will show below, there is ample evidence that this variation was not random. Hence, in my empirical strategy I will consider different approaches to tease out the direct effect of refugee inflows. To guide the empirical analysis and to appropriately address the endogeneity problem, I will now construct a simple general equilibrium model, which links the spatial allocation of refugees to the extent of industrialization and income per capita.

3 Theory

In this section I lay out a simple theoretical framework to study the effect of local labor supply shocks on local GDP per capita and the size of the local manufacturing sector. The model not only highlights the theoretical interaction between the supply of human capital and the endogenous response of productivity but I will also derive the empirical regression equations (and the endogeneity problem) from the model.

3.1 Technology and Preferences

I consider a simple static model of inter-regional trade, where each region $r = 1, .., R$ will correspond to a county in the empirical analysis below. For simplicity I assume that there are no trade costs. Consumers have standard preferences over an agricultural and a manufacturing good, each of which is a CES aggregate of the different R



Notes: This figure depicts the share of refugees (relative to the entire population) for each county of Western Germany in 1950. Counties are harmonized at the level of 1975. See Section 8.6 in the Online Appendix. Source: Statistisches Bundesamt (1955b)

Figure 4: The Allocation of Refugees in Wester Germany: 1950

local varieties. Hence,

$$u = u\left(C_A, C_M\right) = C_A^{\alpha} C_M^{1-\alpha} \text{ where } C_s = \left(\sum_r c_{s,r}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where $c_{s,r}$ is the total amount of consumption of region r's variety in sector s.

The agricultural good is produced using labor and land, i.e.

$$Y_{r,A} = H_{r,A}^{\gamma} T_r^{1-\gamma}, \tag{2}$$

where T_r is the amount of land in county r and $H_{r,A}$ is the total amount of labor (as measured in efficiency units) used in agricultural production in region r. The output elasticity of land γ determines the degree of decreasing returns as T_r is in fixed supply. In contrast, the manufacturing good is potentially subject to increasing returns and hence the source of endogenous local agglomeration forces. For simplicity I follow Krugman (1980b) and assume the local manufacturing good $Y_{r,M}$ to be a composite of differentiated products of the manufacturing firms active in region r, which compete monopolistically. Specifically, I assume that

$$Y_{r,M} = N_r^{\vartheta - \frac{1}{\rho-1}} \times \left(\int_{j=0}^{N_r} m_{j,r}^{\frac{\rho-1}{\rho}} dj\right)^{\frac{\rho}{\rho-1}}, \tag{3}$$

where N_r denotes the number of active manufacturing firms, $m_{j,r}$ is firm j's amount of manufacturing products and ρ is the elasticity of substitution across firms' outputs. Importantly, ϑ parametrizes the extent of aggregate increasing returns - in a symmetric allocation where $m_{j,r} = M/N_r$, (3) implies that $Y_{r,M} = N_r^{\vartheta} \times X$. In case $\vartheta = 0$, aggregate productivity is constant. If $\vartheta > 0$, an increase in the local manufacturing sector N_r increases aggregate productivity in the manufacturing sector. The canonical case of Krugman (1980b) corresponds to $\vartheta = \frac{1}{\rho-1}$.

Manufacturing firms in region r have access to a linear production technology

$$m_{j,r} = Z_r \times h_{j,M,r}, \tag{4}$$

where Z_r parameterizes the productivity of the manufacturing sector in region r and $h_{j,M,r}$ is amount of manufacturing labor employed by firm j. As usual, I assume that the number of firms N_r is determined by free entry and that it takes ζ units of manufacturing workers to set up a firm.

3.2 Local Labor Supply: Natives and Refugees

Each region is inhabited by a measure L_r^N of natives and a measure of L_r^R of refugees and I denote the share of refugees in the population by $\mu_r = \frac{L_r^R}{L_r^N + L_r^R}$. Refugees and natives differ in their marketable skills. In particular, as in Roy (1951), I assume that each individual i draws a skill vector $e^i = \left(e_A^i, e_M^i\right)$, where e_s^i denotes the amount of efficiency units of individual i in sector s. For concreteness I assume that for individual i of group $g = R, N$, e^i is drawn from a Frechet distribution with shape parameter θ and location $\left(Q_{A,r}^g, Q_{M,r}^g\right)$. Hence, the supply of human capital differs both by group and across space. This structure implies that the share of people of group g working in sector s is given by

$$\pi_{r,s}^g = \frac{Q_{M,r}^g w_{M,r}^\theta}{Q_{A,r}^g \left(w_{A,r}\right)^\theta + Q_{M,r}^g w_{M,r}^\theta}. \tag{5}$$

Refugees therefore have a comparative advantage in the manufacturing sector in region r if

$$\frac{\pi_{M,r}^R / \pi_{A,r}^R}{\pi_{M,r}^N / \pi_{A,r}^N} = \frac{Q_{M,r}^R / Q_{A,r}^R}{Q_{M,r}^N / Q_{A,r}^N} \equiv \chi^{Ref} > 1. \tag{6}$$

Furthermore, note that the aggregate supply of human capital to sector s is given by

$$H_r^s\left(\mu_r\right) = L_r \nu \times \left[\mu_r \left(\pi_{M,r}^N\right)^{\frac{\theta-1}{\theta}} \left(Q_{M,r}^N\right)^{\frac{1}{\theta}} + \left(1-\mu_r\right)\left(\pi_{M,r}^R\right)^{\frac{\theta-1}{\theta}} \left(Q_{M,r}^R\right)^{\frac{1}{\theta}}\right], \tag{7}$$

where ν is the usual constant from the gamma function. To the extent that refugees differ from natives in their skills, the aggregate supply of human capital will depends on the share of refugees in the population, μ_r. In particular, it can be shown (see Appendix) that $\frac{H_{M,r}}{H_{A,r}}$ is increasing in μ_r if and only if refugees have a comparative advantage in the manufacturing sector, i.e. if (6) holds true. Finally, average earnings of individuals in group g are

$$\overline{w}_r^g = \nu \times \left(Q_{A,r}^g w_{A,r}^\theta + Q_{M,r}^g w_{M,r}^\theta\right)^{1/\theta}, \tag{8}$$

Hence, average earnings are equalized across both sectors s and depend on the level skills.[4]

While the theory is agnostic about the source of skill differences between refugees and natives, I want to stress that one source could be existing frictions, which prevent refugees from entering the agricultural sector. There is a large historical literature, which stresses refugees' inability to buy agricultural land and which argues that agricultural labor markets were quite underdeveloped with the majority of agricultural workers being family members. In Section 5 below I will provide direct evidence on such frictions using microdata on occupational employment patterns for natives and refugee pre- and post-expulsion. This mechanism could easily fit into the model above. In particular, suppose (as in Hsieh et al. (2013)) that refugees receive only a fraction $1-\tau \in (0,1]$ of

[4] Intuitively, if the wage rate in in sector s increases, the selection into sector s worsens. With these functional form assumptions, these effects exactly cancel out so that the level of earning is independent of the sector where the individual decides to work in.

their wage-income in the agricultural sector. In the case, the expressions above take exactly the same form, where refugees' productivity in the agricultural sector $Q_{A,r}^{R}$ is given by

$$Q_{A,r}^{R} = \tilde{Q}_{A,r}^{R} \times (1-\tau)^{\theta}, \tag{9}$$

where $\tilde{Q}_{A,r}^{g}$ is the actual physical productivity of refugees in the agricultural sector. The structure in (9) is attractive because it implies that even if refugees and natives had the same skills, refugees had a comparative advantage in manufacturing (see (5)) and would receive lower earning (see (8)). I will show below that both of these implications are in line with the data.[5]

3.3 Equilibrium

I will analyze the data as if stemming from the equilibrium of this model. To see how the model generates a link between manufacturing labor supply, manufacturing productivity and therefore regional income per capita, first note that manufacturing firms will set a constant mark-up $\frac{\rho}{\rho-1}$ over marginal costs so that profits are given by $\pi_{j,r} = \frac{1}{\rho-1}\frac{w_{M,r}}{Z_r}m_{j,r}$. Free entry requires that $\pi_{j,r} = \zeta w_{M,r}$, which directly implies that equilibrium firm size is given by $m_{j,r} = (\rho-1) Z_r \zeta$. Hence, total labor demand by the manufacturing sector is therefore given by

$$H_r^M = \int_{j=1}^{N_r} l_{j,r} + \zeta N_r = N_r (\rho - 1)\zeta + \zeta N_r = \rho\zeta N_r, \tag{10}$$

i.e. an increase in the manufacturing workforce will increase the number of varieties in equilibrium. Total production of the regional manufacturing good is therefore given by

$$Y_{M,r} = N_{j,r}^{\vartheta - \frac{1}{\rho-1}} \times m_r N^{\frac{\rho}{\rho-1}} = N_{j,r}^{\vartheta+1}(\rho-1) Z_r \zeta = \chi \times Z_r \left(H_r^M\right)^{\vartheta+1}, \tag{11}$$

where χ is a constant.[6] Hence, the regional manufacturing sector features agglomeration in that an increase in the manufacturing workforce will *increase* manufacturing productivity. Another way to see this is from the equilibrium prices, which are given by

$$P_{A,r} = \left(\frac{w_{A,t}}{\gamma}\right)^{\gamma}\left(\frac{R_t}{1-\gamma}\right)^{\gamma} \text{ and } P_{M.r} = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}\left(\frac{1}{\rho\zeta}H_r^M\right)^{-\vartheta}. \tag{12}$$

While agricultural prices reflect wages and land rents, manufacturing prices are *decreasing* in the size of the manufacturing sector provided that the manufacturing sector. While the current structure of modeling the endogenous innovation response as "entry" is tractable and lends itself directly to the data, one could have derived similar expressions to (11) and (12) in model of directed technological change as in Acemoglu (2002). In particular, (10) captures precisely the intuition that innovation (i.e. N_r) is directed towards the more abundant factor (i.e. the supply of manufacturing human capital H_r^M).

Let us now turn to the general equilibrium of the general economy. An equilibrium allocation is defined in the usual way.

Definition 1. *An equilibrium is a set of wages* $[w_{A,r}, w_{M,r}]_r$, *land rents* $[R_r]_r$, *active manufacturing firms* $[N_r]_r$,

[5] While the expressions in (5) and (8) are consistent with an interpretation of $Q_{A,r}^{R}$ being driven by distortions (as in (9)), the relationship between income per capita and refugee flows does depend on the particular micro foundation - at least as long as the distortion τ is not modeled as a pure welfare loss but there is a "recipient" of the "rent" $(1-\tau)\, w_{A,r}$.

[6] Specifically, $\chi = (\rho-1)\zeta\left(\frac{1}{\rho\zeta}\right)^{1+\vartheta}$.

agricultural and manufacturing prices $[p_{A,r}]_r$ *and* $\left[\left[p^j_{M,r}\right]_{j=1}^{N_r}\right]_r$ *and sector labor supplies* $\left[H^N_{r,M}, H^R_{r,A}, H^R_{r,M}, H^R_{r,A}\right]_r$ *such that (i) consumers and firms behave optimally, (ii) labor markets clear in each region* r*, (iii) good markets clear and (iv) the number of firms* N_r *is consistent with free entry.*

Given the equilibrium relationship between the number of varieties N_r and manufacturing labor supply $H_{r,M}$ (see (10)) and the equilibrium prices (12), the equilibrium is fully characterized by the conditions

$$\zeta_1 \left(H^M_r(\mu_r)\right)^{1+\vartheta} = H^M_r(\mu_r)^{\sigma\vartheta} Z_r^{\sigma-1} w_{M,r}^{-\sigma} P_M^{\sigma-1} \times (1-\alpha) Y \quad (13)$$

$$\left(H^A_r(\mu_r)\right)^{\gamma} T_r^{1-\gamma} = \zeta_2 \left(w_{M,r} \left(\frac{w_{A,r}}{w_{M,r}}\right) \left(\frac{R_{A,r}}{w_{A,r}}\right)^{1-\gamma}\right)^{-\sigma} \times \alpha Y \quad (14)$$

$$\frac{w_{A,r}}{R_{A,r}} \frac{H^A_r(\mu_r)}{T_r} = \frac{\gamma}{1-\gamma}. \quad (15)$$

where $Y = \sum_r Y_r$ denotes aggregate income, P_M is the usual CES price index associated with (1) , ζ_1 and ζ_2 are constants and the human capital supply functions H^M_r and H^A_r are given in (7). It is these supply functions, which directly depend on the number of refugees μ_r. The agricultural good is taken to be the numeraire. (13) and (14) are the market clearing conditions for the regional manufacturing and agricultural good respectively. As expected, higher productivity Z_r and higher prices of competing varieties P_M increase labor demand in region r to the extent that demand is elastic (i.e. $\sigma > 1$). Similarly. higher production costs $w_{M,r}$, $w_{A,r}$ and $R_{A,r}$ reduce the demand for labor in region r. Crucially, the total human capital supply in the manufacturing sector H^M_r determines the *demand* for manufacturing products in equilibrium. To the extent that local productivity is sensitive to changes in local market size, i.e. $\vartheta > 0$, a larger manufacturing sector will increase productivity, lower prices and hence increase demand. Because H^M_r appears on both side of the market clearing condition in (13), inspection of (13) suggests that the Armington elasticity σ will be an important parameters of the analysis. This turns out to be the case - see below. Regions are linked though aggregate "world" income Y, which is determined in general equilibrium. Finally, equation (15) is the agricultural producers' optimality condition for the optimal factor mix between land and agricultural labor. These conditions fully determine the set of equilibrium prices and allocations.

In the empirical application I study the cross-sectional relationship between refugees shares (i.e. μ_r) and endogenous outcomes, i.e. sectoral employment shares and GDP per capita. Hence, consider a log-linear approximation of the equilibrium conditions above, which allows me to express the endogenous variables $\left(H^M_r, w_{M,r}, \frac{w_{A,r}}{w_{M,r}}, \frac{R_{A,r}}{w_{A,r}}\right)$ in terms of regional fundamentals, i.e. manufacturing productivity Z_r, the size of the native population L^N_r, the supply of land T_r and - most importantly - the share of refugees μ_r. To make a closer contact to the data, it is attractive to express the equilibrium directly in terms of observables. In particular, because H^M_r and N_r are proportional (see (10)) and the manufacturing employment share of natives is given by (see (5))

$$\frac{\pi^N_{M,r}}{1-\pi^N_{M,r}} = \left(\frac{w_{M,r}}{w_{A,r}}\right)^{\theta} \frac{Q^N_M}{Q^N_A},$$

i.e. monotone in the prevailing *relative* wage $\frac{w_{A,r}}{w_{M,r}}$, one can characterize the equilibrium directly in terms of $\left(N_r, w_{M,r}, \pi^N_{M,r}, \frac{R_{A,r}}{w_{A,r}}\right)$. Given t

In particular, the equilibrium level of manufacturing human capital $H_{M,r}$ and the equilibrium sectoral wages can be written as (see Section 8.13 in the Appendix)

$$
\begin{aligned}
ln\left(N_r\right) &= GE_N + (1-\gamma)\,\phi_{N,T}T_r + \phi_{N,Z}ln\left(Z_r\right) + \phi_{N,L}ln\left(L_r^N\right) + \beta_N \times \mu_r && (16)\\
ln\left(w_{M,r}\right) &= GE_{w_M} + (1-\gamma)\,\phi_{w,T}T_r + \phi_{w,Z}ln\left(Z_r\right) + \phi_{w,L}ln\left(L_r^N\right) + \beta_w \times \mu_r && (17)\\
\pi_{M,r}^N &= GE_\pi + (1-\gamma)\,\phi_{\pi,T}T_r - \phi_{\pi,Z}ln\left(Z_r\right) + \phi_{\pi,L}ln\left(L_r^N\right) + \beta_\pi \times \mu_r, && (18)
\end{aligned}
$$

where GE collects the aggregate variables Y and P_M, which are common across regions, the $\phi's$ and $\beta's$ are elasticities, which depend on the structural parameters of the model and the equilibrium allocation. In the Appendix, I derive explicit expressions for these elasticities. In particular, I show that the main coefficients of interest - the β's - are given by

$$
\begin{aligned}
\beta_N &= \frac{\phi}{(1-\overline{\mu})\,\overline{\mu}\left(1-\overline{\pi}_M^N\right)\overline{\pi}_M^N}\frac{1}{\theta}\left(\overline{\omega}_M^R\sigma + (\gamma+\sigma(1-\gamma))(\theta-1)\left(\overline{\omega}_M^R\overline{\pi}_M^N + \overline{\omega}_A^R\overline{\pi}_A^N\right)\right) && (19)\\
\beta_w &= \frac{1}{\sigma}\psi_{N,\mu} \times (\vartheta(\sigma-1)-1) && (20)\\
\beta_\pi &= \phi\frac{1}{(1-\overline{\mu})\,\overline{\mu}}\frac{\overline{\omega}_A^R}{\overline{\omega}_M^R} \times \left(\vartheta\overline{\omega}_M^R + \overline{\omega}_A^R(1-\gamma) - \frac{\overline{\omega}_A^R\left(1-\overline{\omega}_M^R\right)}{\sigma-1}\left(\chi^{Ref}-1\right)\right), && (21)
\end{aligned}
$$

where $\phi > 0$ and $\overline{\omega}_s^R$ is the share of human capital refugees account for in sector s.

The elasticities (19)-(21) reflect the direct effect of the cross-sectional variation in the share of refugees μ_r holding fundamentals T_r, Z_r and L_r^N fixed. Equations (16) - (18) are useful in both providing guidance on the expected sign of these elasticities and in concisely clarifying the identification assumptions required to test for these restrictions in the data. The two crucial implications are that the *qualitative* effect of an inflow of refugees onto equilibrium outcomes depends crucially on the interaction between the strength of agglomeration ϑ and the trade elasticity σ. First of all, note that refugee inflows cause an increase in the number of plants, i.e. $\beta_N > 0$. This is not surprising, given the positive relationship between entry ("innovation") and market size.

The effect of such refugees flows on wages in contrast is more subtle. In particular, (20) implies that

$$
\beta_w > 0 \text{ if and only if } \vartheta > \frac{1}{\sigma-1}, \tag{22}
$$

i.e. an inflow of refugees into the economy will *increase* wages in the manufacturing sector if the agglomeration elasticity is sufficiently large. Again, this condition is easier to satisfy if local firms face an elastic demand function. Intuitively, if the local manufacturing can easily expand at the expense of *other* firms, local wages are more likely to increase as a response to the shock (see also Burstein et al. (2017)). The condition in (22) reflects the relative strength of two forces. If $\vartheta = 0$, an inflow of refugees will increase manufacturing output and hence drive down on the regional wage.[7] If, in contrast, ϑ is sufficiently large, the equilibrium wage in the manufacturing sector can increase as a response to the inflow of refugees. In the canonical case of Krugman (1980b), we have $\vartheta = (\rho-1)^{-1}$, so that (22) is simply $\sigma > \rho$, i.e. if the complementarities across manufacturing producers are large ($\rho$is small) and regional varieties are closely substitutable.

Secondly, (21) shows that an inflow of refugees also has an ambiguous effect on the relative wage in the local economy and hence on the employment share of natives' in the manufacturing sector. In particular, a higher share

[7] In particular, (20) and (10) that if $\vartheta = 0$ we have

$$
\frac{dln\left(w_M\right)}{d\mu_r} = -\frac{1}{\sigma}\frac{dln\left(N\right)}{d\mu_r} = -\frac{1}{\sigma}\frac{dln\left(H_M\right)}{d\mu_r}.
$$

of refugees will crowd natives *into* the manufacturing sector if and only if

$$\underbrace{\vartheta \overline{\omega}_M^R}_{\text{Agglomeration}} + \underbrace{\overline{\omega}_A^R (1-\gamma)}_{\text{DRS}} - \frac{\overline{\omega}_A^R \left(1-\overline{\omega}_M^R\right)}{\sigma - 1} \underbrace{\left(\chi^{Ref}-1\right)}_{\text{Comp. Adv.}} > 0. \tag{23}$$

Hence, an inflow of refugees will increase wages in the manufacturing sector *relative* to the competing agricultural sector if (i) agglomeration externalities are sufficiently large ($\vartheta > 0$), (ii) the agricultural sector has decreasing returns ($\gamma < 1$) and (iii) the comparative advantage of refugees in the manufacturing industry is no too strong (χ^{Ref} small). To understand (23), suppose first technology is exogenous (i.e. $\vartheta = 0$) and that refugees and natives have the same skills $\left(\chi^{Ref} = 1\right)$. In that case, the manufacturing sector will expand, to "soak" up the additional people arriving in region r. If refugees have a comparative advantage in manufacturing $\left(\chi^{Ref} > 1\right)$ this does not necessarily happen, as refugees' labor supply is then biased towards the manufacturing sector, which reduces the relative marginal product of labor. Finally, if productivity in the manufacturing sector responds to such population flows, this tends to make the manufacturing sector more attractive for everyone - including the natives.

The important empirical implication of this effect is that inflowing refugees might *increase* the share of natives in the manufacturing sector if the condition in equation (23) is satisfied. Note that the same is true for the population of refugees, as $\pi_{M,r}^R$ is also decreasing in $\frac{w_{A,r}}{w_{M,r}}$ and hence monotone in $\pi_{M,r}^N$. Letting $\pi_{M,r} = \mu_r \pi_{M,r}^R + (1-\mu_r)\pi_{M,r}^N$ be the aggregate manufacturing employment share, one can show that

$$\frac{d\pi_{M,r}}{d\mu_r} = \underbrace{\pi_{M,r}^R - \pi_{M,r}^N}_{\text{Composition Effect}} + \underbrace{\left(\mu_r + (1-\mu_r)\frac{1}{\chi^{Ref}}\left(\frac{\pi_{M,r}^R}{\pi_{M,r}^N}\right)^2\right) \times \frac{\partial \pi_{M,r}^N}{\partial \mu_r}}_{\text{Crowding In/Out}}. \tag{24}$$

Hence, a higher share of refugees increase the manufacturing share "mechanically" if refugees have a comparative advantage in manufacturing (see (6)). Interestingly, if (23) is satisfied, this initial inflow will make the manufacturing competitive *relative* to the agricultural sector and cause a crowding-in, whereby the within-group manufacturing share also increase.

Finally, we can use these results above to link refugee flows to the variation in local income per capita. Income per capita is given by (see Section 8.15)

$$y_r = \frac{1}{\gamma}\left[(1-\mu_r)\overline{w}_r^N\left(1-(1-\gamma)\pi_{M,r}^N\right) + \mu_r \overline{w}_r^R\left(1-(1-\gamma)\pi_{M,r}^R\right)\right],$$

where $\overline{w}_r^N$ and $\overline{w}_r^R$ are the group-specific average earnings. The reason why γ features in this expression is because of the sectoral-differences in the labor share. For simplicity assume that $\gamma = 1$. One can then show that

$$\frac{\partial ln\left(y_r\right)}{\partial \mu_r} = \frac{\partial ln\left(w_{M,r}\right)}{\partial \mu_r} + \frac{\left(\overline{w}_r^R - \overline{w}_r^N\right)}{y_r} - \frac{1}{\theta}\frac{\mu_r \overline{w}_r^R \chi \frac{\pi_{M,r}^R}{\pi_{M,r}^N} + (1-\mu_r)\overline{w}_r^N}{y_r} \times \frac{\partial \pi_{M,r}^N / \pi_{M,r}^N}{\partial \mu_r}. \tag{25}$$

The first term captures the effect of refugees on the level of manufacturing wages. The second term captures the direct human capital effect. In particular, if refugees earn less than natives $\left(\overline{w}_r^R < \overline{w}_r^N\right)$, a higher share of refugees will make the region's human capital deteriorate. This has a direct negative effect on income per-capita. Finally, the last term captures the level of agricultural wages. Intuitively, for a given manufacturing wage $w_{M,r}$, a higher share of native manufacturing employment $\pi_{M,r}^N$ indicates a lower wage in the agricultural sector. This is why this last term enters with a negative sign.

The analysis above shows that *if* natives have an absolute advantage, i.e. $\overline{w}_r^R < \overline{w}_r^N$ (as is the case empirically), a higher share of refugees can only increase income per capita if technological changes desponds positively to the bigger population, i.e. $\vartheta > 0$ (and in fact sufficiently so).[8] For further reference I gather the empirical implications of the theory in the following Proposition.

Proposition 2. *Consider the model above. Suppose that refugees have a comparative advantage in manufacturing (i.e. (6)) hold true and refugee earnings are lower than native earnings, i.e. $\overline{w}^R < \overline{w}^N$ (see (8)). Then the following is true:*

1. *The manufacturing share $\pi_{M,r}$ is increasing in the share of refugees μ_r,*
2. *The number of manufacturing plants N_r is increasing in the share of refugees μ_r,*
3. *If (23) is satisfied, i.e. agglomeration forces are sufficiently strong, there is crowding-in, i.e. $\pi^g_{M,r}$ is increasing in the share of refugees μ_r,*
4. *If agglomeration forces are sufficiently strong, income per capita is increasing in the share of refugees μ_r.*

In the remainder of the paper, I will look for these implications in the data. The historical variation in the allocation of refugees μ_r is crucial to do so. The equilibrium system in (16)-(18) implies that one needs empirical variation in the cross-sectional allocation of refugees, which is orthogonal to (i) the productivity of the manufacturing sector Z_r, (ii) the aggregate supply of human capital L_r^N and (iii) the availability of other complementary production factors T_r. At the same time, I want to stress that (16)-(18) are general equilibrium relationships. This implies that the cross-sectional variation in μ_r should be large to plausibly be informative about the effects of agglomeration. The historical setting analyzed in this paper is useful on both accounts: the aggregate supply shock is large and the historical setting allows me to identify spatial variation in the share of refugees, which is arguably orthogonal to local fundamentals. This is where I turn now.

4 Empirical Analysis

4.1 Data

In this paper I use a variety of datasets. The majority of the analysis exploits spatial variation and links refugee flows in 1950 to economic outcomes at the county-level in the 50s, 60s and 70s. To perform this analysis, I constructed a panel dataset at the German county-level spanning the time-period from 1933 to 1970. The dataset was constructed by digitizing a host of historical publications. In contrast to many other countries there are, to best of my knowledge, no records of the historical micro census data with sufficient regional breath to calculate outcomes at the level of the roughly 500 Western German counties. However, the local statistical offices did publish summary statistics of the respective census at the county-level at the time. I therefore got access to the respective publications and digitized the respective data.

The basis of dataset is comprised of the population censuses for the years 1933, 1939, 1950, 1961 and 1970. For each of these years, the publications report a variety of outcomes at the county-level. Most importantly, they contain the level of population, sectoral employment shares, occupational shares, sex ratios and some other characteristics at the county-level. I then augmented this dataset with four additional pieces of information. The first concerns the regional allocation of refugees, which I digitized from a special statistical publication published in

[8] According to (23), the last term is positive if $\chi^{Ref} > 1$. One can show, however, that the whole expression in (25) is negative if $\vartheta = 0$.

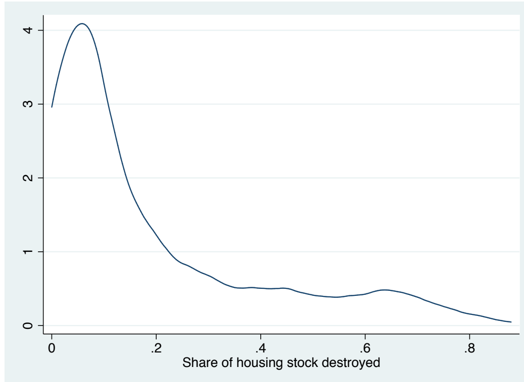
1953. Secondly, I require a measure of regional economic development. I was not able to find data on wages at the county-level for the time period before 1975. However, in the 50s, 60s and 70s, the different statistical offices from the respective German states instituted a commission to construct measures of GDP at the county-level. These are results were published and could be digitized. Third, I also digitized the county-level results for three waves of the manufacturing census in 1933, 1939 and 1956. The manufacturing census reports the number of plants by industry at the county-level. This allows me to measure entry, which - according to theory - is a measure of the endogenous technological response to changes in labor supply. Finally, I also exploit the information on war time destruction and housing supply, which I digitized from the historical housing census conducted in 1950. This census contains information on the extent of war damages for each county and detailed information on living conditions of refugees and natives. I want to stress that this data is different from the one used in Brakman et al. (2004) and Burchardi and Hassan (2013). These papers mostly focus on the extent of war-time destruction in cities. The housing census contains information on war damages for each county covering the entire landmass of Germany.

To build the final dataset it is important to realize that Germany went through numerous administrative boundary changes between 1933 and 1970. Hence, I used GIS referenced maps for the respective years to aggregate the information in a time-consistent way. I will present my empirical results at the geographical resolution that was present at the time, i.e. results for the say sectoral employment patterns in 1950 are presented in 1950 county borders, while the results for income per capita, which I measure in 1961, is reported using the borders in 1961. However, I also did the entire analysis in the borders of 1975 with almost identical results.

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

4.2 Determinants of the cross-sectional variation in refugees

To test the theoretical predictions outlined above, I require variation in the share of refugees μ_r, which is orthogonal to the other factors determining the equilibrium allocations. As stressed in theory, such confounding factors could come in two forms. First of all refugee flows could be correlated with regional fundamentals like manufacturing productivity Z_r, the existing stock of human capital L_r^N or the quantity of land T_r. Secondly, refugee flows could be selected. Recall that the structural error terms $u_{j,r}$ in the equilibrium system are functions of refugees' skills $Q_{M,r}^R$ and $Q_{A,r}^R$. Hence, if refugees are selected in that the underlying skills of refugees are correlated with share of refugees μ_r, the variation in μ_r does not only capture the labor supply effect holding human capital fixed (which is the effect stressed in the theory) but is also correlated with regional variation in the human capital of the new-comers. In this section, I will show that (1) refugee flows are correlated with spatial fundamentals, (2) how this correlation can be addressed and (3) that concerns about refugees being selected on their skills are unlikely to be important. I will also explicitly address concerns about refugees' migration decisions and show direct evidence that migratory



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

Figure 5: The Distribution of wartime destruction

responses are limited.

The spatial allocation of refugees in Western Germany in 1950 Both the historical literature and the available government documents from the Official Military Government of the US in Germany (OMGUS) clearly indicate that the spatial allocation of refugees was not random but correlated with regional fundamentals. At the same time these sources also suggest that the concern about selection is limited. In particular, the existing evidence shows very clearly that an orderly allocation of the refugees across localities in Western Germany was impossible at the time. The main reason was the administrative burden. While migration flows of this magnitude in such a sort time span would present a formidable challenge for any functioning state authority, the challenge in Germany in 1945 was enormous.

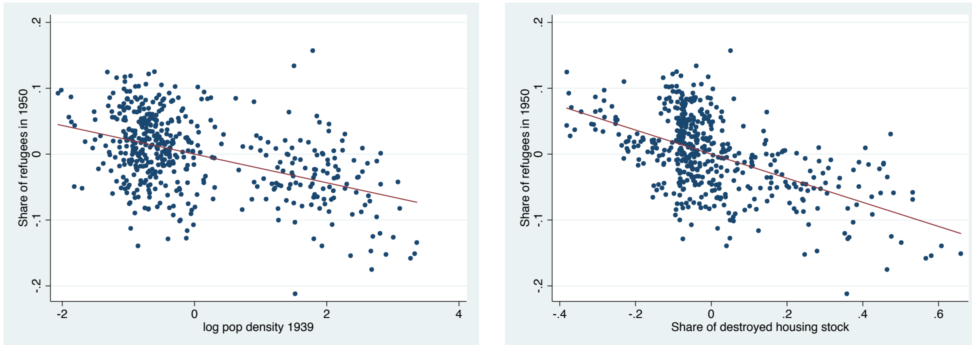
The dominant problem for respective military governments[9] were the acute shortages in housing supply. While the population in Western Germany *exceeded* its pre-war level despite the losses in the war, the housing stock was substantially lower. Data from German housing census at the county-level shows that about 23% of the aggregate housing stock was damaged. Moreover, there is considerable heterogeneity in the extent to which Western German counties were affected by the Allied bombing campaign. In Figure 5 I depict the cross-sectional distribution of war-time destruction, i.e. the share of the housing stock, which was damaged in the war. It is clearly seen that there are many counties, where more than 60% of their housing stock was damaged during the war. Moreover, the extent of war-time damages is strongly correlated with pre-war population density.[10]

Werner Nellner, one of the leading post-war economic historians, describes the situation as follows: "In the midst of the chaotic post-war circumstances arrived the refugee transports. The entirely confusing political and economic situation paired with the abruptness of this pouring-in simply did not allow a sensible distribution of the expellees into areas where they could find work. The ultimate goal was to find shelter for those displaced persons, even though in the majority of cases the situation was very primitive and many had to dwell in the tightness of refugee camps for years" (Nellner, 1959, p. 73) and some observers even concluded that most refugees were "dumped into Western Germany and settled where they could" (Petersen, 1964, p. 420).

This focus on sending the incoming refugees to places where it was easy to find shelter suggests that the

[9] Recall that until 1949 Germany did not have a federal administration, but that each occupied zone was governed by its own provisory military government.

[10] A simple bivariate regression of the share of damaged houses on the log of population density in 1939 has a R^2 of 0.45. The coefficient is 0.1 with a standard error of 0.005.



Notes: The left panel shows the relationship between the (log) population density in 1939 and the share of refugees in 1950. The right panel shows the relationship between the share of housing destroyed during the war and the share of refugees in 1950. Both plots show residual variation after taking out a set of state fixed effects.

Figure 6: Spatial determinants of the allocation of refugees: Population density and wartime destruction

new-comers settled in places with a historically low population density and in places, which saw little war-time destruction. This is exactly what we see in the data. In Figure 6, I depict the correlation of population density in 1939 (left panel) and the extent of war-time destruction (right panel) with the share of refugees (relative to the county population) in 1950. Both of these variables are strong predictors of the cross-sectional variation in refugees. In terms of the model, this implies that the share of refugees μ_r is not orthogonal but correlated with regional fundamentals.

To better understand this correlation, let us go back to theory. The underlying spatial heterogeneity is summarized by the vector $\left(Z_r, L_r^N, T_r\right)$. The patterns in Figure 6 show that the share of refugees μ_r is correlated with T_r and L_r^N. What Figure 6 does not show is that the allocation of refugees is *un*correlated with measures of manufacturing productivity Z_r once these two observables are controlled for. In particular, the model implies that the manufacturing employment share $\pi_{M,r}$ and the equilibrium number of plants N_r is positively correlated with Z_r holding L_r^N and T_r fixed. I therefore consider regressions of the form

$$\mu_r = \delta_s + \beta \times x_{cs} + \phi \times \ln \text{pop dens}_r^{39} + \varphi \times \text{war destr}_r + \eta' g_r + v_r, \tag{26}$$

where μ_r is the share of refugees in county r in state s in 1950, δ_s is a set of state fixed effects, x_{cs} are different pre-war characteristics, which are monotone in Z_r, pop dens^{39} and war destr are the population density in 1939 and the extent of wartime destruction and g_r contain two geographic controls, namely the distance to the inner german border and a dummy variable wether or not the respective county is a border region.[11] The results of (26) are contained in Table 2 below.

The first columns replicates the results depicted in Figure 6: There is a strong correlation between the share of refugees and pre-war population density and the extent of war-time destruction. The next four columns show, that conditional on these characteristics, refugee flows are uncorrelated with measures of pre-war manufacturing activity. In particular, neither the share of manufacturing employment in 1939 (column 2) or 1933 (column 3), nor the number of plants in general (column 4) or manufacturing plants in particular (column 5) in 1939 is correlated with the share of refugees. Columns 6 and 7 also show that the allocation of refugees is uncorrelated with pre-war trends in population growth and the extent of county-level urbanization in 1939, which I measure as the share of the population living in cities with less than 2000 inhabitants.[12]

[11] The latter is important to control for later changes in market access through the division of Germany in the spirit of **?**.

[12] Naturally, the latter measure is strongly correlated with population density in 1939.

	Dep. Variable: Share of refugees in 1950 (μ_r)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln pop dens. 1939	-0.010***	-0.008***	-0.011***	-0.010***	-0.011***	-0.010***	-0.006
	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Share of housing stock damaged	-0.135***	-0.137***	-0.134***	-0.134***	-0.134***	-0.136***	-0.138***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Manufacturing share in 1939		-0.026					
		(0.024)					
Manufacturing share in 1933			0.016				
			(0.021)				
ln num of plants (1939)				-0.000			
				(0.004)			
ln num of manufac. plants (1939)					0.001		
					(0.004)		
Population growth 1933-39						-0.002	
						(0.007)	
Pop share in small cities 1939							0.019
							(0.012)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	487	487	486	471	471	486	487
R^2	0.756	0.757	0.757	0.766	0.766	0.757	0.757

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. "Share of housing stock damaged" is the share of the housing stock, which was damaged during the war. "Manufacturing share in 1939" and "Manufacturing share in 1933" is the county-level manufacturing employment share in 1939 and 1933. "ln num of plants (1939)" is the log of the number of plants from the Manufacturing Census in 1939. "ln num of manufac. plants (1933)" is the log of the number of manufacturing plants from the Manufacturing Census in 1933. "Pop share in small cities 1939" is the share of the county population, which in 1939 was living in cities with less than 2000 inhabitants. "Population growth 1933-39" is the growth rate of the county population between 1933 and 1939. "State FE" indicates wether the regression controls for state fixed effects. "Distance" indicates that the regression controls for the log of the distance to the inner german border. "Border" indicates that the regression contains a set of fixed effects for wether a county is a border county.

Table 2: The Initial Allocation of Refugees Across German Counties

These results are consistent with the view that refugees were systematically allocated towards regions with high housing availability (i.e. low density and little destruction) but uncorrelated with regional manufacturing productivity Z_r and natives' human capital $Q^N_{r,s}$, holing housing supply fixed.[13] This lack of correlation was in fact not only known to the military government, but was also 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 OMGUS concludes that "both the planning and the execution of the support measures for German expellees was conducted entirely under welfare perspectives. The people in charge at the Military Government are social service officials. Similarly on the side of the German civil government, the department in charge is the social service agency. Entire communities are moved so that the population of some counties is increased by 25-30% and the agency in charge was founded to support the elderly, disabled people and the poor. Neither in Stuttgart, nor in Wiesbaden[14] we could find any information that either the Military Government or the German civil government even attempted to set up coordinated plans to receive, support or distribute the expellees. At no point, representatives of the Ministries of Labor, Industry, Nutrition or Education were consulted concerning the servicing or the distribution of expellees. The whole problem has not been handled as one of settlements of entire communities but as an emergency problem of supporting the poor." (Grosser and Schraut, 2001, p. 85). In a similar vain, a government official in Bavaria recalls that "there was no sensible planning whatsoever....The expelled peasants, worker and the industrialists, merchants, doctors, lawyers etc. had lost everything. Was there any chance that they could ever live independently again? That was their most pressing concern. The main problem for government officials and refugee commissars however was to provide provisional shelter during the winter of 1946/47" (Kornrumpf, 1979, p. 27). My first empirical strategy will therefore use the residual variation in refugee flows after controlling for pre-war trends in population density and the extent of war time destruction.

Alternatively, I will also exploit an instrumental variable strategy. The geographic patterns of the refugee allocation shown in Figure 4 suggest a strong correlation between refugee inflows and the distance to the expulsion region, which - as shown in Figure 1 - were exclusively in the East. To exploit this variation as an instrument, I calculate, for each Western German county in 1950, the distance to each county in the Eastern Territories and the Sudetenland, i.e. to all the regions, where the German population got expelled. I then take the respective county population size in 1939 to calculate a population-weighted distance for each "potential" receiving county in 1950. Formally, this measure is given by

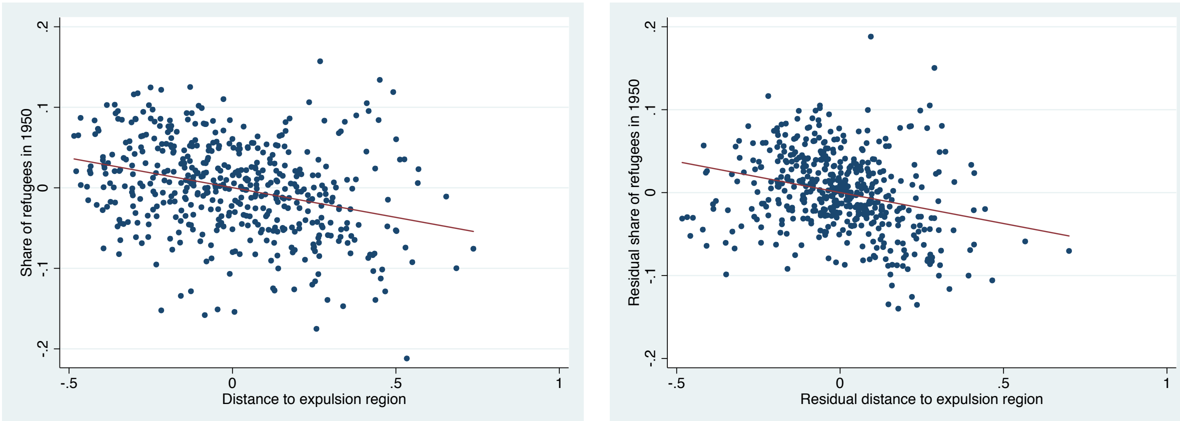
$$d_c = \sum_{r \in ER} d_{c,r} \times pop_r^{1930}, \tag{27}$$

where ER is the set of regions, where the expulsions take place, $d_{c,j}$ is the geographical distance between county j and r and pop_r^{1930} is the size of the population in 1939.

Figure 7 shows that there is a strong negative gradient between refugee flows and the pre-war distance. In the left panel I show the within-state cross-sectional relationship between refugee shares and the distance measure in (27). Hence, the negative slope does not only reflect the "East-West" comparison displayed in Figure 4, but it shows that within states, geographical distance is a key component of refugee flows. In the right panel I depict the same relationship after controlling for population density in 1939, the extent of war time destruction and the geographic controls g_r (i.e. a within-border fixed effect and the distance to the inner german border). This residual variation is very similar. My second empirical strategy will therefore use this measure as an instrument for the spatial allocation of refugees.

[13] Recall that Z_r reflects both manufacturing productivity and workers' human capital as we normalized natives' skills to unity.

[14] Stuttgart and Wiesbaden are the state capital of Baden-Wurttemberg and Hesse, both states within the US occupied zone.



Notes: The left figure shows the relationship between the (log) distance to expulsion region calculated according to (27) and the share of refugees in 1950 after taking out a set of state fixed effects. The right figure shows the relationship after controlling for the log of population density in 1939, the share of war time destruction, a set of border of fixed effects and the distance to the inner german border.

Figure 7: Refugees and the Distance to expulsion regions

Selection I will now turn to the concern of selection, i.e. the fact that the share of refugees μ_r is correlated with refugees' skills. The "lack of planning" highlighted above suggests that this form of selection was very limited. In fact, the very absence of a sensible distribution of refugees were already felt in the fall of 1947. In an official economic report by the OMGUS, it is argued that "expellees of Eastern Europe have been settled, not where their skills could be best utilized, but in accordance with the availability of food and housing to meet their needs. Thus, labor supply is often remote from the centers of labor demand ... Most of these have been placed without regard to their skills, and many where there is no demand or material for them to pursue their trades."(Office of the Military Government for Germany, 1947, p. 4-5) With a special reference to industrial workers the report notes that "the industrial population, as pointed out previously, is not at all times resident in or near the labor market. Skilled laborers from the eastern territories are living in agricultural communities far from sources of employment. And even those who have been fortunate enough to find housing and food close to industrial centers have not always had the good fortune to find a demand for their special skills" (Office of the Military Government for Germany, 1947, p. 9). This excess supply of workers with the wrong skill set induced both job degrading and unemployment in the population of refugees. Julius Isaac of the National Institute of Economic and Social Research in London worked as a consultant for the OMGUS Civil Administration Division with a special reference to the problem of assimilation of refugees. He reports that "precious skilled labor has to work in unskilled occupations because there are no opportunities to work in their jobs.... Often they are located in rural areas where it is almost impossible to find appropriate employment" (Isaac, 1948). Furthermore, "practically the only unemployment among able-bodied male workers in western Germany, other than that of the white-collar group, is among the refugees. This is because the distribution of refugees was based on housing potentialities rather than on employment possibilities and the labor market." (Barton, 1948, p, 27).

These historical assessments suggest that the variation of refugees' skills, Q^R_{rs}, was uncorrelated with existing spatial fundamentals. Using the data, I can test this assumption to some extent. While I do not have individual-level data on the allocation of refugees across space, I do observe some observable characteristics. In particular, I know the share of refugees within each county coming from the Sudetenland (versus the Eastern Territories) and the religious affiliation. In Table 3 I show that the share of refugees within these subgroups is uncorrelated with the above measures of manufacturing productivity, i.e. manufacturing employment shares (both in 1939 and 1933) and the number of industrial plants (the only exception being the one positive correlation between 1933

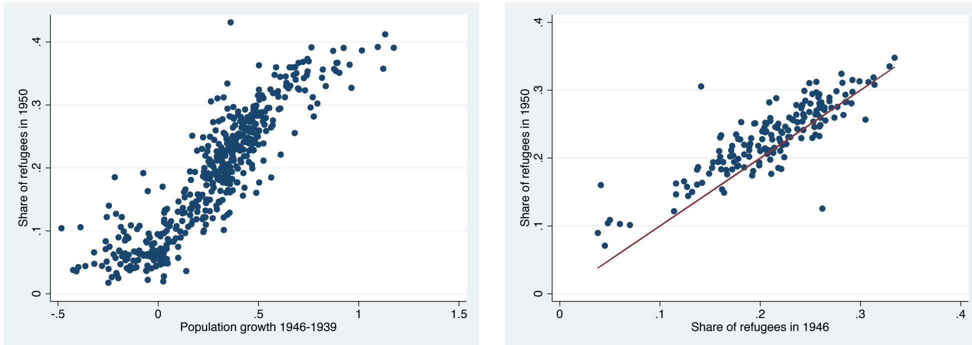
	Share of refugees from CSSR in 1950			Share of protestant refugees in 1950		
Manufacturing share in 1939	0.054			0.079		
	(0.058)			(0.054)		
Manufacturing share in 1933		0.060			0.100^{**}	
		(0.052)			(0.047)	
ln num of manufac. plants (1939)			0.003			0.009
			(0.010)			(0.008)
ln pop dens 1939	Yes	Yes	Yes	Yes	Yes	Yes
Wartime destr.	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	486	485	470	486	485	470
R^2	0.804	0.804	0.801	0.745	0.745	0.745

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 in column 1 - 3 is the share of refugees in county r, who come from the CSSR. The dependent variable in column 4 - 6 is the share of refugees in county r, who are protestant. "Manufacturing share in 1939" and "Manufacturing share in 1933" is the county-level manufacturing employment share in 1939 and 1933. "ln num of plants (1939)" is the log of the number of plants from the Manufacturing Census in 1939. "ln pop dens 1939" indicates that the regression controls for the (log of) population density in 1939. "Wartime destr." indicates that the regression controls for the share of the housing stock, which was damaged during the war. "State FE", "Distance" and "Border FE" indicate that the regression controls for state fixed effect, the log of the distance to the inner german border and for a set of fixed effects for wether a county is a border county.

Table 3: The Initial Allocation of Refugees Across German Counties: Selection

manufacturing employment and the share of protestant refugees). In particular, the first three columns are telling. In 1939 the Sudetenland was much more industrialized than the Easter Territories of the German Reich. This is for example seen in Figure 2, where I show that the population density in the Eastern Territories is quite low and the agricultural employment share very high. This is in start contrast to the Sudetenland. The fact the region of origin is uncorrelated with pre-war manufacturing productivity at the county level suggests that the extent to which refugees were allocated across space according to their pre-war occupation was small.

Migration between 1946 and 1950 Even if the initial allocation of refugees across space was indeed uncorrelated with manufacturing productivity (conditional on population density and the extent of war-time destruction), a natural concern is of course the migratory response of refugees. More specifically, to what extent does the observed allocation in 1950 already reflect refugees' endogenous location choices since 1947? It turns out that migration prior to 1950 was quite rare. First of all, it is important to note that labor mobility was severely restricted in the post-war period. In 1945, the refugee committee of the Occupying Forces decided to deploy armed forces at the state boundaries to prevent internal migration (Fluechtlingsausschuss des Laenderrats (1945)) and William H. Draper, Director of the Economic Division of the OMGUS, notes that "Germany has been virtually cut into four Zones of Occupation - with the Zone borders not merely military lines, but almost air-tight economic boundaries" (Office of the Military Government for Germany, 1945, p. 10). Additionally, the incentives to migrate were also arguably low. Due to above mentioned housing shortage, the majority of expellees either lived in refugee camps or with the native population who were forced by the Military Governments to accept expelled families in their homes. This political support however was only provided in the locations refugees were initially assigned to. Similarly, there



Notes: The left panel shows the relationship between the population growth rate from 1939-1946 and the share of refugees in 1950. The right panel shows the relationship for the share of refugees in 1946 and 1950 for the state of Bavaria.

Figure 8: Spatial Mobility between 1946 and 1950

were restriction to receive food stamps without being officially registered (Grosser and Schraut, 2001, p. 83).[15]

In Figure 8 I provide evidence for this absence of a large-scale migratory response. While the census in 1950 is - to the best of my knowledge - the only dataset with comprehensive data on the allocation of refugees across all counties in Western Germany, the census in 1946 has information on the population at the county level. The inflow of refugees in the years until 1947 was arguably the major source of variation in population growth between 1939 and 1946.[16] If migration between 1947 and 1950 would be a major concern, the allocation of refugees in 1950 should not be strongly correlated with the rate of population growth prior to 1946. As seen in the right panel of Figure 8 however, the opposite the case - counties with high population growth from 1939 to 1946 still have a much higher share of refugees in 1950. For one state, Bavaria, I actually did find information on the allocation of refugees in 1946. Hence, for this county, I can directly look at the correlation between the allocation of refugees in 1946 and 1950. As seen in the left panel of Figure 8. these track each other very closely. Again, this is consistent with a view that migration prior to 1950 did not play a very important role.

This absence of spatial mobility is in fact often alluded to in the contemporary sources. For example the economic reports of the OMGUS themselves argue that high levels of unemployment are accompanied by labor shortage because "the mobility of labor is limited. Hence there is little possibility of an early change in the distribution of labor. For example. 46% of the job openings in Bavaria in March 1947 were in the major cities of Munich, Nuremberg and Augsburg, while the majority of immigrant labor resided in rural districts. In consequence, the economic absorption of immigrants is greatly hampered" (Office of the Military Government for Germany, 1947, p. 10). Despite the large heterogeneity in the economic burden the refugees imposed on individual states, there were hardly any organized transfers. The main reason for this was the reluctance of states to accept refugees and it was only after the foundation of the Federal Republic of Germany in 1949, that there was enough political pressure for a centralized solution for the redistribution to be found. Before 1950 however, "all proposals for specific reallocation rules were rejected. The main argument was that the available data about the housing situation was not comparable across states. Hence, it was demanded that a general housing census was needed, which was only conducted in 1950 in conjunction with the general census. ... Measures concerning a population redistribution within the area of Germany had only a minor influence until the end of 1950" (Nellner, 1959, p. 39, 43).[17] Gerhard Reichling,

[15] Additionally the Allied governments were very reluctant to allow for free migration as they were afraid that expellees would form non-integrated sub populations along former village or city lines. To achieve this, various organizations trying to track down friends and family members were forbidden until the early 1950s (Nellner, 1959, p. 75).

[16] In fact, the rate of population growth between 1939 and 46 is strongly correlated with the distance to the expulsion regions in Eastern and Central Europe.

[17] It is hard to tell if this was a real problem or if those states that were about the receive expellees from Bavaria, Schleswig-Holstein

Employment share of refugees relative to natives in ...			
Sectors		Occupations	
Agiculture	8.1	Self-employed	5.4
Manufacturing	14.8	Family worker	1.9
Services	12.1	Employees	13.5
		Unskilled workers	22.3

Note: This table reports the share of refugees in total employment for the respective sectoral or occupational group, i.e. $L_s^R / \left(L_s^R + L_s^N\right)$.
Source: Federal Statistical Office (1953, p. 47-49)

Table 4: The Importance of Refugees by Sectors and Occupations

a German historian who was in charge of the main statistical analysis about the integration of the refugees, also concludes that "there is no aspect where the Federal Republic of Germany shows a similar degree of heterogeneity as in the absorption and distribution of expellees. The occupying forces decided about the contingents for their occupation zones without any German involvement....The fact that there is no central German state authority during the time of the expulsion and the missing solidarity of the states with respect to the expellees prevents an equitable and adequate distribution of burdens caused by the influx of expellees for years" (Reichling, 1958, p. 17).

4.3 Economic Outcomes: Refugees, Industrialization and Economic Development

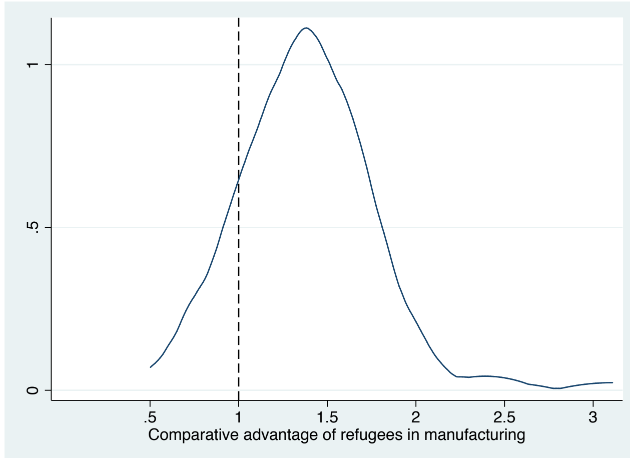
I will now turn to the study of the outcomes of this historical experiment. I start by providing direct evidence to the premise of Proposition 2, i.e. that (i) refugees indeed have a comparative advantage in manufacturing but (ii) had a lower level of skills in that they had lower earnings. I will then use the cross-sectional variation across counties to establish a positive relationship between the allocation of refugees and local GDP per capita. Together with the fact the refugees earn less in the cross-section, this hard to rationalize without the existence of some agglomeration forces. Finally, I exploit the structure of the theory to provide additional evidence in favor of agglomeration forces and the underlying mechanism. I first show that refugees indeed triggered entry of manufacturing plants. I then study the relationship between refugees and manufacturing employment, both in the short to medium run, i.e. at the 3 year horizon, and in the long-run, i.e. at the 10 year horizon, and show that refugee inflows caused a substantial increase in the manufacturing sector. Finally, I directly focus on the implications on crowding-in. I find a positive relationship between refugee inflows and within-group manufacturing employment shares, which is again consistent with the presence of agglomeration externalities.

4.3.1 Skill Endowments: Absolute and Comparative Advantage between Refugees and Natives

The majority of my analysis relies on the variation in refugees across localities to test the predictions in Proposition 2. This proposition, however, relies on (i) refugees having a comparative advantage in the manufacturing industry and (ii) refugees having lower earning than natives. In this section I provide evidence that this is indeed the case.

Consider first the case of comparative advantage. In Table 4 I report the share of refugees in different sectoral and occupational groups. Refugees are most common to work in the manufacturing sector. Moreover, they are most common in the occupational class of unskilled workers. Conversely, there is a suspicious absence of refugees in the agricultural sector. Most importantly, there are few self-employed refugees or working family-members. In Section 5 below I will present additional micro-data to argue the employment patters in Table 4 are likely to reflect the inability of refugees to acquire agricultural land. However, for the main analysis I do not have to take a stand

and Lower Saxony used this argument as an excuse to postpone large scale population transfers.



Notes:: This figure plots the distribution of $\lambda_r \equiv \frac{\pi^R_{r,M}/\pi^R_{r,A}}{\pi^N_{r,M}/\pi^N_{r,A}}$ for all counties in Bavaria in 1950. Through the lens of the model, $\lambda_r = \frac{Q^R_{M,r}/Q^R_{A,r}}{Q^N_{M,r}/Q^N_{A,r}}$ (see (6)). Hence, comparative advantage is equalized if $\lambda_r = 1$.

Figure 9: Refugees' Comparative Advantage in Manufacturing

wether this comparative advantage literally reflects refugees' superior skills in industrial occupations or wether they entered manufacturing jobs because of potential entry barriers in the agricultural sector.

While Table 4 already shows that refugees were more likely to work in the manufacturing sector, it actually *understates* the extent to which this is the case. The reason is that refugees mostly lived and worked in rural areas, where manufacturing employment was less important (as seen in Table 2). Though the lens of the model, we can identify the extent of comparative advantage directly from the observed group-specific sectoral employment shares *within* localities $\pi^g_{r,s}$. In particular, refugees have a comparative advantage in manufacturing in region r if $\lambda_r \equiv \frac{\pi^R_{r,M}/\left(1-\pi^R_{r,M}\right)}{\pi^N_{r,M}/\left(1-\pi^N_{r,M}\right)} > 1$ (see 6). While the majority of census files do not contain sectoral employment patters at the county level separately for refugees and natives, the state Bavaria published a special employment report in 1950, where such data is included. Hence, for this state we can directly calculate λ_r for each county. In Figure I plot the cross-sectional distribution of λ_r. Because λ_r exceeds unity for the majority of regions, refugees have indeed a substantial comparative advantage in the manufacturing sector.

While these patterns are informative about relative skills, it is also important to know wether refugees had superior human capital than the native population. If that was the case, the relationship between regional GDP per capita and the number of refugees would reflect both potential agglomeration economies and a positive inflow in human capital (per capita). In Table 5 I show that this is not the case. In particular, I use the EVS microdata to run regressions of the form

$$\overline{w}_i = \beta \times Refugee_i + \alpha' x_i + \delta_s + \delta_{Ind} + \delta_{City} + \delta_{Occ} + \nu_i,$$

where $\overline{w}_i$ denotes earning of individual i, $Refugee$ indicates the refugee status, x_i controls for demographic characteristics and $\delta_s, \delta_{Ind}, \delta_{City}$ and δ_{Occ} are state, industry, city and occupation fixed effects. The results in Table 5 show that if anything refugees had lower level of marketable human capital than their native counterparts. On average, refugee earnings were about 10% lower in 1962. Industrial and occupational sorting patterns account for some of these lower wages. However, even within industry-location-occupation pairs, refugees earned about 3.5% less on average. As for the patterns of comparative advantage, these results could be driven by distortions, whereby

	Dep. Variable: log earnings ($\ln(\bar{w}_i)$)			
Refugee	-0.098***	-0.100***	-0.071***	-0.036***
	(0.010)	(0.008)	(0.007)	(0.007)
Demographics	No	Yes	Yes	Yes
State FE	No	Yes	Yes	Yes
City size FE	No	Yes	Yes	Yes
City structure FE	No	Yes	Yes	Yes
Industry FE	No	No	Yes	Yes
Occupation FE	No	No	No	Yes
Observations	32584	32573	32573	32573
R^2	0.003	0.323	0.401	0.488

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The dep. variable is the log of annual earnings. "Refugee" is an indicator for wether the individual is a refugee. "Demographics" control for sex and for a quadratic polynomial in age. "State FE" indicates that the regression control for a set of state fixed effects. "City size FE" and "City structure FE" indicate wether the regression controls for a set of five city size fixed effects and five city structure ("Urban center", "Urban fringe", "industrial area", "rural community", "mixed zone") fixed effects. "Industry FE" control for a set of 11 industry fixed effects. "Occupation FE" control for a set of 10 occupation fixed effects. The data stems from 1962 cross-section of the EVS (Einkommens- und Verbrauchsstichprobe).

Table 5: Refugees vs Natives: Earnings in 1962

refugees faced barriers to enter the agricultural sector.[18]

4.3.2 Refugees, Agglomeration and Local Economic Development

The most important implication of the theory is that in the presence of agglomeration forces an increase in the labor force can increase GDP per capita even though the marginal worker might be less skilled than the average worker. To study this question, I now relate the inflow of refugees to local productivity. To measure productivity, I exploit measures of GDP at the county level in 1961. My measure of productivity is total GDP relative to the economically active population. This population measure corrects the local population by measures of commuting. In the context of this analysis, this is correction is important as commuting flows at the county level are large. In Section 8.9 in the Appendix I describe the construction of the data in more detail. In what follows I will for simplicity refer to this measure as "GDP per capita". My main empirical specification takes the form

$$\ln y_r^{1961} = \delta_s + \beta \times \ln \mu_r^{1950} + \zeta \times Z_r + \phi \times \ln \text{pop dens}_r^{39} + \varphi \times \text{war destr}_r + \eta' g_r + \upsilon_r, \quad (28)$$

where y_{cs}^{1961} denotes GDP per capita in region r in 1961, μ_r^{1950} denotes the share of refugees in 1950, pop dens^{39} and war destr are the population density in 1939 and the extent of wartime destruction, g_r contains the geographic controls and Z_r are different measure of pre-war manufacturing productivity. As a baseline specification, I will estimate β in (28) using OLS. In terms of the theory, this assumes that Z_r, the pre-war population density and the extent of war destruction control for the spatial variation in fundamentals $\left(T_r, Z_r, L_r^N\right)$. As an alternative I will also estimate β using the expulsion distance as an instrument. The results are contained in Table 6.

In the first column I report the simple cross-sectional relationship between GDP per capita in 1961, the pre-war measures of spatial fundamentals and the share of wartime destruction. As expected, regions with a higher

[18] While the theory implies that such distortions would lower average earning (see (8) and (9)), the particular functional forms used in the theory predict that such distortions would reduce earning equally in all sectors. This is of course a very special property of the Frechet distribution.

	Dep Variable: log GDP per capita ($\ln y_r$)							
	OLS						IV	
ln refugee share 1950		-0.062**	0.119***	0.110***	0.124***	0.111***	0.093*	
		(0.028)	(0.027)	(0.027)	(0.026)	(0.026)	(0.049)	
Population growth 1939-46								0.375**
								(0.163)
ln pop dens. 1939	0.054***		0.057***	0.003	0.297***	-0.006	0.002	0.018
	(0.010)		(0.010)	(0.016)	(0.086)	(0.016)	(0.016)	(0.018)
Share of housing stock damaged	0.009		0.153**	0.178***	0.195***	0.147**	0.158*	0.260**
	(0.057)		(0.071)	(0.067)	(0.064)	(0.064)	(0.083)	(0.112)
Manufacturing share in 1939	0.730***		0.750***	0.605***	0.428	0.519***	0.596***	0.597***
	(0.084)		(0.084)	(0.172)	(0.701)	(0.168)	(0.174)	(0.171)
Manufacturing share in 1933				0.027	0.961	0.004	0.034	-0.029
				(0.192)	(0.663)	(0.184)	(0.194)	(0.200)
Agricultural share in 1933				-0.043	0.078	-0.056	-0.043	-0.069
				(0.083)	(0.111)	(0.082)	(0.083)	(0.087)
Pop share in small cities 1939				-0.230***	-0.191***	-0.281***	-0.231***	-0.217***
				(0.049)	(0.050)	(0.053)	(0.050)	(0.050)
ln num of manufac. plants (1939)					0.107***	0.089***		
					(0.020)	(0.013)		
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Higher order terms	No	No	No	No	Yes	No	No	No
City FE	No	No	No	No	Yes	No	No	No
Population weights	No	No	No	No	No	Yes	No	No
Observations	486	486	486	485	469	469	485	482
R^2	0.481	0.225	0.500	0.528	0.598	0.577	0.527	0.515

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The regression is at the county level. The dependent variable is the log of GDP per capita in 1961. "ln refugee share 1950" is the log of the share of refugees in 1950. "Population growth 1939-46" is the growth rate of the county population between 1939 and 1946. "ln pop dens 1939" is the log of the population density in 1939. "Share of housing stock damaged" is the share of the housing stock, which was damaged during the war. "Manufacturing share in 1939" and "Manufacturing share in 1933" is the county-level manufacturing employment share in 1939 and 1933. "Agricultural share in 1933" is the county-level agricultural employment share in 1933. "Pop share in small cities 1939" is the share of the county population, which in 1939 was living in cities with less than 2000 inhabitants. "ln num of plants (1939)" is the log of the number of plants from the Manufacturing Census in 1939. "State FE", "Distance" and "Border FE" indicates wether the regression controls for state fixed effects, the log of the distance to the inner german border and fixed effect for wether a county is a border county. "Higher order terms" indicates wether the regression contains a full set of quadratic terms and interactions of the pre-war sectoral employment shares. "City FE" and "Population weights" indicate wether the regression contains a set of city fixed effects and wether each observation is weighed with the log of the population in 1933. In the IV specifications (columns 7 and 8) I instrument the share of refugees (column 7) and the rate of population growth (column 8) with the distance to the expulsion regions (see (27)).

Table 6: Refugees and Economic Development: 1961

population density and a larger manufacturing sector in 1939 are richer 20 years later. In terms of the theory, I interpret these correlations to reflect the persistence in regional human capital $\left(L_r^N\right)$ and manufacturing productivity (Z_r). Note that there is no relationship between war time destruction and income per capita. In column two I report the *unconditional* relationship between refugee-flows and economic development. This relationship is strongly negative because - as shown above - refugees were not assigned randomly but allocated to rural and hence poorer areas. This is seen column three. Once the observable pre-war heterogeneity and the extent of war-time destruction is controlled for, the effect of refugees is positive. Note also that the inclusion of the refugee share changes the relationship between GDP per capita and population density and wartime-destruction. In particular, places with more war-time destruction are richer in 1961, presumably because they were already more developed during the war and hence a target for the Allied bombing campaign.

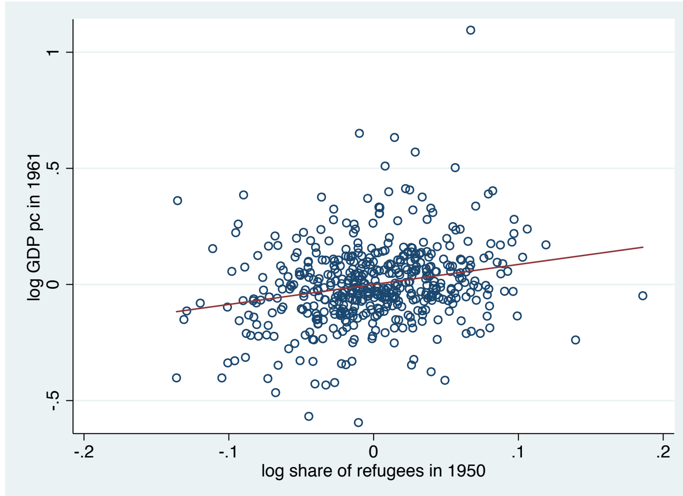
Columns 4, 5 and 6 show that this effect of refugees on local economic development is robust to controlling for richer controls in the pre-war sectoral structure and alternative measures of urbanization (column 4), for city fixed effects and higher-order terms in the pre-war control variables (column 5) and for weighting regions by their population size in 1939. Doing so leaves both the coefficient and the standard error essentially unchanged. Finally, columns 8 and 9 contain two different IV specifications. In column 8, I instrument the refugee share by the pre-war population-weighted distance to the expulsion regions (see (27)). The IV estimate is very close to the OLS even though the the standard error doubles. In the last column I consider a complementary strategy. Instead of using the share of refugees μ_r as the main independent variable, I use the population growth rate between 1939 and 1946 and instrument it with the same distance-based instrument. As shown in Figure 8 above, the allocation of refugees in 1950 is the main predictor of population growth in 1946. To the extent that the distance based instrument captures the part of the variation in population growth, which is due to the inflow of refugees, this is another way to relate differences in income per capita to the importance of refugees in the population. It also has the benefit that it is measured in October 1946, i.e. closer to the date of the actual expulsions. Table 6 shows that this measure is also significantly correlated with manufacturing employment at the county-level. Finally, in Figure 10 I show the relationship between refugees and GDP per capita estimated in column 4 graphically. The figure shows a robust positive relationship, which is not driven by particular outliers.

The economic magnitudes implied by the estimates in Table 6 are economically meaningful. Changing the share of refugees from 15% to 25% corresponds to a change of 0.5 log points.[19] Given the elasticity of 0.1, this corresponds to a change in income per capita by 5%. This estimate is also consistent with the elasticity between GDP per capita and population growth reported in the last column. Suppose that population growth between 1939 and 1946 was only driven by the inflow of refugees. It is then easy to verify that $\frac{L_r^{46}}{L_r^{39}} = \frac{1}{1-\mu_r}$. For the above example of a change of μ_r from 15% to 25%, the estimated elasticity of 0.375 also generates differences in GDP per capita of 5%. The fact that these different strategies generate estimates, which are quantitatively identical is reassuring in that the results in Table 6 reflect the direct effect of refugee inflows holding fundamentals fixed. Together with the fact that refugees earn *less* than refugees, Table 6 is difficult to rationalize in a model where local technology does not respond positively to the size of the local population.

4.3.3 The Mechanism: Refugees and Local Manufacturing

The theory makes clear predictions on the economic mechanism, which underlies Table 6. With agglomeration forces strong enough to result in an increase local income per capita, the theory implies that the inflow of refugees should *also* be positively correlated with manufacturing employment and the entry of manufacturing plants. In this section I will test both of these predictions.

[19] The empirical standard deviation of $ln\left(\mu_r\right)$ is 0.66.



Notes: The figure shows the relationship between the (log of) share of refugees in 1950 and (the log of) GDP per capita in 1961 estimated from specification 4 in Table 6, i.e. after controlling for (i) log population density in 1939, (ii) the extent or wartime destruction, (iii) the manufacturing share in 1939 and 1933, (iv) the agricultural share in 1933), (v) the extent of county-level pre-war urbanization, (vi) state fixed effects, (vii) the distance to the inner-german border and (viii) a dummy variable for wether or not the county is a border county.

Figure 10: Refugees and Economic Development in 1961

Refugees and Manufacturing Employment Shares in 1950 Consider first the effect of refugee inflows and manufacturing employment in 1950, i.e. roughly 3 years after the initial shock. I consider the same empirical strategy as for the case of GDP per capita. More specifically, my main specification is

$$\pi_{M,r}^{1950} = \delta_s + \beta \times \mu_r^{1950} + \zeta \times Z_r^{pw} + \phi \times \ln \text{ pop dens}_r^{39} + \varphi \times \text{war destr}_r + \eta' g_r + v_r, \tag{29}$$

where $\pi_{M,r}^{1950}$ is the regional manufacturing share in 1950 and the remaining variables are defined as above. In particular, Z_r^{pw} again contains various measures of pre-war manufacturing productivity in region r and I will estimate β both by OLS and by the IV strategy used above.

The coefficient of interest β captures both the direct composition effect, whereby a higher share of refugees, which predominantly work in the manufacturing industry will increase the manufacturing shares, and the endogenous effect of crowding-in (see (24)). Because the condition on the strength of agglomeration ϑ for there to be crowding-in is strictly weaker than for manufacturing wages and income per capita to increase (see (22) and (23)), the theory implies that there is crowding-in and that β exceeds the pure composition effect.

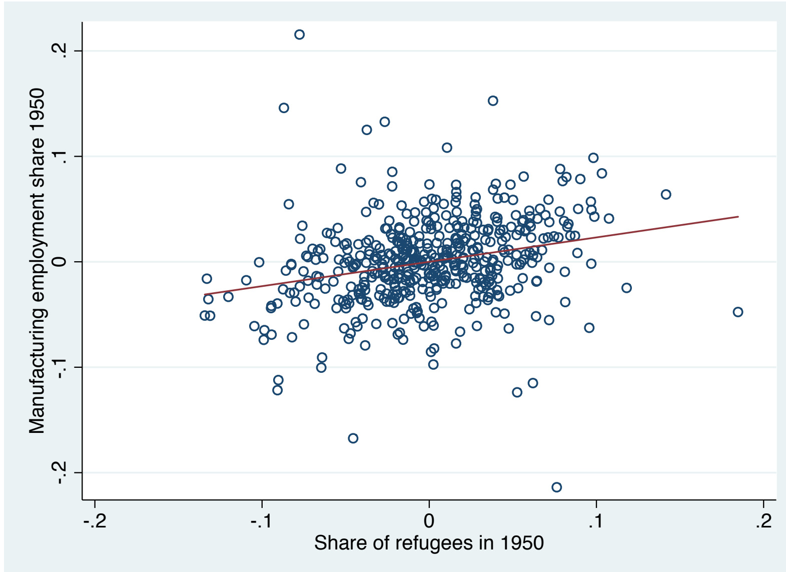
The results of estimating (29) are contained in Table 7. In column 1 I first show that - as already suggested by Table 2 above - there is no relationship between the share of refugees in 1950 μ_{cs} and manufacturing share in 1939 conditional on the controls. Columns 2 runs the exact same specification 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 manufacturing employment by 2.2 percentage points. If this effect was *only* due to the composition effect, it would need to be that $\pi_{M,r}^{N} - \pi_{M,r}^{R} = 0.22$, i.e. the manufacturing share among refugees would need to be 22 percentage points larger than for natives. While - as shown above - refugees indeed have a comparative advantage in manufacturing, the empirically observed variation in manufacturing shares are much smaller.

Columns 3 shows that this effect is unchanged, when I control for pre-war manufacturing productivity as proxied for by the manufacturing employment share in both 1933 and 1939. Note that the standard error declines substantially as manufacturing employment shares are persistent at the regional level. Note also that the R^2 of the

	Manufacturing employment share in						Employment share in	
	in 1939	in 1950					Agricult.	blue coll. occ.
	OLS	OLS			IV		OLS	
Share of refugees in 1950	-0.060	0.218**	0.231***	0.161***	0.275***		-0.169*	0.179***
	(0.086)	(0.085)	(0.049)	(0.048)	(0.078)		(0.101)	(0.040)
Population growth 1939-46						0.144***		
						(0.040)		
ln pop dens. 1950				0.072***				
				(0.016)				
ln pop dens. 1939	0.031***	0.030***	-0.002	-0.070***	-0.002	0.005	-0.042***	-0.005
	(0.007)	(0.007)	(0.004)	(0.016)	(0.004)	(0.004)	(0.008)	(0.003)
Share of housing stock damaged	-0.057*	-0.034	0.026*	0.047***	0.031*	0.074***	0.013	-0.020
	(0.032)	(0.031)	(0.015)	(0.016)	(0.018)	(0.028)	(0.027)	(0.015)
Manufacturing share in 1939			0.644***	0.640***	0.648***	0.631***	-0.325***	0.004
			(0.082)	(0.079)	(0.083)	(0.077)	(0.080)	(0.034)
Manufacturing share in 1933			0.246***	0.236***	0.241***	0.248***	-0.138**	0.022
			(0.068)	(0.064)	(0.067)	(0.062)	(0.068)	(0.023)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urbanization in 1939	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Manufacturing share 1950	No	No	No	No	No	No	No	Yes
Male share in 1950	No	No	No	No	No	No	No	Yes
Observations	487	487	486	486	485	482	486	413
R^2	0.527	0.538	0.900	0.907	0.899	0.899	0.877	0.909

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The regression is at the county level. The dependent variable is the manufacturing employment share in 1939 (column 1), in 1950 (columns 2 - 6), the agricultural employment share in 1950 (column 7) and the employment share in blue-collar occupations in 1950 (column 8). “Share of refugees in 1950” is the the share of refugees in 1950. “Population growth 1939-46” is the growth rate of the county population between 1939 and 1946. “ln pop dens 1939” and “ln pop dens 1950” is the log of the population density in 1939 and 1950. “Share of housing stock damaged” is the share of the housing stock, which was damaged during the war. “Manufacturing share in 1939” and “Manufacturing share in 1933” is the county-level manufacturing employment share in 1939 and 1933. “State FE”, “Distance” and “Border FE” indicates wether the regression controls for state fixed effects, the log of the distance to the inner german border and fixed effect for wether a county is a border county. “Urbanization in 1939” controls for the share of the county population, which in 1939 was living in cities with less than 2000 inhabitants. “Manufacturing share in 1950” and “Male share in 1950” controls for the 1950 manufacturing share and the share of males in the population. In the IV specifications (columns 5 and 6) I instrument the share of refugees (column 5) and the rate of population growth (column 6) with the distance to the expulsion regions (see (27)).

Table 7: Refugees and Manufacturing Employment in 1950



Notes: The figure shows the relationship between the share of refugees in 1950 and the share of workers working in the manufacturing industry implied by column 3 in Table 7, that is conditional on (i) log population density in 1939, (ii) the extent or wartime destruction, as measured by the share of the housing stock destroyed, (iii) the manufacturing share in 1939 and 1933, (iv) state fixed effects, (v) the extent of county-level urbanization, as measured by the share of the population living in cities with less than 2000 inhabitants, (vi) the distance to the inner-german border and (vii) a dummy variable for wether or not the county is a border county.

Figure 11: Refugees and Manufacturing Employment in 1950

regression increased to 0.9. Column 5 asks to what extent the relationship between refugees and manufacturing employment are purely a function of population density. Conditional on population density in 1939, high refugee-share regions are obviously larger in 1950. Column 4 shows that the size of the manufacturing sector is positively correlated with population density in 1950 but that refugees are "special" in that a higher share of refugees increases manufacturing employment holding the size of the population fixed. Note that this is exactly what is predicted by the theory. With decreasing returns in agriculture, population growth per se will increase manufacturing employment.[20] However, to the extent that refugees have a comparative advantage in manufacturing, variation in μ_r will increase the relative supply of manufacturing human capital and hence manufacturing employment (see (7)). Columns 5 and 6 contain the two IV estimates. As was the case for income per capita, the IV point estimate is very similar to the OLS but has a larger standard error. If I again use the instrumented rate of population growth from 1939 or 1946, I also find a significant positive relationship with the size of the manufacturing sector. Again, this coefficient is also *quantitatively* consistent with the estimates in columns 2 to 5. If refugees were the main determinant of the cross-sectional variation in population growth, the estimate of 0.144 implies that an increase in the share of refugees by 10 percentage points increase the manufacturing share by 2.2%.[21] Finally, Figure 11 again displays the empirical relationship estimated in column 3. As was the case for GDP per capita, the empirical relationship is robustly positive and does not seem to driven by particular outliers.

In the last two columns of Table 7 I consider regression where I use the agricultural employment share (column 7) and the share of blue collar workers (column 8) as dependent variables. These show that the increase in manufacturing employment stems to a large extent from a decline in agricultural employment and that refugee-rich counties also have a sizable larger share of blue-collar workers. Note that this is not simply an artifact of these

[20] Formally, the coefficient $\phi_{u,L}$ in equation (18) is negative as long as $\gamma < 1$. Hence, a higher number of people will increase the *relative* wage in manufacturing, which will increase the manufacturing share.

[21] If population growth was mainly determined through refugees, then $L_r^{46}/L_r^{39} = (1-\mu_r)^{-1}$. An increase in μ_r from 0.15 to 0.25 implies that L_r^{46}/L_r^{39} changes by 0.16. Multiplying this with 0.144 yields 0.022.

counties having larger manufacturing shares as column 8 explicitly controls for the overall share of manufacturing workers. Both these results are consistent with the results on sectoral allocation patterns of refugees reported in Table 4 and analysis of the micro data contained in Section 5 below.

Refugees and Plant Entry The key theoretical mechanism for the results in Tables 6 and 7 is an endogenous technological response of productivity and refugee inflow. In the theory laid out above, this response is captured by the interaction of plants entry and agglomeration externalities. While I want to stress that there is a variety of mechanisms by which local productivity can respond to the size of the local labor force, the focus on plant entry has the benefit that it is directly observable in the manufacturing census. Coincidentally, this mechanism also appears explicitly in the historical sources. In 1949, M. Bold, the Deputy Director of the US Military Government in Bavaria for example notes that "since refugees and bombed-out Bavarians now living in rural areas cannot move nearer to industrial jobs, such jobs must go to them. In fact many world famous industries wanting to reestablish in Bavaria have already sought locations in non-industrial areas near idle workers".

From the digitized historical manufacturing census files for 1933, 1939 and 1956, I can measure the number of manufacturing plants at time t, N_r^t. This allows me to relate the growth rate of manufacturing plants $g_{N,r} = ln\left(N_r^{1956}/N_r^{1939}\right)$ to the inflows of refugees according to

$$g_{N,r} = \delta_s + \beta \times \mu_r^{1950} + \zeta \times Z_r^{pw} + \phi \times \ln \text{pop dens}_r^{39} + \varphi \times \text{war destr}_r + \eta' g_r + v_r, \tag{30}$$

I will refer to this growth rate as the extent of entry.

Table 8 shows that there is a strong positive relationship, whereby refugee inflows trigger plant entry at the local level. Columns 1 and 2 show that an increase in the refugee share by 10 percentage points increases the number of manufacturing plants by 10% and that this correlation does not depend on whether one controls the pre-war manufacturing share. Columns 3 and 4 allow the level of manufacturing plants in 1939 to affect subsequent entry. Again, this does not change the estimates at all. Finally, columns 5 to 7 consider three IV approaches. Columns 5 and 6 use - as above - the instrumented refugee share and extent of regional population growth. Both show a significant relationship between entry and refugee flows, even though the coefficient in column 5 is quite a bit larger. In the last column, I focus directly on the number of refugees holding the size of the population in 1939 fixed. The estimated elasticity implies that a 10% increase in the number of refugees is associated with a 2.3% increase in the number of manufacturing plants.

The Long Run: Refugees and the Size of the Manufacturing Sector in 1961

In Table 7 I focused on the size of the manufacturing sector in 1950. Income per capita, however, is measured in 1961. The theory is a static model, which implies a relationship between the contemporaneous manufacturing share and the income per capita. In this section I therefore show that the spatial variation in the late 1940s predicts the size of the manufacturing sector in 1961. To do so, I follow the exact same approach as above, i.e. I consider regressions of the form of (29), except that I will use the the manufacturing share in 1961, $\pi_{M,r}^{1961}$ as the dependent variable. For brevity I only report the results with a full set of pre-war controls.

In the first two columns of Table 9 I show that the refugee share in 1950 is significantly correlated with the manufacturing share in 1961, i.e. about 10-15 years after the initial settlement. Comparing the results in Table 9 with the ones from 1950 in Table 7 it is seen that the coefficients are roughly of the same magnitude. In columns 3 and 4 I conduct the same exercise except that I use the share of refugees in 1961 as the dependent variable. Of course, during the 1950s mobility in Germany is in principle unrestricted. Hence, if refugees are more likely

	Growth of manufacturing plants 1939-56		log of manufacturing plants 1956				
	OLS		OLS		IV		
Share of refugees in 1950	1.033**	1.056**	1.050**	1.056**	1.939***		
	(0.459)	(0.440)	(0.436)	(0.428)	(0.602)		
Population growth 1939-46						0.843***	
						(0.302)	
ln refugees 1950							0.233***
							(0.077)
ln num of manufac. plants (1939)			0.953***	0.826***	0.880***	0.922***	0.782***
			(0.040)	(0.083)	(0.064)	(0.067)	(0.091)
ln num of manufac. plants (1933)				0.136*	0.070	0.054	0.037
				(0.082)	(0.050)	(0.053)	(0.047)
ln pop dens. 1939	-0.028	-0.096***	-0.091**	-0.087**	-0.066**	-0.038	-0.057*
	(0.037)	(0.035)	(0.036)	(0.034)	(0.032)	(0.037)	(0.033)
Share of housing stock damaged	-0.270	-0.142	-0.135	-0.116	-0.095	0.110	-0.082
	(0.189)	(0.180)	(0.181)	(0.178)	(0.162)	(0.223)	(0.185)
ln pop 1939							-0.065
							(0.097)
Manufacturing share in 1939		1.528***	1.560***	1.422***	1.481***	1.295***	1.398***
		(0.418)	(0.419)	(0.404)	(0.402)	(0.404)	(0.406)
Manufacturing share in 1933		0.312	0.326	0.463	0.334	0.417	0.425
		(0.371)	(0.375)	(0.377)	(0.367)	(0.374)	(0.374)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Urbanization	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	471	470	470	467	466	465	466
R^2	0.319	0.401	0.789	0.789	0.796	0.792	0.801

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 growth of the number of manufacturing plants between 1939 and 1956, $\ln\left(N_r^{59}/N_r^{39}\right)$, in column 1 and 2 and the log of manufacturing plants in 1956, $\ln\left(N_r^{56}\right)$, in columns 3 - 7. "Share of refugees in 1950" is the the share of refugees in 1950. "Population growth 1939-46" is the growth rate of the county population between 1939 and 1946. "ln refugees 1950" is the log of the number of refugees in 1950. "ln num of manufac. plants (1939)" is the log of the number of manufacturing in 1939 and for 1933 analogously. "ln pop dens 1939" and "ln pop 1939" is the log of the population (density) in 1939. "Share of housing stock damaged" is the share of the housing stock, which was damaged during the war. "Manufacturing share in 1939" and "Manufacturing share in 1933" is the county-level manufacturing employment share in 1939 and 1933. "State FE", "Distance" and "Border FE" indicates wether the regression controls for state fixed effects, the log of the distance to the inner german border and fixed effect for wether a county is a border county. "Urbanization in 1939" controls for the share of the county population, which in 1939 was living in cities with less than 2000 inhabitants. In the IV specifications (columns 5-7) I instrument the share of refugees (column 5), the rate of population growth (column 6) and the (log of the) number of refugees (column 7) with the distance to the expulsion regions (see (27)).

Table 8: Refugees and the Entry Manufacturing Plants

to move than the rest of the population and regions, which experience growth in their manufacturing sector, to be more attractive places to live, we would expect the coefficient on the 1961 refugee share to be upward biased. This is exactly what I find in data. Column 3 reports the OLS estimate. This estimate is about 50% higher than the OLS estimate using the 1950 refugee share reported in column 1. Of course, we can still exploit the distance to the pre-war expulsion regions to instrument for the 1961 refugee share. Despite the fact that refugees had 15 years to relocate, there is still a sizable first stage relationship, which is highly significant.[22] If we only use the variation in the allocation of refugees in 1961, which is predicted by the instrument, the coefficient is almost numerically identical to the IV estimate using the instrumented refugee share in 1950. This suggests that refugees' migration decisions between 1950 and 1961 are correlated with changes in the size of the manufacturing sector. This spatial co-movement between refugees and manufacturing employment is also seen in column 5, where I show that *changes* in the share of refugees between 1950 and 1961 are positively correlated with changes in the manufacturing employment share. I want to stress that this coefficient does *not* have a causal interpretation even if the initial allocation of refugees was as good as random conditional on the set of pre-war characteristics.[23]

5 Sources of Comparative Advantage: Micro evidence on Skills and Frictions

The above analysis showed a positive relationship between refugee-inflows, manufacturing employment and income per capita at the regional level. This analysis in particular suggests that incoming refugees increased the relative supply of human capital in the manufacturing sector, induced manufacturing firms to enter and thereby spurred regional manufacturing productivity through local gains from variety. In this last section of the paper I will provide some more micro-evidence on *why* refugees were "biased" towards the manufacturing sector, that is why refugees were much more likely to work in manufacturing as seen in Table 4 and Figure 9. In particular, I will present additional micro-evidence, which suggests that frictions for refugees to enter the agricultural sector were important. In the notation of theory: an important source of refugees' comparative advantage in manufacturing might have been a distorted agricultural labor market. While such distortions were obviously disadvantageous for refugees (holding prices fixed), they might have been beneficial for the regional economy as a whole if the endogenous productivity response in the manufacturing sector was important.

To provide some suggestive evidence for the importance of frictions, I will exploit the information from the 1971 census, which contains micro-data on long-run employment histories. In particular, that census asked every respondent in 1971 where he/she lived in 1939 and in which occupation/sector cell he/she worked in 1939, 1950, 1960 and 1971. By analyzing this time-series of retrospective questions, I can therefore measure the life-cycle of employment patterns for both refugees and natives. Importantly, the data spans the time of the expulsion in the mid 1940s. Hence, I can exactly see how refugees' employment patterns change *relative* to natives between 1939 and 1950.

A first look at this data is contained in Table 10. The first two columns contain sectoral and occupational employment shares for natives and to-be refugees in 1939, i.e. *prior* to the expulsion. Consistent with the higher agricultural employment shares in the Eastern Territories (see Figure 2), individuals living in Western Germany in 1971 but having lived in the expulsion regions in 1939 were more likely to work in agriculture and less likely to work

[22] The coefficient on the distance to the expulsion regions in the first stage equation is given by -0.054 with a standard error of 0.016. For comparison, the first stage relationship with the share of refugees in 1950 has a coefficient of -0.083 with almost the same standard error.

[23] In future work I plan use the initial variation in the spatial allocation of refugees in quantitative model of spatial mobility to study the co-movement between the industrial structure and labor mobility.

	Dep. Variable: Manufacturing share in 1961 $\pi_{r,M}^{1961}$				
	OLS	IV	OLS	IV	OLS
Share of refugees in 1950	0.204***	0.263***			
	(0.059)	(0.094)			
Share of refugees in 1961			0.330***	0.263***	
			(0.086)	(0.084)	
Change in refugee share 1950-1961					0.259**
					(0.102)
Manufacturing share in 1950					0.928***
					(0.051)
ln pop dens. 1939	-0.018***	-0.018***	-0.020***	-0.020***	-0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Share of housing stock damaged	0.003	0.011	0.007	0.001	-0.030**
	(0.019)	(0.023)	(0.020)	(0.020)	(0.013)
Manufacturing share in 1939	0.751***	0.756***	0.726***	0.728***	0.094**
	(0.094)	(0.095)	(0.088)	(0.088)	(0.046)
Manufacturing share in 1933	0.184**	0.179**	0.175**	0.180**	-0.045
	(0.074)	(0.075)	(0.070)	(0.071)	(0.030)
ln num of manufac. plants (1939)	-0.005	-0.005	-0.002	-0.002	0.012**
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)
ln num of manufac. plants (1933)	0.006	0.006	0.004	0.004	-0.007
	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)
State FE	Yes	Yes	Yes	Yes	Yes
Distance	Yes	Yes	Yes	Yes	Yes
Border FE	Yes	Yes	Yes	Yes	Yes
Urbanization	Yes	Yes	Yes	Yes	Yes
Observations	468	468	463	463	463
R^2	0.818	0.818	0.835	0.834	0.921

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 1961. "Share of refugees in 1950 (1961)" is the the share of refugees in 1950 (1961). "Change in refugee share 1950-1961" is the change in the refugee share between 1950 and 1961, i.e. $\mu_r^{1961} - \mu_r^{1950}$. "Manufacturing share in 1950" is the manufacturing employment share in 1950. "ln pop dens 1939" is the log of the population density in 1939. "Share of housing stock damaged" is the share of the housing stock, which was damaged during the war. "Manufacturing share in 1939" and "Manufacturing share in 1933" is the county-level manufacturing employment share in 1939 and 1933. "ln num of manufac. plants (1939)" is the log of the number of manufacturing in 1939 and for 1933 analogously. "State FE", "Distance" and "Border FE" indicates wether the regression controls for state fixed effects, the log of the distance to the inner german border and fixed effect for wether a county is a border county. "Urbanization in 1939" controls for the share of the county population, which in 1939 was living in cities with less than 2000 inhabitants. In the IV specifications (columns 2 and 4) I instrument the share of refugees in 1950 (column 2) and the share of refugees in 1961 (column 4) with the distance to the expulsion regions (see (27)).

Table 9: Refugees and Manufacturing Employment in 1961

	Pre expulsion		Post expulsion					
	1939		1950		1960		1971	
	Natives	Refugees	Nat.	Ref.	Nat.	Ref.	Nat.	Ref.
Sectoral composition of employment								
Agriculture	18.7	**26**	17.4	**12.6**	12.8	**4.9**	10	2.5
Manufacturing	43.6	**35.9**	46.5	**51.9**	47.8	**56.6**	45.3	52
Services	24.7	23.2	23.9	21	25.2	22	26	24.7
Public Sector	12.9	14.9	12.1	14.5	14.1	16.4	18.7	20.8
Occupational composition of employment								
Self-employed (Agricult.)	17.1	**18.8**	16.6	**3.6**	13.6	**2.8**	11.3	2.6
Skilled Employee	5	5.4	5.4	5.3	7	7	9.6	9.6
Unskilled Employee	12.1	12.3	10.5	10.3	11.8	11.8	12.4	13
Skilled Worker	2.2	1.9	2.4	2	1.9	1.9	2.3	2.5
Unskilled Worker	30.3	**29.7**	27.9	**46.3**	26.1	**37.5**	23.3	31.6

Note: This table reports sectoral and occupational employment shares for the 1939, 1950, 1960 and 1971 by refugee status. The data stems from the MZU 71 (Mikrozensus Zusatzerhebung 1971), i.e. the data for the years prior to 1971 is based on individuals' responses in 1971.

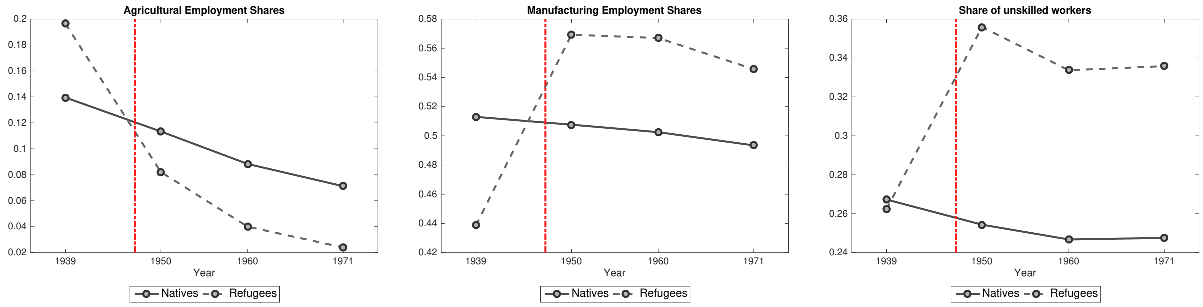
Table 10: Sectoral and Occupational Mobility from 1939 to 1971

in manufacturing. In terms of their occupational standing, they were about as likely as their native peers to be self-employed in agriculture and there is no difference in the likelihood to work in an unskilled occupation. The next two columns show the same data in 1950, i.e. immediately *after* the expulsion. While the employment patterns for native almost the same as in 1939, they are vastly different for refugees in Western Germany. In particular, their employment share in agriculture declines by more than 50%. At the same time, manufacturing employment among refugees increases dramatically, exceeds 50% and is now higher than for natives. The occupational data in the lower panel has additional information on these reallocation patterns: the decline in agricultural employment is essentially accounted for by a decline in self-employed farmers, i.e. famers who lost their land when being expelled. After the expulsion, these individuals take unskilled jobs, which are mostly in the manufacturing sector. The remaining columns in Table 10 show that these reallocation patterns in 1950 are by no means transitory. In contrast, they persist all throughout the 1950s and 1960s.

While Table 10 is indicative of the expulsion being a large disruption on individual employment histories, note that the composition of individuals changes as, for example, only few individuals who worked in 1939 are still in the labor force in 1971. To see more directly that Table 10 is in fact representative of a typical life-cycle, consider Figure 12, where I depict the same information for the cohort of workers born between 1915 and 1919. Hence, this cohort is 20-25 years old in 1939 and in their late twenties or early thirties at the time of the expulsion around 1946. In 1971, this cohort is 50-55 years old, i.e. still in the labor force. The three panels in the Figure show the agricultural employment share (left panel), the manufacturing employment share (middle panel) and the cohort's share of unskilled workers (right panel). The red vertical line indicates the time of the expulsion.

Following this cohort of individuals shows a very similar expulsion experience. Among refugees, 20% use to work in the agricultural sector in 1939. After the expulsion and their resettlement to Western Germany, only 8% still did so. In contrast, the share of manufacturing employment within the same cohort of individual increases from 44% to almost 60%. Moreover, the majority of the increase is accounted for by an increase in the employment of unskilled workers. Again, this is very different for the cohort of natives, where the time of the expulsion is hardly noticeable.

The patterns in Table 10 and (especially) Figure 12 are suggestive of distortions in the agricultural sector. First of all, refugees were *more* likely to work in agriculture in 1939. Secondly, as seen above, refugees were much *more* likely to actually live in rural communities. Hence, refugees left the agricultural sector *despite* living and working



Notes:The figure shows the agricultural employment share (left panel), the manufacturing share (middle panel) and the share of unskilled workers (right panel) for the cohort of workers born between 1915 and 1919 by refugee status. The experience of natives (refugees) is depicted in solid (dashed) lines. The expulsion, taking place in 1947, is drawn as the red, vertical line. The data stems from the MZU 71 (Mikrozensus Zusatzerhebung 1971), i.e. the data for the years prior to 1971 is based on individuals' responses in 1971.

Figure 12: The Life-Cycle of the 1915-1919 Cohort

in regions with a comparative advantage in the production of agricultural goods. While it is of course possible for agricultural skills to be highly specialized and for example to be soil-specific, it seems plausible that many refugees simply had difficulties to find agricultural work despite their agricultural human capital.

In fact, the historical literature is in wide agreement on why the assimilation of refugees in the agricultural sector would be difficult. In 1950 Germany, the majority of agricultural employment was still very much concentrated in small, family-run farms. According to the agricultural census, which reports the average farm size for each county in Germany, the average farm size is on the order of magnitude of 10-15 hectares.[24][25] Hence, the demand for outside agricultural workers was quite limited or - in terms of the theory - the returns to scale in the agricultural sector were low. Even the Military Government of the US point out in 1947 that of the immigrants "well over half a million, were farmers. But agricultural acreage [...] cannot be expanded significantly. Within the US Zone the possibility of increasing settlement by changing the size and structure of farms is very small".

To see that more directly, consider Table 11, where I show the occupational employment distribution for native and refugee workers for both the agricultural and the manufacturing sector. Consider first the agricultural sector. The first column shows that 75% of all native workers in agriculture are either self-employed or family members. This relative absence of hired hands is of course the consequence of most farms being small. Now consider the case of refugees. Not only are very few refugees employed in agriculture to begin with, but *conditional* on actually working in the agricultural sector, almost all of them are in fact hired workers. The reason is obviously that few refugees were able to acquire land. Western Germany (in contrast to the Soviet Occupied Zone) did not have a land reform, where refugee were compensated for their land losses during the expulsion. Additionally, refugees naturally had very limited means to acquire land - both because they did not have any assets and because the supply of land for sale prior to the currency reform in 1949 was very limited. The combination of these features might have made the manufacturing sector a very natural sector to seek employment in.

6 Conclusion

In this paper I used a particular historical setting to empirically estimate wether local technology responds to changes in local factor supplies. I focused on the expulsion of the ethnic German population from their domiciles in Central and Eastern Europe and their subsequent resettlement in Western Germany. In the three years after the

[24] In Section 7.1 in the Appendix I provide more analysis of the micro-data. See in particular Figure 14.

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

Occupation	Agriculture		Manufacturing	
	Natives	Refugees	Natives	Refugees
Self-employed	24.4	2.5	18.2	9.7
Family employment	49.9	5.1	5.6	1.5
Employee	0.6	1.3	11.1	8.9
Workers	25.1	90.9	65.2	79.9

Note: This table reports sectoral and occupational employment shares for the 1939, 1950, 1960 and 1971 by refugee status. The data stems from the MZU 71 (Mikrozensus Zusatzerhebung 1971), i.e. the data for the years prior to 1971 is based on individuals' responses in 1971.

Table 11: Occupational Distribution within Sectors

Second World War, roughly 8m people arrived in Western Germany. At the time, this amounted to an increase in population by about 20%. Furthermore, counties in Western Germany differed substantially in the extent to which they participated in the incoming refugee flows. Using both the policies of the US and UK Military Government in Post-War Germany and the pre-war geographic distance from counties in Western Germany to the expulsion regions to isolate the exogenous component in refugee flows, I study the effect of labor supply shocks on long-run economic development across 500 counties in Western Germany.

I find a positive relationship between refugee inflows and local productivity and between refugee inflows and the size of the manufacturing sectors. I also present direct evidence on the theoretical mechanism generating such agglomeration effects, namely a positive relationship between refugee inflows and the entry of new manufacturing plants. Together with the fact that refugees' earning were lower then the ones of natives, these cross-county results are consistent with models featuring endogenous technological change but hard to reconcile with a neoclassical model, where productivity is exogenous.

References

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

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

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

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

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "Immigration and the Rise of American Ingenuity," 2017. Working Paper.

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

Barton, Betty, *The problem of 12 million German refugees in today's Germany*, Philadelphia: American Friends Service Commitee, 1948.

Besser, Christoph, "Die deutschen Vertriebenen in der Bundesrepublik Deutschland von 1945 bis 1969," Historical Statistic of Germany.

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

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

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

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

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

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

de Zayas, Alfred-Maurice, *The German Expellees: Victims in War and Peace*, Palgrave Macmillan, 1993.

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

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

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

Fajgelbaum, Pablo and Stephen J Redding, "External integration, structural transformation and economic development: Evidence from argentina 1870-1914," 2014. Working Paper.

Fuchs-Schündeln, Nicola and Matthias Schündeln, "Precautionary savings and Self-Selection: Evidence from the german reunification "Experiment"," *The Quarterly Journal of Economics*, 2005, *120* (3), 1085–1120.

_ **and Tarek Hassan**, "Natural Experiments in Macroeconomics," in "The Handbook of Macroeconomics," Elsevier, 2015.

fuer Datenverarbeitung und Statistik Nordrhein-Westfalen, Landesamt, "Bruttoinlandsproduct und Bruttowertschoepfung der kreisfreien Staedte und Landkreise in der Bundesrepublik Deutschland," Technical Report Heft 8, Volkswirtschafticher Gesamtrechnung der Laender 1978.

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.

Hanlon, W Walker, "Necessity is the mother of invention: Input supplies and Directed Technical Change," *Econometrica*, 2015, *83* (1), 67–100.

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

Hsieh, Chang-Tai, Eric Hurst, Peter Klenow, and Charles Jones, "The Allocation of Talent and U.S. Economic Growth," 2013.

Isaac, Julius, *German Refugees in the U.S. Zone 1948/49* 1948.

Kornrumpf, Martin, *In Bayern angekommen. Die Eingliederung der Vertriebenen: Namen, Daten, Fakten*, Muenchen: Olzog, 1979.

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

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

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

Nagy, Dávid Krisztián, "City location and economic development," 2016. Working Paper.

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

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

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

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

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

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

Petersen, William, "Review," *American Sociological Review*, 1964, *29* (3), 419–421.

Redding, Stephen and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," 2016. Working Paper.

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.

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

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

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

Roy, Andrew D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3*, 135–146.

Statistisches Bundesamt, *Statistisches Taschenbuch über die Heimatvertriebenen in der Bundesrepublik Deutschland und in West-Berlin*, Wiesbaden: Statistisches Bundesamt, 1953.

_ , “Die Vertriebenen und Fluechtlinge in der Bundesrepublik Deutschland in den Jahren 1946 bis 1953,” in “Statistik der Bundesrepublik Deutschland,” Vol. 114 1955.

_ , “Die Vertriebenen und Fluechtlinge in der Bundesreublik Deutschland 1946-1953,” in “Statistik der Bundesrepublik Deutschland,” Vol. 114 Kohlhammer Stuttgart 1955.

Steinberg, Heinz Günter, *Die Bevölkerungsentwicklung in Deutschland im Zweiten Weltkrieg*, Bonn: Kulturstiftung der deutschen Vertriebenen, 1991.

7 Appendix

7.1 Additional Empirical Results

Eastern Territories of the German Reich		East and Southeast Europe	
Silesia	4.6	Danzig	0.4
Brandenburg	0.6	Baltic States	0.3
Pomerania	1.9	Poland	1
East Prussia	2.5	Czechoslovakia	3.5
		Hungary	0.6
		Yugoslavia	0.7
		Romania	0.8

Notes: The table shows the ethnic German population in different regions in East and Central Europe in 1939. Source: Federal Statistical Office (1953)

Table 12: The German Population in Central and Eastern Europe in 1939

7.2 Characterization of the equilibrium

In this section we characterize the equilibrium of the model laid out in section 3.

The manufacturing sector Consider first the manufacturing sector in region r. Let $w_{M,r}$ be the manufacturing wage in region r. Because firms set a constant markup $\frac{\rho}{\rho-1}$ over their marginal costs, (4) implies that the profits of firm j in region r are given by$\pi_{j,r} = p_{j,r}m_{j,r} - \frac{w_{M,r}}{Z_r}m_{j,r} = \frac{1}{\rho-1}\frac{w_{M,r}}{Z_r}m_{j,r}$. Free entry requires that $\pi_{j,r} = \zeta w_{M,r}$, so that

$$m_{j,r} = m_r = (\rho - 1) Z_r \zeta. \tag{31}$$

Hence, equilibrium employment of firm j in region r is given by $l_{j,r} = (\rho - 1)\zeta$. Total labor demand by the manufacturing sector is therefore given by

$$H_r^M = \int_{j=1}^{N_j} l_{j,r} + \zeta N_j = \rho\zeta N_j, \tag{32}$$

the number of varieties in equilibrium is proportional to the number of workers in the manufacturing sector. Furthermore, (3) and (31) imply that

$$Y_{M,r} = N_{j,r}^{\vartheta - \frac{1}{\rho-1}} \times m_r N^{\frac{\rho}{\rho-1}} = N_{j,r}^{\vartheta+1}(\rho - 1) Z_r\zeta = \bar{\zeta} Z_r \left(H_r^M\right)^{\vartheta+1}, \tag{33}$$

where

$$\bar{\zeta} = \frac{(\rho - 1)\zeta}{(\zeta\rho)^{\vartheta+1}}. \tag{34}$$

Hence, the degree of increasing returns is determined by ϑ.

We can also calculate the corresponding price index. By symmetry, a consumer spends a fraction $\frac{X}{N_j}$ on each variety if aggregate spending is X. Given the equilibrium price $p_j = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}$, the consumer buys $m_j = \frac{X}{N_r}\frac{1}{p_j} = \frac{X}{N_r}\frac{\rho-1}{\rho}\frac{Z_r}{w_{M,r}}$. The total manufacturing service flow is hence given by

$$Y = N_r^{\vartheta+1} m_j = N_r^{\vartheta+1}\frac{X}{N_j}\frac{\rho - 1}{\rho}\frac{Z_r}{w_{M,r}} = N_r^{\vartheta} X \frac{\rho - 1}{\rho}\frac{Z_r}{w_{M,r}}.$$

Hence, the implied price index is given by

$$P_{M,r} \equiv \frac{X}{Y} = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}N_r^{-\vartheta} = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}\left(\frac{1}{\rho\zeta}H_r^M\right)^{-\vartheta} = \frac{1}{\bar{\zeta}}\frac{w_{M,r}}{Z_r}\left(H_r^M\right)^{-\vartheta}. \tag{35}$$

The agricultural sector Given the production function in (2), the price of the agricultural variety in region r is given by

$$p_{A,r} = \left(\frac{w_{A,r}}{\gamma}\right)^{\gamma}\left(\frac{R_r}{1-\gamma}\right)^{1-\gamma}, \tag{36}$$

where $w_{A,r}$ and R_r is the agricultural wage and the land rent in region r. Moreover, profit maximization of agricultural producers implies that labor demand is given by.

$$H_{A,r} = \frac{R_r}{w_{A,r}}\frac{\gamma}{1-\gamma}T_r. \tag{37}$$

Total agricultural production can therefore be written as

$$Y_{A,r} = H_{A,r}^{\gamma}T_r^{1-\gamma} = \left(\frac{R_r}{w_{A,r}}\frac{\gamma}{1-\gamma}\right)^{\gamma}\times T_r.$$

Labor supply Aggregate labor supply in region r stems from individuals' sectoral choice problem. Consider a refugee i in region r. Given wages $w_{A,r}$ and $w_{M,r}$, this refugee decides to work in the manufacturing sector as long as

$$w_{M,r}e_M^i \geq w_{A,r}(1-\tau)e_A^i.$$

The share of refugees in region r working in the manufacturing sector is therefore given by

$$\pi_{M,r}^R = \int_{e_M}\left[\int_{e_A\leq\frac{w_{M,r}}{w_{A,r}(1-\tau)}e_M} dF(e_A)\right]dF(e_M).$$

As e_A and e_M are independently Frechet distributed, i.e.

$$F(e_s) = exp\left(-Q_s e_s^{-\theta}\right) \text{ and } f(e_s) = Q_s\theta e_s^{-\theta-1}exp\left(-Q_s e_s^{-\theta}\right)$$

we get that

$$\begin{aligned}
\pi_{M,r}^R &= \int_{e_M} P\left(e_A \leq \frac{w_{M,r}}{w_{A,r}(1-\tau)}e_M\right)dF(e_M)\\
&= \int_{e_M} exp\left(-Q_{A,r}^R\left(\frac{w_{M,r}}{w_{A,r}(1-\tau)}e_M\right)^{-\theta}\right)Q_{M,r}^R\theta e_M^{-\theta-1}exp\left(-Q_{M,r}^R e_M^{-\theta}\right)de_M\\
&= Q_{M,r}^R\int_{e_M} exp\left(-\left[\frac{Q_{A,r}^R\left(w_{A,r}(1-\tau)\right)^{\theta}+Q_{M,r}^R w_{M,r}^{\theta}}{w_{M,r}^{\theta}}\right]e_M^{-\theta}\right)\theta e_M^{-\theta-1}de_M.
\end{aligned}$$

Defining

$$\delta \equiv \frac{Q_{A,r}^R\left(w_{A,r}(1-\tau)\right)^{\theta}+Q_{M,r}^R w_{M,r}^{\theta}}{w_{M,r}^{\theta}} \tag{38}$$

we get that

$$\begin{aligned}
\pi^R_{M,r} &= \frac{Q^R_{M,r}}{\delta}\int_{e_M} exp\left(-\delta e_M^{-\theta}\right)\delta\theta e_M^{-\theta-1}de_M = \frac{Q^R_{M,r}}{\delta} \\
&= \frac{Q^R_{M,r}w^\theta_{M,r}}{Q^R_{A,r}\left(w_{A,r}\right)^\theta + Q^R_{M,r}w^\theta_{M,r}}. \qquad (39)
\end{aligned}$$

Similarly, we get that

$$\pi^N_{M,r} = \frac{Q^N_{M,r}w^\theta_{M,r}}{Q^N_{A,r}\left(w_{A,r}\right)^\theta + Q^N_{M,r}w^\theta_{M,r}}. \qquad (40)$$

Now, let us solve for the total amount of efficiency units provided. Consider the refugees in region r. The distribution of e^i_M conditional on choosing the manufacturing sector is

$$\begin{aligned}
H(m) &= P\left(e_M \leq m | \text{working in manufacturing}\right) \\
&= \frac{P\left(e_M \leq m \text{ and } e_A \leq \frac{w_{M,r}}{w_{A,r}(1-\tau)}e_M\right)}{P\left(e_A \leq \frac{w_{M,r}}{w_{A,r}(1-\tau)}e_M\right)} \\
&= \frac{1}{\pi^R_{M,r}}\int_{e_M=0}^m exp\left(-Q^R_{A,r}\left(\frac{w_{M,r}}{w_{A,r}\left(1-\tau\right)}e_M\right)^{-\theta}\right)Q^R_{M,r}\theta e_M^{-\theta-1}exp\left(-Q^R_{M,r}e_M^{-\theta}\right)de_M \\
&= \frac{1}{\pi^R_{M,r}}\int_{e_M=0}^m exp\left(-\delta e_M^{-\theta}\right)Q^R_{M,r}\theta e_M^{-\theta-1}de_M \\
&= \frac{1}{\pi^R_{M,r}}\frac{Q^R_{M,r}}{\delta}\int_{e_M=0}^m exp\left(-\delta e_M^{-\theta}\right)\delta\theta e_M^{-\theta-1}de_M \\
&= exp\left(-\delta m^{-\theta}\right).
\end{aligned}$$

Hence, the total labor supply in the manufacturing sector is[26]

$$H^R_{M,r} = L_{R,r}\times\pi^R_{M,r}\times\delta^{1/\theta}\times\Gamma\left(1-\frac{1}{\theta}\right),$$

where $\Gamma(.)$ is the gamma function. Using that $\delta = \frac{Q^R_{M,r}}{\pi^R_{M,r}}$ (see (38) and (39)) we get that

$$H^R_{M,r} = L_{R,r}\times\left(\pi^R_{M,r}\right)^{\frac{\theta-1}{\theta}}\times\left(Q^R_{M,r}\right)^{1/\theta}\times\nu,$$

where $\nu \equiv \Gamma\left(1-\frac{1}{\theta}\right)$.

Hence, the total supply of efficiency units in the manufacturing sector is given by

$$\begin{aligned}
H_{M,r} &= H^N_{M,r} + H^R_{M,r} \\
&= L^N_r\times\left(\pi^N_{M,r}\right)^{\frac{\theta-1}{\theta}}\times\nu\times\left(Q^N_{M,r}\right)^{\frac{1}{\theta}} + L^R_r\times\left(\pi^R_{M,r}\right)^{\frac{\theta-1}{\theta}}\times\nu\times\left(Q^R_{M,r}\right)^{\frac{1}{\theta}} \\
&= L^N_r\nu\left(\pi^N_{M,r}\right)^{\frac{\theta-1}{\theta}}\left(Q^N_{M,r}\right)^{\frac{1}{\theta}}\left\{1+\frac{\mu_r}{1-\mu_r}\times\left(\frac{\pi^R_{M,r}}{\pi^N_{M,r}}\right)^{\frac{\theta-1}{\theta}}\times\left(\frac{Q^R_{M,r}}{Q^N_{M,r}}\right)^{\frac{1}{\theta}}\right\}. \qquad (41)
\end{aligned}$$

[26] Recall that of $F(x) = e^{-\left(\frac{x}{s}\right)^{-\alpha}} = e^{-s^\alpha(x)^{-\alpha}}$ we have that $E[x] = s\times\Gamma\left(1-\frac{1}{\alpha}\right)$.

Similarly, agricultural labor supply is given

$$H_{A,r} = L_r^N \nu \left(\pi_{A,r}^N\right)^{\frac{\theta-1}{\theta}} \left(Q_{A,r}^N\right)^{\frac{1}{\theta}} \left\{ 1 + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\pi_{A,r}^R}{\pi_{A,r}^N}\right)^{\frac{\theta-1}{\theta}} \times \left(\frac{Q_{A,r}^R}{Q_{A,r}^N}\right)^{\frac{1}{\theta}} \right\}. \tag{42}$$

Goods demand Given consumers' preferences in (1), the goods market clearing conditions are

$$Y_{M,r} = \overline{\zeta} Z_r \left(H_r^M\right)^{\vartheta+1} = \frac{1}{p_{M,r}} \left(\frac{p_{M,r}}{P_M}\right)^{1-\sigma} (1-\alpha) \times Y \tag{43}$$

$$Y_{A,r} = H_{A,r}^{\gamma} T_r^{1-\gamma} = \frac{1}{p_{A,r}} \left(\frac{p_{A,r}}{P_A}\right)^{1-\sigma} \alpha \times Y \tag{44}$$

where Y denotes aggregate income, $p_{A,r}$ and $p_{M,r}$ are the regional prices given in (35) and (36) and P_M and P_A are the usual CES price indices

$$P_s = \left(\sum_{r=1}^{R} p_{s,r}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{45}$$

(43) and (44) simply stem from the fact that consumers spend a fraction α $(1-\alpha)$ on manufacturing (agricultural) products and the usual demand relationship stemming from the CES structure of the Armington aggregators. Finally, aggregate income is given by

$$Y = \sum_{r=1}^{R} Y_r = \sum_{r=1}^{R} \left(T_r R_r + w_{M,r} H_{M,r} + w_{A,r} H_{A,r}\right). \tag{46}$$

Equilibrium Given the unknowns $[R_r, w_{A,r}, w_{M,r}, H_{M,r}, H_{A,r}]$ we can calculate (Y, P_A, P_M) - see (46), (45), (35) and (36). Hence, we have $5 \times R$ unknowns. The five corresponding equilibrium conditions are the $2 \times R$ equilibrium conditions (43) and (44), the $2 \times R$ labor market clearing conditions (41) and (42) and the optimality condition for agricultural inputs (37). Note also that we can still pick a numeraire. If we multiply all wages and rental rates by a constant c, aggregate income Y is scaled by c (see (46)), the goods market clearing conditions (43) and (44) are unaffected, labor supply is unaffected (as only relative wages matter) and so are the optimality conditions (37). Hence, we can impose the normalization (see (45))

$$P_A = \left(\sum_{r=1}^{R} p_{s,r}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \left(\sum_{r=1}^{R} \left(\left(\frac{w_{A,r}}{\gamma}\right)^{\gamma} \left(\frac{R_r}{1-\gamma}\right)^{1-\gamma}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = 1. \tag{47}$$

Because (37) relates the endogenous demand for agricultural labor directly to the amount of available land and the agricultural wage (relative to the rental rate), it is useful to define the two relative prices

$$x_r \equiv \frac{R_r}{w_{A,r}} \text{ and } u_r = \frac{w_{A,r}}{w_{M,r}}. \tag{48}$$

Hence, we can write aggregate income Y_r and the aggregate prices P_M and P_A, i.e. (46), (35) and (36), as

$$\begin{aligned} Y &= \sum_{r=1}^{R}\left(T_r R_r + w_{M,r}H_{M,r} + w_{A,r}H_{A,r}\right) = \sum_{r=1}^{R} w_{M,r}\left(H_{M,r} + \frac{1}{1-\gamma}T_r u_r x_r\right) && (49)\\ P_A &= \left[\sum_r \left(\Gamma w_{M,r}u_r x_r^{1-\gamma}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} && (50)\\ P_M &= \left(\sum_{r=1}^{R}\left(\frac{1}{\overline{\zeta}}\frac{w_{M,r}}{Z_r}\left(H_r^M\right)^{-\vartheta}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, && (51)\end{aligned}$$

where $\Gamma \equiv \left(\frac{1}{\gamma}\right)^{\gamma}\left(\frac{1}{1-\gamma}\right)^{1-\gamma}$ and $\overline{\zeta}$ is given in (34). Using the demand equation (43) and the definition of the price index we get

$$\overline{\zeta}Z_r\left(H_r^M\right)^{\vartheta+1} = p_{M,r}^{-\sigma}\left(1-\alpha\right)\times YP_M^{\sigma-1} = \left(\frac{1}{\overline{\zeta}}\frac{w_{M,r}}{Z_r}\left(H_r^M\right)^{-\vartheta}\right)^{-\sigma}\left(1-\alpha\right)\times YP_M^{\sigma-1},$$

so that

$$\left(H_r^M\right)^{1-\vartheta(\sigma-1)} = \overline{\zeta}^{\sigma-1}Z_r^{\sigma-1}w_{M,r}^{-\sigma}\left(1-\alpha\right)\times YP_M^{\sigma-1}.$$

Hence, we can write the equilibrium system as (see (43), (44), (41) and (42)) as

$$\begin{aligned} \left(H_r^M\right)^{1-\vartheta(\sigma-1)} &= \overline{\zeta}^{\sigma-1}Z_r^{\sigma-1}w_{M,r}^{-\sigma}P_M^{\sigma-1}\times\left(1-\alpha\right)Y && (52)\\ \left(\frac{\gamma}{1-\gamma}\right)^{\gamma}\times x_r^{\gamma}\times T_r &= \left(\Gamma w_{M,r}u_r x_r^{1-\gamma}\right)^{-\sigma}\times\alpha Y\\ H_{M,r} &= \nu L_r^N\times\Lambda_M\left(\mu_r,u_r\right)\\ x_r\frac{\gamma}{1-\gamma}T_r &= \nu L_r^N\times\Lambda_A\left(\mu_r,u_r\right), && (53)\end{aligned}$$

where

$$\begin{aligned} \Lambda_M\left(\mu_r,u_r\right) &= \left(\frac{1}{1+\psi_r^N u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}}\left(Q_{M,r}^N\right)^{\frac{1}{\theta}} + \frac{\mu_r}{1-\mu_r}\times\left(\frac{1}{1+\psi_r^R u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}}\left(Q_{M,r}^R\right)^{\frac{1}{\theta}} && (54)\\ \Lambda_A\left(\mu_r,u_r\right) &= \left(\frac{\psi_r^N u_r^{\theta}}{1+\psi_r^N u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}}\left(Q_{A,r}^N\right)^{\frac{1}{\theta}} + \frac{\mu_r}{1-\mu_r}\times\left(\frac{\psi_r^R u_r^{\theta}}{1+\psi_r^R u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}}\left(Q_{A,r}^R\right)^{\frac{1}{\theta}}, && (55)\end{aligned}$$

and Y and P_M are given in (49) and (51). To derive these equations we used that the agricultural good is the numeraire and that

$$\begin{aligned} \pi_{M,r}^g &= \frac{Q_{M,r}^g w_{M,r}^{\theta}}{Q_{A,r}^g\left(w_{A,r}\left(1-\tau^g\right)\right)^{\theta} + Q_{M,r}^g w_{M,r}^{\theta}} = \frac{1}{1+\psi_r^g u_r^{\theta}}\\ \pi_{A,r}^g &= 1-\pi_{M,r}^g = \frac{\psi_r^g u_r^{\theta}}{1+\psi_r^g u_r^{\theta}}\\ \psi_r^g &\equiv \frac{Q_{A,r}^g}{Q_{M,r}^g}. && (56)\end{aligned}$$

This equilibrium system can be written in the following simplified form.

Lemma 3. *Consider (52)-(53) and (54) and (55). Define, the scaled fundamentals*

$$\tilde{L}_r^N = L_r^N \times \left(Q_{M,r}^N\right)^{\frac{1}{\theta}} \quad \textit{and} \quad \tilde{T}_r = T_r \left(Q_{M,r}^N / Q_{A,r}^N\right)^{\frac{1}{\theta}} \quad \textit{and} \quad \tilde{Z}_r^{\sigma-1} = Z_r^{\sigma-1} \left(Q_{M,r}^N / Q_{A,r}^N\right)^{(\sigma-1)/\theta} \tag{57}$$

and the scaled manufacturing wage

$$\tilde{w}_{M,r}^{-\sigma} \quad = \quad w_{M,r}^{-\sigma} \left(\frac{Q_{A,r}^N}{Q_{M,r}^N}\right)^{(\sigma-1)/\theta}, \tag{58}$$

the equilibrium vector $\left(N_r, \pi_{M,r}^N, \tilde{w}_{M,r}, x_r\right)$ *is given by*

$$\begin{aligned}
\left(\rho\zeta N_r\right)^{1-\vartheta(\sigma-1)} &= \bar{\zeta}^{\sigma-1} \tilde{Z}_r^{\sigma-1} \tilde{w}_{M,r}^{-\sigma} P_M^{\sigma-1} \times (1-\alpha) Y \\
\left(\frac{\gamma}{1-\gamma}\right)^{\gamma} \times x_r^{\gamma+\sigma(1-\gamma)} \times \tilde{T}_r &= \Gamma^{-\sigma} \tilde{w}_{M,r}^{-\sigma} \left(\frac{1-\pi_{M,r}^N}{\pi_{M,r}^N}\right)^{-\sigma/\theta} \times \alpha Y \\
\rho\zeta N_r &= \nu \tilde{L}_r^N \times \lambda_M \left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right) \\
x_r \frac{\gamma}{1-\gamma} \tilde{T}_r &= \nu \tilde{L}_r^N \times \lambda_A \left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right)
\end{aligned}$$

where

$$\lambda_M \left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right) \quad = \quad \left(\pi_{M,r}^N\right)^{\frac{\theta-1}{\theta}} + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\pi_{M,r}^N}{\left(1-\pi_{M,r}^N\right) \times \chi_r + \pi_{M,r}^N}\right)^{\frac{\theta-1}{\theta}} (q_r)^{\frac{1}{\theta}} \tag{59}$$

$$\lambda_A \left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right) \quad = \quad \left(1-\pi_{M,r}^N\right)^{\frac{\theta-1}{\theta}} + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\chi_r \left(1-\pi_{M,r}^N\right)}{\left(1-\pi_{M,r}^N\right) \times \chi_r + \pi_{M,r}^N}\right)^{\frac{\theta-1}{\theta}} (q_r)^{\frac{1}{\theta}} (\chi_r)^{\frac{1}{\theta}}, \tag{60}$$

and

$$\chi_r = \frac{Q_{M,r}^N / Q_{A,r}^N}{Q_{M,r}^R / Q_{A,r}^R} \textit{ and } q_r = \frac{Q_{M,r}^R}{Q_{M,r}^N}. \tag{61}$$

Proof. See Section 8.11 in the Online Appendix. □

Using the system in Lemma 3, we can write it as

$$\begin{aligned}
(1-\vartheta(\sigma-1)) \times ln(N_r) &= c_1 + (\sigma-1) ln(Z_r) - \sigma ln(w_{M,r}) + (\sigma-1) ln(P_M) + ln(Y) && (62) \\
(\gamma + \sigma(1-\gamma)) ln(x_r) + ln(T_r) &= c_2 - \sigma ln(w_{M,r}) - \frac{\sigma}{\theta} ln\left(\frac{1-\pi_{M,r}^N}{\pi_{M,r}^N}\right) + ln(Y) && (63) \\
ln(N_r) &= c_3 + ln\left(L_r^N\right) + ln\left(\lambda_M\left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right)\right) && (64) \\
ln(x_r) + ln(T_r) &= c_4 + ln\left(L_r^N\right) + ln\left(\lambda_A\left(\mu_r, \pi_{M,r}^N, \chi_r, q_r\right)\right), && (65)
\end{aligned}$$

where we for collect all constants in the c_j terms and drop the "~" over the scaled variables.

Now note that

$$\begin{aligned}
ln\left(\frac{1-\pi_{M,r}^N}{\pi_{M,r}^N}\right) &\approx ln\left(\frac{1-\bar{\pi}_M^N}{\bar{\pi}_M^N}\right) - \frac{1}{\left(1-\bar{\pi}_M^N\right)\bar{\pi}_M^N}\left(\pi_{M,r}^N - \bar{\pi}_M^N\right) \\
&= \psi_\pi - \frac{1}{\left(1-\bar{\pi}_M^N\right)\bar{\pi}_M^N} \pi_{M,r}^N.
\end{aligned}$$

As we show in Section 8.12 below, the log-linear approximation of $ln\left(\lambda_{M,r}\right)$ and $ln\left(\lambda_{A,r}\right)$ is given by

$$
\begin{aligned}
ln\left(\lambda_M\left(\mu_r,\pi_{M,r}^N\right)\right) &\approx \psi_{\lambda_M}+\frac{\overline{\omega}_M^R}{\left(1-\overline{\mu}\right)\overline{\mu}}\times\mu_r+\frac{\theta-1}{\theta}\frac{\overline{\varphi}_M}{\overline{\pi}_M^N\left(1-\overline{\pi}_M^N\right)}\times\pi_{M,r}^N\\
ln\left(\lambda_A\left(\mu_r,\pi_{M,r}^N\right)\right) &\approx \psi_{\lambda_A}+\frac{\overline{\omega}_A^R}{\left(1-\overline{\mu}\right)\overline{\mu}}\times\mu_r-\frac{\theta-1}{\theta}\frac{\overline{\varphi}_A}{\overline{\pi}_M^N\left(1-\overline{\pi}_M^N\right)}\times\pi_{M,r}^N,
\end{aligned}
$$

where

$$
\begin{array}{lll}
\omega_M^R=\frac{H_M^R}{H_M} \text{ and } & \overline{\varphi}_M=\left(1-\overline{\omega}_M^R\right)\overline{\pi}_A^N+\overline{\omega}_M^R\overline{\pi}_A^R\\
\omega_A^R=\frac{H_A^R}{H_A} \text{ and } & \overline{\varphi}_A=\left(1-\overline{\omega}_A^R\right)\overline{\pi}_M^N+\overline{\omega}_A^R\overline{\pi}_M^R
\end{array},
$$

and ψ_{λ_M} and ψ_{λ_A} do not vary across regions. The equilibrium system in (62)-(65) can then be written as

$$
\begin{aligned}
\left(1-\vartheta\left(\sigma-1\right)\right)\times ln\left(N_r\right) &= c_1+\left(\sigma-1\right)ln\left(Z_r\right)-\sigma ln\left(w_{M,r}\right)+\left(\sigma-1\right)ln\left(P_M\right)+ln\left(Y\right) && (66)\\
\left(\gamma+\sigma\left(1-\gamma\right)\right)ln\left(x_r\right)+ln\left(T_r\right) &= c_2-\sigma ln\left(w_{M,r}\right)+\frac{\sigma}{\theta}\frac{1}{\left(1-\overline{\pi}_M^N\right)\overline{\pi}_M^N}\pi_{M,r}^N+ln\left(Y\right) && (67)\\
ln\left(N_r\right) &= c_3+ln\left(L_r^N\right)+\frac{\overline{\omega}_M^R}{\left(1-\overline{\mu}\right)\overline{\mu}}\times\mu_r+\frac{\theta-1}{\theta}\frac{\overline{\varphi}_M}{\overline{\pi}_M^N\left(1-\overline{\pi}_M^N\right)}\times\pi_{M,r}^N && (68)\\
ln\left(x_r\right)+ln\left(T_r\right) &= c_4+ln\left(L_r^N\right)+\frac{\overline{\omega}_A^R}{\left(1-\overline{\mu}\right)\overline{\mu}}\times\mu_r-\frac{\theta-1}{\theta}\frac{\overline{\varphi}_A}{\overline{\pi}_M^N\left(1-\overline{\pi}_M^N\right)}\times\pi_{M,r}^N. && (69)
\end{aligned}
$$

Solving for the endogenous variables in terms of fundamentals yields the following elasticities (see Section 8.13 in the Online-Appendix for the detailed derivation):

$$
\begin{aligned}
\frac{\partial N_r}{\partial\mu_r}=\psi_{N,\mu} &= \frac{\phi}{\left(1-\overline{\mu}\right)\overline{\mu}\left(1-\overline{\pi}_M^N\right)\overline{\pi}_M^N}\frac{1}{\theta}\left(\overline{\omega}_M^R\sigma+\left(\gamma+\sigma\left(1-\gamma\right)\right)\left(\theta-1\right)\left(\overline{\omega}_M^R\overline{\pi}_M^N+\overline{\omega}_A^R\overline{\pi}_A^N\right)\right)\\
\frac{\partial w_{M,r}}{\partial\mu_r}=\psi_{w,\mu} &= \frac{1}{\sigma}\psi_{N,\mu}\times\left(\vartheta\left(\sigma-1\right)-1\right)\\
\frac{\partial\pi_{M,r}^N}{\partial\mu_r}=\psi_{\pi,\mu} &= \phi\frac{1}{\left(1-\overline{\mu}\right)\overline{\mu}}\frac{\overline{\omega}_A^R}{\overline{\omega}_M^R}\times\left(\vartheta\overline{\omega}_M^R+\overline{\omega}_A^R\left(1-\gamma\right)-\frac{\overline{\omega}_A^R\left(1-\overline{\omega}_M^R\right)}{\sigma-1}\left(\chi^{Ref}-1\right)\right).
\end{aligned}
$$

Finally, we can turn to the relationship between refugees and income per capita. This analysis is contained in Section 8.15 in the Online-Appendix.