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

Michael Peters*

June 2017

**Abstract**

Can increases in the size of the local workforce raise productivity and spur economic development? This paper uses a particular historical episode to study this question empirically. After the Second World War, between 1945 and 1948, about 8m Ethnic Germans were expelled from their domiciles in Middle and Eastern Europe and transferred to Western Germany. At the time, this inflow amounted to almost 20% of the Western German population. 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 outcomes in the 50s and 60s. 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. These patterns are consistent with theories of local agglomeration and endogenous technological change but hard to rationalize in a neoclassical framework with exogenous technology. To quantify the strength of such agglomeration forces, I calibrate a multi-region general equilibrium trade model to the cross-sectional data from 1939 and 1950. In order to replicate the observed relationship between refugee inflows and changes in manufacturing employment, the model requires an elasticity of manufacturing productivity with respect to the local workforce of about 0.2.

*eMail: m.peters@yale.edu. I am very grateful to Abhijit Banerjee, Penny Goldberg, Tarek Hassan, Virgiliu Midrigan, Steve Redding, Gianluca Violante, Joachim Voth and especially Chad Jones, Pete Klenow and Alwyn Young for their comments. I also thank seminar participants at Chicago Booth, Fordham, Harvard, MIT, NYU, Princeton, Stanford, 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. This paper studies a particular historical episode to provide direct empirical evidence for the importance of such mechanisms.

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 forced population movements in world history. By 1950, about 8m people had been transferred to Western Germany. Given the population at the time, this amounted to an increase in the total population of 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.

For this study 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. Two features of the historical context make this empirical variation very suitable for the question of interest. The first is the sheer size of the initial population shock. As many of the theoretical mechanisms above stress the importance of general equilibrium effects, one requires changes in factor supplies, which are sufficiently large to plausibly have aggregate consequences. The second concerns the 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 availabity of housing rather than the structure of the local economy. Moreover, the Military Governments implemented tight mobility controls, which essentially ruled out refugees' spatial adjustment before 1950.

These aspects of the historical context imply 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 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 novel data on the extent of war-time destruction for all Western 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. I also consider an instrumental variable strategy, where I exploit the distance to the pre-war population centers in Eastern and Central Europe interacted with properties of the local housing supply after the war.

To perform my analysis, I construct a novel dataset 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 measure 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 these publications contains information on sectoral and occupational employment shares, sex-ratios, population density and other county 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, 1950 and 1956 waves of the German census of manufacturers.

Using this data, my results suggest that local manufacturing productivity did increase in response to the inflow of refugees. First of all, I establish a positive relationship between the allocation of refugees and subsequent manufacturing growth. In particular, refugee-receiving counties experience faster increases in manufacturing employment in the 1950s and 60s and experienced higher rates of plant entry after the Second World War. Secondly, I provide evidence for a positive relationship between refugee inflows in 1950 and local GDP per capita in 1961. At the same time, I also show refugees earned less than natives, suggesting that the inflow of refugees did not increase the average human capital in a region. Taken together these patterns are consistent with an endogenous local productivity response but hard to rationalize in a neoclassical framework with constant technologies.

The reason why an inflow of refugees might increase income per capita despite the decline in average regional human capital is the interaction between external agglomeration forces and workers' occupational choice. As firms' do not internalize the effect of manufacturing hiring on local productivity, equilibrium wages in the manufacturing sector are too low from a social point of view. Local labor supply shifts, which increase manufacturing employment can therefore be beneficial.

To quantify the strength of the underlying productivity response, I then calibrate a canonical multi-region general equilibrium trade model to these cross-county patterns. 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, in particular their absolute and comparative advantage. While the agricultural sector uses a fixed factor ("land") and is subject to decreasing returns, productivity in the manufacturing sector is endogenous and depends on the size the manufacturing sector. One interpretation of this relationship posits that 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, and - crucially - in the number of refugees in the population. Furthermore, regional productivity also depends directly on other aspects of the post-war period, which are correlated with the allocation of refugees, in particular the extent of war-time destruction.

The model makes tight predictions on the spatial co-movement between refugees, population density, manufacturing employment, the extent of war-time destruction and income per capita. I calibrate the model to the cross-sectional data in 1939 and 1950. The key moment to discipline the strength of agglomeration is the estimated cross-sectional relationship between manufacturing employment and the allocation of refugees. The calibrated model implies a sizable productivity response, whereby the elasticity of local manufacturing productivity with respect to local labor supply is about 0.2.

Finally, I also present direct micro evidence on why refugees' labor supply was biased towards the manufacturing sector. 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 that covers the time of the 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 faced barriers to work in agriculture. As the agricultural sector in Germany was widely dominated by small, family farms, such barriers plausibly took the form of frictions in the agricultural land market. 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 models of trade and 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); Faber and Gaubert (2016); Kucheryavyy et al. (2016); Ramondo et al. (2016) 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 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

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

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) 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 variation 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 describe in detail the initial allocation of refugees across counties in Western Germany. Section 4 contains the empirical analysis. I first provide a very simple theoretical model, which motivates the empirical analysis. Using this framework, I then analyze the relationship between the inflow of refugees, income per capita, manufacturing employment and the entry of new plants. In Section 5 I then provide a richer general equilibrium model, which nests the simple model in Section 4 as a special case, and calibrate the structural parameters to the regression evidence from the historical experiment. Section 6 concludes.

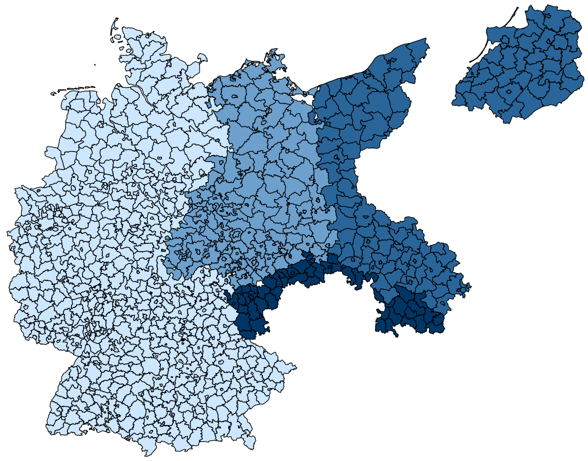
## 2 The Historical Setting

**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 sizable German minorities in other countries of Eastern Europe, most importantly the so-called Sudetenland in Czechoslovakia. This region in the north of Czechoslovakia has a long tradition of German settlements and was annexed by the Nazi Government in 1938.

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

[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 annexed by Germany in 1938. During the Potsdam Conference in 1945, Germany lost the Eastern Territories of the German Reich and the Sudetenland.

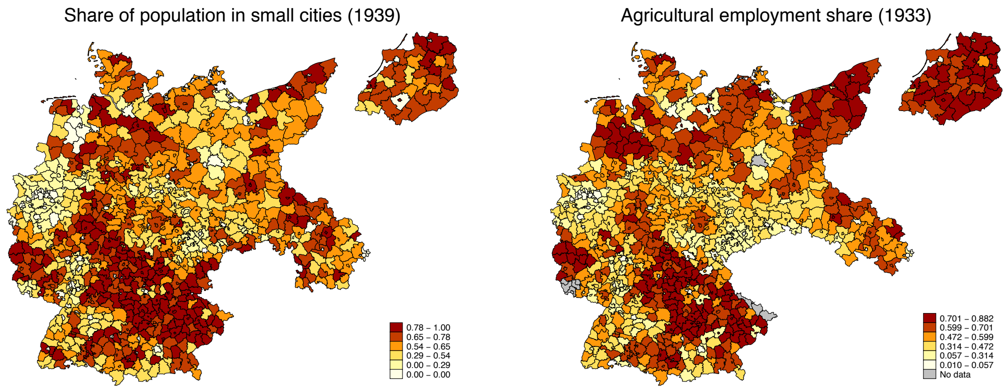
Figure 1: The German Reich in 1939

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 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 14 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 map on the right depicts the distribution of regional agricultural employment shares 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, which is shown 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.

### 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



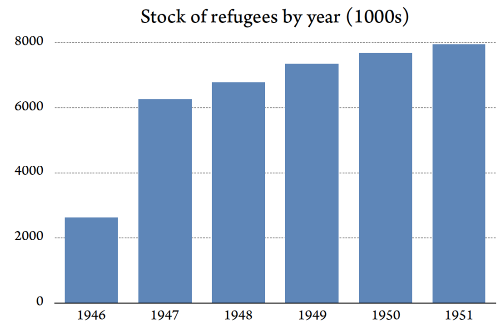

Notes: The left map shows a measure of urbanization in 1939, namely the share of the county population living in cities with less than 2000 inhabitants. The map on the right displays the agricultural employment share in 1933.

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

| | 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




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

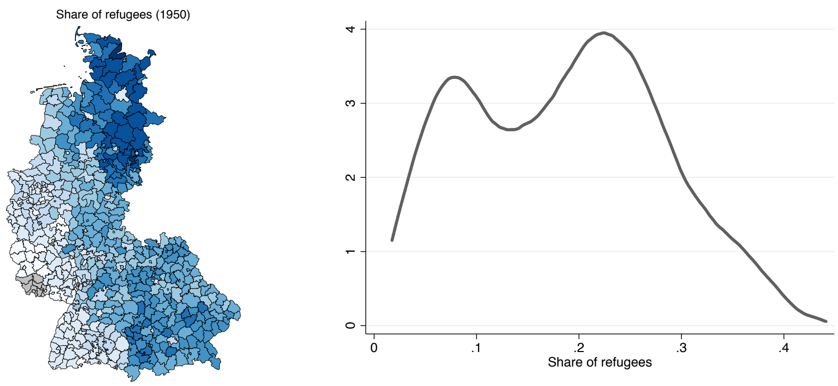
the Eastern Territories were allocated to both Poland and Russia respectively. Int that process, the governments of Russia, the UK and US decided to expel the German population from these territories. The official protocol of the Potsdam conference reads:

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

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, 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

[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: 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. In the right panel I show the marginal distribution of the share of refugees. Source: Statistisches Bundesamt (1955)

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

amounted to roughly 20% of the population living in Western Germany at the time. Despite the casualties during the war, the population of Western Germany had therefore increased substantially from 1939 to 1950 (Steinberg, 1991).

In this paper I will exploit the cross-sectional variation in refugee flows across counties in Western Germany, which I display in Figure 4. In the left panel I depict the geographic allocation of refugees across Western German counties in 1950. The right panel displays the corresponding marginal distribution. 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 systematically correlated with pre-existing differences across counties. This is where I turn now.

# 3 Refugees in Western Germany

## 3.1 Data

In this paper I use a variety of datasets. The majority of the analysis exploits the spatial variation and links refugee flows in 1950 to economic outcomes at the county-level in the 50s, 60s and 70s. Depending on the time period, there are roughly 500 counties in Western Germany. 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 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 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 the entry of manufacturing plants at the regional level.

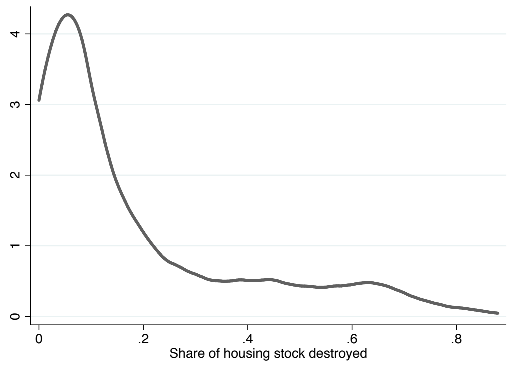
Finally, I also exploit the information on war time destruction and housing supply, which I digitized from the historical housing census conducted in 1950. This census contains information on the extent of war damages for each county and detailed information on living conditions of refugees and natives. 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.

### 3.2 The Initial Allocation of Refugees Across Counties

To use the cross-sectional variation in refugee flows as an empirical shifter of local labor supply, the initial allocation of refugees should be orthogonal to other county characteristics, in particular local productivity. The historical context is crucial to tease out a component of the cross-sectional distribution of refugees, which is exogenous with respect to other local supply conditions. Two features are important. First of all, until the early 1950s, refugees'




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

labor mobility was severely limited. Not only was the vast majority of refugees allocated to particular localities upon their arrival in Western German by the respective Military Governments[4], but moving restrictions prevented any substantial reallocation of refugees until after 1950. Secondly, the historical sources overwhelmingly suggest that an orderly allocation of the refugees across localities in Western Germany was impossible at the time. The main reason was the administrative burden. In fact, the striking East-West trajectory shown in Figure 4 already suggests the Military Governments of the US and the UK did not manage to redistribute the incoming refugees across localities in Western Germany.

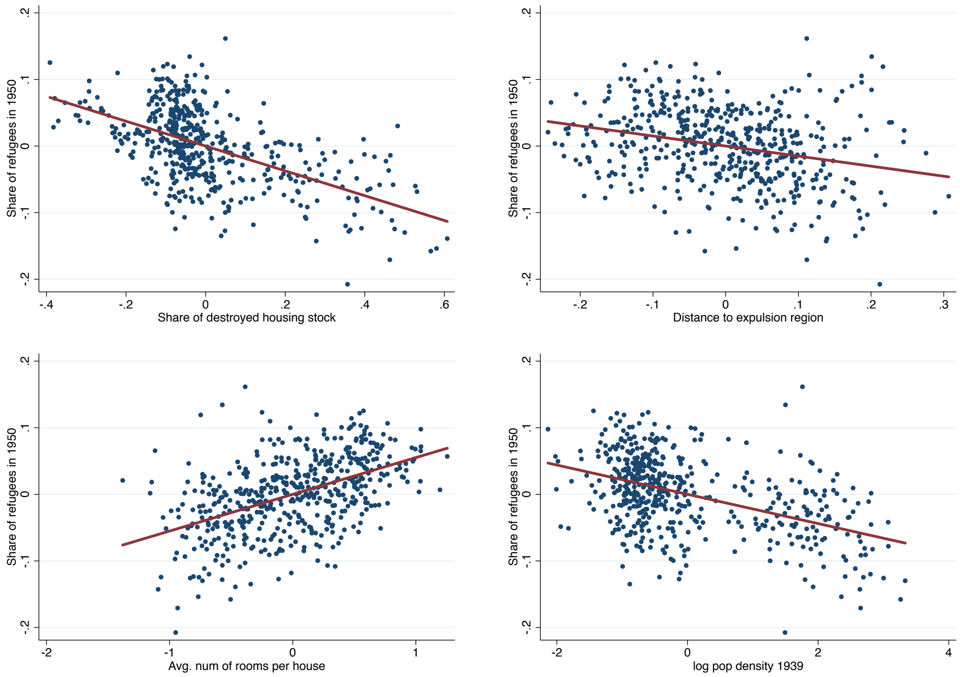
Werner Nellner, one of the leading post-war economic historians, describes the situation as follows: "In the midst of the chaotic post-war circumstances arrived the refugee transports. The entirely confusing political and economic situation paired with the abruptness of this pouring-in simply did not 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 the availability of housing supply leaves particular trails in the data. Most importantly, refugees were naturally sent to places where it was easies to find shelter, i.e. places with historically low population density. Secondly, data from the German housing census at the county-level shows that about 23% of the aggregate housing stock was damaged during the Allied bombing campaign. Moreover, there is considerable heterogeneity in the extent to which Western German counties were affected. 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.[5]

Thirdly, counties differed substantially in the structure of their houses, in particular the number of rooms *per* house. The reason why the existence of such marginal rooms within natives' houses were an important determinant of the allocation of refugees is that initially many refugees were housed *within* the apartments of natives whenever spare rooms were available. While a family of four was often forced to accept refugees into their three bed-room

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

[5]Expectedly, the extent of war-time damages is strongly correlated with pre-war population density. 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 figure depicts the relationship between the share of refugees and the extent of war-time destruction (upper left panel), the average number of rooms per house (lower left panel), the population density in 1939 and the population-weighted distance to the expulsion region calculated according to (1) (upper right panel). All plots show residual variation after taking out a set of state fixed effects.

Figure 6: Regional housing supply and the Allocation of Refugees

apartment, this was less the case if only 2 bedrooms were available. Finally, and as already hinted at above, the geographic location of a particular county played a key role the expected inflow of refugees: counties close the expulsion regions in the East, received far more refugees than the average county.

This is exactly what we see in the data. In Figure 6, I depict the correlation of the share of refugees with (i) the extent of war-time destruction (upper left panel), (ii) the average number of rooms per house (lower left panel), (iii) the population density in 1939 (lower right panel) and (iv) the population-weighted distance to the expulsion region (upper right panel). This distance measure is calculated as

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

where $d_{c,j}$ is the geographical distance between county $j$ and $r$ and $pop_r^{1930}$ is the size of the population in 1939. The correlations are as expected: rural counties close to the Eastern border with little war time destruction and large houses were natural places for refugees to be allocated to.

These systematic correlations imply that refugee inflows were correlated with regional productivity as productive locations tend to have higher population density or were more severely targeted by the bombing campaign. Similarly, a region's distance to the settlements in Eastern Europe might have direct effects on its effective productivity if - as argued by Redding and Sturm (2008) - the division of Germany after the Second World War caused a decline in market access. This suggests that the allocation of refugees was negatively correlated with local productivity, a conjecture, which I confirm in the calibrated model in Section 5.

What Figure 6 does not show is that the allocation of refugees is *un*correlated with a regions' extent of sectoral specialization once the initial level of pre-determined population density is controlled for. To see this, I consider regressions of the form

$$y_r = \delta_s + \beta \times \mu_r + \phi \times \ln \text{pop dens}_r^{33} + \varphi \times \text{exp\_dist}_r + \upsilon_r, \tag{2}$$

where $y_r$ is a set of pre-war outcomes between 1933 and 1939 at the regional level, $\mu_r$ is the share of refugees in county $r$ in 1950, $\delta_s$ is a set of state fixed effects and pop dens$^{33}$ is the population density in 1933, i.e. predetermined with respect to all outcomes $y_r$. Finally, $\text{exp\_dist}_r$ is the expulsion distance measure calculated in (1). The results of (2) are contained in Table 2 below.

The results are reassuring in that the allocation of refugees is uncorrelated with other county outcomes in the pre-war period. In particular, there is no correlation with manufacturing or agricultural employment shares prior to the war (columns 1 - 3) or the number of manufacturing plants in either 1933 or 1939 (columns 4 and 5). Similarly, there is no correlation with the growth rate of manufacturing employment once population growth is controlled for (column 7). The latter is important, because refugee flows are systematically correlated with the level of population density in 1939 and hence with the population growth rate for a given level of population density in 1933 (column 8).

These results are consistent with the view that refugees were not systematically allocated towards regions with a particular economic structure holding population density fixed. In my first empirical strategy I will therefore use the residual variation in refugee flows after controlling for the determinants of refugee flows displayed in Figure 6. This strategy is valid if this residual variation is uncorrelated with other factors of county productivity. The historical accounts of that episode suggest that this is the case. Not only was the uncoordinated assignments of refugees 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. ... 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, 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)

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 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 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 stark 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

| | Empl. share in | | | ln man. plants | | Change in man. share | | Pop. growth |
|---|---|---|---|---|---|---|---|---|
| | Man. (1933) | Ag. (1933) | Man. (1939) | 1933 | 1939 | 1933-1939 | | 1933-1939 |
| Share of refugees in 1950 | 0.045 | 0.074 | -0.119 | 0.079 | -0.311 | -0.164** | -0.014 | -1.011*** |
| | (0.095) | (0.151) | (0.086) | (0.558) | (0.539) | (0.072) | (0.069) | (0.351) |
| ln pop dens. 1933 | 0.065*** | -0.155*** | 0.047*** | 0.056* | 0.082** | -0.018*** | | -0.032* |
| | (0.005) | (0.009) | (0.004) | (0.034) | (0.033) | (0.004) | | (0.017) |
| lnExpulsionDistance | -0.114** | -0.002 | -0.138*** | 0.176 | 0.153 | -0.025 | 0.000 | -0.006 |
| | (0.049) | (0.081) | (0.043) | (0.257) | (0.259) | (0.028) | (0.029) | (0.115) |
| Population growth 1933-39 | | | | | | | -0.040** | |
| | | | | | | | (0.017) | |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 510 | 510 | 509 | 507 | 494 | 509 | 509 | 510 |
| $R^2$ | 0.507 | 0.630 | 0.468 | 0.227 | 0.277 | 0.102 | 0.052 | 0.037 |

Standard errors in parentheses

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

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 refugees in 1950" is the share of refugees in 1950. "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. "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.

Table 2: The Initial Allocation of Refugees Across German Counties

| | 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

across space according to their pre-war occupation was unlikely to be quantitatively important.

**An Instrumental Variable Strategy** Alternatively, I will also exploit a complimentary instrumental variable strategy. In Figure 4 I showed a strong correlation between refugee flows and the distance to the expulsion regions and the average number of rooms per house. As explained above, the number of rooms per existing house is a determinant of refugee flows in as far as the native population was forced to accept refugees into their homes. This margin of housing supply was, however, only tapped into when other options to house refugees were exhausted. Hence, the average size of existing houses was a more important determinant of refugee flows the higher the potential of a given county to receive refugee inflows. Because this potential was particularly high for counties geographically close to the expulsion regions, one can expect the *interaction* between the average house size and the distance to the expulsion regions to be a powerful predictor of refugee flows.

In Table 4 I show that this is the case. Specifically, I regress the share of refugees on the four variables depicted in Figure 6 and a set of state fixed effects. Column 1 shows that the average number of rooms positively predicts the share of refugees and the expulsion distance negatively correlated with the share of refugees. The next four columns run the exact same regression for counties close and far away from border. In particular, I group counties according to quantiles of the distance to the expulsion regions. For counties closer than the median or 25% county (columns 2 and 4), the average number of rooms is a significant predictor of the share of refugees. For counties far away (column 3 and 5) this is not the case. This is consistent with the view that private homes were only a source of housing supply when the stream of potential refugees was particularly large. In the last 3 columns, I again use the full sample but allow for an interaction between the average number of rooms and the expulsion distance. As expected this interaction term is negative and highly significant. Note also that the main effect of the expulsion distance is no longer significant. In the last column I consider a different measure for the supply of marginal housing units: the share of houses with at least four rooms. Again, I find that this positively predicts the share of refugees and particularly so in areas close to the expulsion regions.

I will use these results as an instrumental variable strategy. In particular, under the assumption that regions with small and large houses are not affected *differentially* by their geographical location after the war, the *interaction* between the distance to the expulsion region and average room size is a valid instrument for the allocation of refugees.

### 3.3 Migration and Persistence

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? And for how long did these local labor supply shocks last?

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 as the political support for the housing allocation was only provided in the locations refugees were initially assigned to. Similarly, there were restriction to receive food stamps without being officially registered (Grosser and Schraut, 2001, p. 83).[6] 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).

In Table 5 I provide direct evidence that migratory responses were limited and for the persistence of initial population shocks. 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, I was able to find a information on the share of refugees in 1946 for a subset of states. In column 1, I shows that the within-state correlation between the refugee share in 1950 and 1946 is close to unity. Columns 2 and 3 repeat this exercise for the share of refugees in 1955 and 1961. Even 10-15 years after the expulsion, there is a strong correlation between the initial allocation of refugees and the refugees still residing in a particular county. Finally, in columns 4 and 5 I show that the share of refugees is a strong predictor of population growth since 1939, i.e. the initial refugee shock had a long persistence on the spatial allocation of economic activity across Western German counties.

[6] 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).

| | Share of refugees | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Full Sample | | Expulsion Distance | | | | Full Sample | |
| | | $< q^{50}$ | $> q^{50}$ | $< q^{25}$ | $> q^{75}$ | | | |
| Avg num of rooms (rooms/apt) | 0.037*** | 0.028*** | 0.014 | 0.049*** | 0.007 | 0.248*** | 0.260*** | |
| | (0.006) | (0.008) | (0.009) | (0.011) | (0.007) | (0.044) | (0.045) | |
| Avg rooms * Exp. Dist. | | | | | | -0.045*** | -0.047*** | |
| | | | | | | (0.009) | (0.009) | |
| Share of apts with more than 4 rooms | | | | | | | | 5.228** |
| | | | | | | | | (2.050) |
| Multiroom share * Exp. Dist. | | | | | | | | -0.251** |
| | | | | | | | | (0.100) |
| lnExpulsionDistance | -0.108*** | -0.017 | -0.442*** | -0.165*** | -0.299*** | -0.035 | -0.029 | 0.032 |
| | (0.024) | (0.032) | (0.043) | (0.063) | (0.101) | (0.027) | (0.028) | (0.063) |
| ln pop dens. 1939 | 0.001 | -0.002 | -0.008 | 0.004 | -0.009** | -0.001 | -0.001 | -0.005* |
| | (0.003) | (0.004) | (0.005) | (0.005) | (0.004) | (0.003) | (0.003) | (0.003) |
| Share of housing stock damaged | -0.123*** | -0.177*** | -0.064*** | -0.149*** | -0.034* | -0.123*** | -0.132*** | -0.125*** |
| | (0.015) | (0.026) | (0.018) | (0.044) | (0.018) | (0.014) | (0.016) | (0.015) |
| Share of housing stock built after 1945 | | | | | | | 0.136 | |
| | | | | | | | (0.097) | |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 511 | 255 | 255 | 127 | 127 | 511 | 511 | 511 |
| $R^2$ | 0.778 | 0.662 | 0.796 | 0.616 | 0.657 | 0.785 | 0.787 | 0.770 |

Note: Robust Standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The regression is at the county level. "Share of housing stock damaged" is the share of the housing stock, which was damaged during the war. "Avg num of rooms" denotes the average number of rooms per house within the county. "ln Expulsion Distance" denotes the distance to the expulsion regions (see (1)). "Avg rooms * Exp. Dist." denotes the interaction between "ln Expulsion Distance" and "Avg num of rooms". "Multiroom share * Exp. Dist" denotes the interaction between "ln Expulsion Distance" and "Share of apts with more than 4 rooms". $q^{\tau}$ denotes the $\tau$th quantile of the expulsion distance. Hence, columns 2 to 5 condition the sample by the expulsion distance. "State FE" indicates wether the regression controls for state fixed effects.

Table 4: The Allocation of Refugees and Housing Supply

| | Share of refugees in | | | Pop. growth | |
|---|---|---|---|---|---|
| | 1946 | 1955 | 1961 | 1939-46 | 1939-61 |
| Share of refugees in 1950 | 0.970*** | 0.708*** | 0.563*** | 1.823*** | 1.075*** |
| | (0.039) | (0.030) | (0.038) | (0.107) | (0.161) |
| ln pop dens. 1939 | | | | -0.054*** | 0.007 |
| | | | | (0.006) | (0.007) |
| State FE | Yes | Yes | Yes | Yes | Yes |
| Observations | 363 | 511 | 456 | 508 | 463 |
| $R^2$ | 0.880 | 0.843 | 0.688 | 0.812 | 0.224 |

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

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 5: Refugees and Population Growth

# 4 Economic Consequences in the 50s and 60s

I will now turn to the study of the outcomes of this historical experiment, i.e. how did the local economy respond to this inflow of people. I first show that the in-pour of refugees triggered an increase in the local manufacturing sector, both in the short and the long-run. Furthermore, I provide direct micro-evidence why refugees acted as a local supply shifter for the manufacturing sector. Then I show empirical evidence of a positive cross-sectional relationship between refugees and income per capita. Using microdata on relative earnings between refugees and native I also directly show that refugees earn less than natives. Though the lens of a simple model, these patterns are hard to rationalize without the existence of some agglomeration forces. To quantify the strength of these forces, I therefore calibrate a richer structural model in Section 5 to the cross-county evidence.

## 4.1 A (Very) Simple Theoretical Framework

To study the economic consequences of this inflow of refugees into the local economy, consider the following simple theoretical framework. In Section 5 below I consider a richer version of this model, which however nests the framework considered here as a special case. Consider a county within Western Germany. There are two sectors of production, agriculture and manufacturing. While agricultural production requires land to produce and is hence subject to decreasing returns, the manufacturing sector only requires labor an input. The production functions are given by

$$Y_A = H_A^{1-\gamma} T^{\gamma} \text{ and } Y_M = Z_M H_M,$$

where $T$ is the amount of land, which is non-traded, in fixed supply and rented at price $R$, and $H_j$ denotes the allocation of human capital across industries. The total supply of human capital is given by

$$H = L^N + QL^R,$$

where $L^N$ and $L^R$ denotes the number of natives and refugees and $Q$ parametrizes the differences in skills. Importantly, productivity in the manufacturing sector can potentially depend on the size of the local workforce, i.e.

$$Z_M = \overline{Z}_M \times H^{\vartheta}.$$

Hence, productivity is increasing in the size of the population if $\vartheta > 0$.[7] Furthermore, suppose that the county is small (and hence cannot affect the respective prices $p_A$ and $p_M$) and that factor markets are competitive. How does an inflow of refugees affect the sectoral employment structure of the economy and income per capita?

Consider first the equilibrium manufacturing employment share $\pi_M = \frac{H_M}{H}$. Because the marginal product of labor is equalized across sectors, it is easy to show that

$$\frac{d\pi^M}{dL^R} = \left(1 - \pi^M\right) \times \frac{\gamma + \vartheta}{\gamma} \frac{Q}{H}. \tag{3}$$

Hence, an inflow of refugees increases the manufacturing employment *share* if the agricultural sector is subject to decreasing returns, i.e. $\gamma > 0$. Furthermore, this effect is stronger, the stronger the productivity benefits of agglomeration, $\vartheta$. Now consider income per capita $y = \frac{wH+RT}{L^N+L^R}$. Because equilibrium wages are given by $w = Z_M p_M$ and optimality in the agricultural sector requires that $RT = \frac{\gamma}{1-\gamma} w H_A$, it follows that the change in income per capita is given by

$$\frac{dln\left(y\right)}{dL^R} = \frac{dln\left(w\right)}{dL^R} + \frac{dln\left(\frac{1-\gamma\pi^M}{1-\gamma}\right)}{dL^R} + \frac{dln\left(\frac{H}{L}\right)}{dL^R} = \frac{Q}{H}\left(\vartheta - \left(\gamma+\vartheta\right)\frac{1-\pi^M}{1-\gamma\pi^M} - \left(\frac{1-Q}{Q}\right)\left(1-\mu\right)\right), \tag{4}$$

where $\mu = \frac{L^R}{L}$ is the share of refugees in the population. Hence, the inflow of refugees affects income per capita though three distinct channels. First of all, there could be direct productivity effects to the extent that $\vartheta > 0$. Secondly, refugee inflows tend to reduce income per capita as one sector of the economy is subject to decreasing returns to scale. Finally, the effect on income per capita depends on the relative supply of human capital: if refugees are less skilled than the native population, i.e. $Q < 1$, an inflow of refugees reduces income per capita through a deterioration of the average local human capital supply. In particular, if there was no response of local productivity to the size of the population, i.e. $\vartheta = 0$, and refugees were less skilled (and hence earn less) than the native population, (3) and (4) necessarily imply that an inflow of refugees would increase manufacturing employment but decrease income per capita.

## 4.2 Refugees and Local Manufacturing

**Refugees and Manufacturing Employment in 1950** To establish the link between refugee inflows and the size of the manufacturing sector, I start by studying the relationship between the share of refugees and the manufacturing employment share in 1950, i.e. roughly 3 years after the initial shock. The empirical specification of interest is given by

$$\pi^{50}_{M,r} = \delta_s + \beta \times \mu^{50}_r + \alpha \times \pi^{39}_{M,r} + \phi \times \ln \text{pd}^{39}_r + \varphi \times \text{wd}_r + \tau \times \text{exp\_dist}_r + \chi \times h_r + \zeta \times C_r + v_r, \tag{5}$$

where $\pi^t_{M,r}$ denotes the manufacturing employment share in time $t$, $\mu^{50}_r$ is the share of refugees in 1950 and $\text{pd}^{39}$, $\text{wd}_r$, $\text{exp\_dist}_r$ and $h_r$ denote population density in 1939, the extent of wartime destruction, the distance to the expulsion

[7] Here I assume for simplicity that $Z_M$ depends on $H$. In the richer model in Section 5 I assume that $Z_M$ depends on $H_M$, which is endogenously determined.

regions and the average house size. Furthermore, $C_r$ denotes additional control variables. Because I explicitly control for $\pi_{M,r}^{39}$, the coefficient of interest $\beta$ captures the effect of refugees on the growth of manufacturing employment. I will estimate $\beta$ both via OLS and using the IV strategy outlined above, i.e. using the interaction between the expulsion distance and the regional housing supply as the excluded instrument. The results of estimating (5) are contained in Table 6.

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 $\mu_{cs}$ and the manufacturing share in 1939. Columns 2 runs the exact same specification using the 1950 manufacturing employment share as the dependent variable. Now there is a sizable positive effect: an increase in the share of refugees by 10 percentage points increases manufacturing employment by 3 percentage points. Column 3 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 3 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. Below I will provide direct evidence why this is the case. Finally, column 4 shows that this relationship is robust to a host of additional control, in particular a set of fixed effect for border counties, a flexible polynomial for the distance to the inner german border, the manufacturing and agricultural share in 1933, population density in 1933 and a measure of pre-war urbanization, namely the share of the county population living in small villages.

The next column contains the IV estimate. While the point estimate is larger than the OLS estimate, in particular the standard error increases substantially. In the next two columns, I use the refugee share in 1946 instead of 1950 and estimate (5) for the subset of counties for which I was able to gather the data in 1946. The results are qualitatively similar, even though the estimated coefficients - both the OLS and the IV - are somewhat smaller. This differences is, however, most likely due to sampling error. In column 8 I rerun (5) for the sample of counties, where the 1946 data is available. For these counties, the estimated effect of the 1950 refugee share is essentially the same as when using the 1946 data. Finally, in the last two columns of Table 6, I consider the same regression where I use the agricultural employment share (column 9) and the share of blue collar workers (column 10) 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.[8]

In Figure 7 I display the empirical relationship estimated in column 3 graphically. The figure shows a robust positive relationship, which is not driven by particular outliers. Recall that (3) highlights that importance of the parameter $\vartheta$ for the strength of this cross-county relationship. In fact, it is precisely this slope of the cross-county relationship between refugees and manufacturing employment which I will use to identify the endogenous response of manufacturing productivity in the quantitative model calibrated in Section 5.

**Refugees and the Entry of Manufacturing Plants** The simple model above, highlights the potential for an endogenous productivity response in the manufacturing sector captured by $\vartheta$. Standard industry equilibrium models (most notably Krugman (1980b)) pinpoint to a particular microfoundation for the existence of such agglomeration externalities: the entry of new producers. In fact, 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

[8]Note that this is not simply an artifact of these counties having larger manufacturing shares as column 10 explicitly controls for the overall share of manufacturing and male workers.

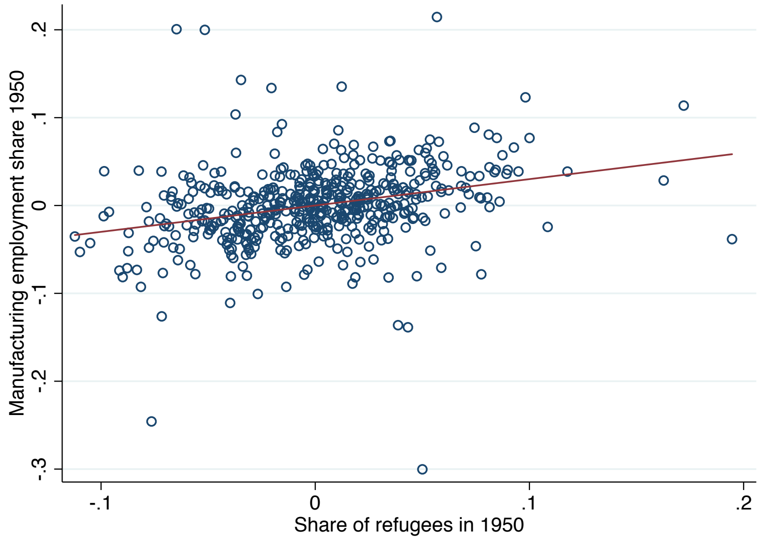
| | Manufacturing share | | | | | | | | Ag. share | Blue collar share |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1939 | | | | 1950 | | | | | 1950 |
| | | OLS | | | IV | OLS | IV | | OLS | |
| Share of refugees in 1950 | -0.117 | 0.300*** | 0.238*** | 0.228*** | 0.453** | | | 0.190*** | -0.233* | 0.264*** |
| | (0.071) | (0.055) | (0.054) | (0.050) | (0.198) | | | (0.070) | (0.128) | (0.045) |
| Population growth 1939-46 | | | | | | | | | | |
| Share of refugees in 1946 | | | | | | 0.165*** | 0.320* | | | |
| | | | | | | (0.061) | (0.173) | | | |
| ln pop dens. 1950 | | | 0.071*** | | | | | | | |
| | | | (0.017) | | | | | | | |
| ln pop dens. 1939 | -0.008** | 0.007* | -0.062*** | -0.016* | -0.013 | 0.007* | 0.008* | 0.006 | -0.100*** | 0.007** |
| | (0.004) | (0.003) | (0.017) | (0.009) | (0.009) | (0.004) | (0.004) | (0.004) | (0.007) | (0.003) |
| Share of housing stock damaged | -0.002 | 0.011 | 0.036** | 0.028* | 0.053* | 0.010 | 0.034 | 0.014 | 0.115*** | -0.022 |
| | (0.018) | (0.017) | (0.018) | (0.017) | (0.027) | (0.023) | (0.033) | (0.024) | (0.038) | (0.017) |
| Avg num of rooms (rooms/apt) | -0.028*** | -0.021*** | -0.022*** | -0.010* | -0.020** | -0.013* | -0.017** | -0.016** | 0.014 | -0.008 |
| | (0.008) | (0.006) | (0.006) | (0.006) | (0.009) | (0.007) | (0.007) | (0.007) | (0.013) | (0.005) |
| lnExpulsionDistance | -0.050* | 0.033* | 0.025 | -0.006 | 0.027 | 0.035 | 0.047** | 0.031 | -0.080* | 0.020 |
| | (0.026) | (0.019) | (0.019) | (0.021) | (0.033) | (0.022) | (0.021) | (0.021) | (0.043) | (0.018) |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Distance | No | No | No | Yes | Yes | No | No | No | No | No |
| Border FE | No | No | No | Yes | Yes | No | No | No | No | No |
| Manufacturing share 1939 | No | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 1933 Controls | Yes | No | No | Yes | Yes | No | No | No | No | No |
| Urbanization 1939 | No | No | No | Yes | Yes | No | No | No | No | No |
| Housing share post 1945 | No | No | No | Yes | Yes | No | No | No | No | No |
| Manufacturing share 1950 | No | No | No | No | No | No | No | No | No | Yes |
| Male share in 1950 | No | No | No | No | No | No | No | No | No | Yes |
| Observations | 509 | 510 | 510 | 509 | 509 | 362 | 362 | 362 | 510 | 510 |
| $R^2$ | 0.813 | 0.885 | 0.893 | 0.907 | 0.901 | 0.887 | 0.884 | 0.887 | 0.803 | 0.950 |

Standard errors in parentheses

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

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 (1)).

Table 6: 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 6, 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 7: Refugees and Manufacturing Employment in 1950

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

$$g_{N,r} = \delta_s + \beta \times \mu_r^{50} + \alpha \times \pi_{M,r}^{39} + \phi \times \ln \text{pd}_r^{39} + \varphi \times \text{wd}_r + \tau \times \text{exp\_dist}_r + \chi \times h_r + \zeta \times C_r + v_r, \qquad (6)$$

I will refer to this growth rate as the extent of entry. The results are contained in Table 7.

In the first column I show that - reassuringly - there is no relationship between the allocation of refugees and plant entry between 1933 and 1939. Running the same specification for the growth rate between 1939 and 1950 yields a sizable positive effect, whereby an increase in the share of refugees by 10 percentage points increases the number of manufacturing plants in 1950 by 10%. The next columns show that this effect is essentially unaffected by additional controls. In particular controlling for the number of manufacturing plants in 1933 and 1939 and the 1939 manufacturing share leaves the coefficient unchanged. Column 5 shows that I get the same results when I use the number of plants in 1956 instead of 1950 or the log share of refugees. The last two columns contain the IV estimates, which are very imprecisely estimated. In particular, the results are very sensitive as to wether I control for the number of plants in the pre-war period.

### Refugees and Manufacturing Employment in 1961

Finally, I can also use the 1961 census to measure the effects of the initial allocation of refugees on manufacturing employment in that time period. I consider the exact same specification as before but for brevity I only report a subset of the results. In the first two columns of Table 8 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 8 with the ones from 1950 in Table 6 it is seen that the coefficients are roughly of the same

| | Growth in manufacturing plants | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1933-39 | | 1939-50 | | 1939-56 | | 1939-50 | |
| | | | OLS | | | | | IV |
| Share of refugees in 1950 | 0.017 | 0.849*** | 0.804*** | 0.943*** | 1.109** | | -0.941 | 2.265*** |
| | (0.328) | (0.268) | (0.217) | (0.204) | (0.479) | | (0.889) | (0.847) |
| ln refugee share 1950 | | | | | | 0.160*** | | |
| | | | | | | (0.033) | | |
| ln pop dens. 1939 | 0.063*** | 0.009 | 0.037** | 0.201* | 0.223** | 0.210* | 0.008 | 0.215* |
| | (0.022) | (0.019) | (0.016) | (0.120) | (0.113) | (0.122) | (0.019) | (0.119) |
| Share of housing stock damaged | 0.167 | 0.155 | 0.199* | 0.170* | -0.339** | 0.228** | -0.061 | 0.323** |
| | (0.132) | (0.111) | (0.103) | (0.099) | (0.165) | (0.105) | (0.161) | (0.149) |
| Avg num of rooms (rooms/apt) | 0.095* | -0.101** | -0.065* | -0.085*** | -0.111 | -0.082*** | -0.035 | -0.142*** |
| | (0.049) | (0.042) | (0.033) | (0.031) | (0.068) | (0.030) | (0.052) | (0.051) |
| lnExpulsionDistance | -0.106 | 0.223* | 0.224** | -0.168 | -0.218 | -0.131 | 0.022 | 0.021 |
| | (0.146) | (0.129) | (0.111) | (0.113) | (0.251) | (0.115) | (0.182) | (0.182) |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Distance | No | No | No | Yes | Yes | Yes | No | Yes |
| Border FE | No | No | No | Yes | Yes | Yes | No | Yes |
| Manufac. plants 1933, 1939 | No | No | Yes | Yes | Yes | Yes | No | Yes |
| Manufacturing share 1939 | No | No | No | Yes | Yes | Yes | No | Yes |
| 1933 Controls | No | No | No | Yes | Yes | Yes | No | Yes |
| Urbanization 1939 | No | No | No | Yes | Yes | Yes | No | Yes |
| Housing share post 1945 | No | No | No | Yes | Yes | Yes | No | Yes |
| Observations | 491 | 495 | 491 | 490 | 490 | 490 | 495 | 490 |
| $R^2$ | 0.063 | 0.140 | 0.319 | 0.468 | 0.442 | 0.471 | 0.066 | 0.429 |

Standard errors in parentheses

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

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 (1)).

Table 7: Refugees and the Entry Manufacturing Plants

magnitude. In column 3 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 to move than the rest of the population and regions, which experience growth in their manufacturing sector, turn out 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 the data as the coefficient is almost twice as high as the OLS estimate using the 1950 refugee share reported in column 2. The last three columns contain the IV specifications. While the results using the 1950 refugee share are qualitatively similar to the OLS results, they are imprecisely estimated and hence not significantly differentially similar from zero. Interestingly, the IV estimate using the 1961 data is very similar to the 1950 IV estimate. This suggests that refugees' migration decisions between 1950 and 1961 are correlated with changes in the size of the manufacturing sector.

**Labor Market Opportunities: Refugees and Manufacturing Employment**

The above analysis showed a positive relationship between refugee-inflows, manufacturing employment and the entry of manufacturing plants at the regional level. According to the simple model above, this cross-sectional relationship is driven by a combination of decreasing returns to scale in agriculture and agglomeration externalities in the manufacturing sector. However, for simplicity the model abstracted from differences in *comparative* advantage between refugees and natives. In this section I will provide some direct micro-evidence on that refugees were indeed "biased" towards the manufacturing sector and plausibly acted as a local supply shifter. Such differences in human capital are important for the calibration of the structural model below.

Recall that refugees came from areas, which - historically - tended to specialize in agriculture (see Figure 2). Also recall that refugees were predominantly allocated to rural areas, which - in 1939 - had a low employment share in manufacturing.[9] Upon their arrival in Western Germany, however, refugees did *not* enter the agricultural sector but instead predominantly worked in the manufacturing sector. This reallocation pattern is clearly visible in a unique special supplement to the 1971 census, which contains micro-data on long-run employment histories. In particular, that census asked every respondent in 1971 where he/she lived in 1939 and in which occupation/sector cell he/she worked in 1939, 1950, 1960 and 1971. By analyzing this time-series of retrospective questions, I can therefore measure the life-cycle of employment patterns for both refugees and natives. Importantly, the data spans the time 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 9. 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 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

[9] Recall that the regression reported in Table 2 controls for population density ((in 1933). The unconditional correlation between the share of refugees and the manufacturing employment share is significantly negative.

| | Manufacturing share in 1961 | | | | | |
|---|---|---|---|---|---|---|
| | OLS | | | IV | | |
| Share of refugees in 1950 | 0.241*** | 0.182*** | | 0.183 | 0.064 | |
| | (0.061) | (0.058) | | (0.202) | (0.213) | |
| Share of refugees in 1961 | | | 0.451*** | | | 0.103 |
| | | | (0.060) | | | (0.259) |
| ln pop dens. 1939 | -0.013*** | -0.017** | -0.021** | -0.013*** | -0.018** | -0.022** |
| | (0.004) | (0.009) | (0.008) | (0.004) | (0.009) | (0.009) |
| Share of housing stock damaged | -0.011 | -0.017 | 0.015 | -0.018 | -0.031 | -0.018 |
| | (0.020) | (0.021) | (0.022) | (0.031) | (0.032) | (0.032) |
| Avg num of rooms (rooms/apt) | -0.007 | -0.002 | -0.014* | -0.005 | 0.003 | 0.002 |
| | (0.007) | (0.007) | (0.007) | (0.009) | (0.009) | (0.011) |
| lnExpulsionDistance | 0.027 | 0.044 | 0.047 | 0.021 | 0.027 | 0.018 |
| | (0.026) | (0.032) | (0.030) | (0.029) | (0.041) | (0.036) |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Distance | No | Yes | Yes | No | Yes | Yes |
| Border FE | No | Yes | Yes | No | Yes | Yes |
| Manufacturing share 1939 | Yes | Yes | Yes | Yes | Yes | Yes |
| 1933 Controls | No | Yes | Yes | No | Yes | Yes |
| Urbanization 1939 | No | Yes | Yes | No | Yes | Yes |
| Housing share post 1945 | No | Yes | Yes | No | Yes | Yes |
| Observations | 510 | 509 | 455 | 510 | 509 | 455 |
| $R^2$ | 0.784 | 0.804 | 0.840 | 0.783 | 0.802 | 0.827 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

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 (1)).

Table 8: 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 9: Sectoral and Occupational Mobility from 1939 to 1971

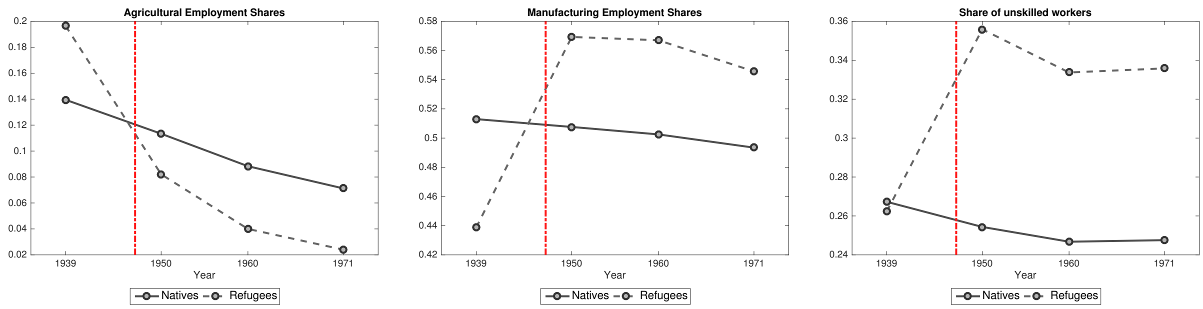
the expulsion, these individuals take unskilled jobs, which are mostly in the manufacturing sector. The remaining columns in Table 9 show that these reallocation patterns in 1950 are by no means transitory. In contrast, they persist all throughout the 1950s and 1960s.

While Table 9 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 9 is in fact representative of a typical life-cycle, consider Figure 8, 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. For them, the time period of the expulsion is hardly noticeable.

The patterns in Table 9 and (especially) Figure 8 are suggestive of distortions in the agricultural sector as they show that refugees left the agricultural sector *despite* living and working 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.[10] Hence, the demand for outside agricultural workers was quite limited. Even the Military Government

[10]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.




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 8: The Life-Cycle of the 1915-1919 Cohort

| 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 10: Occupational Distribution within Sectors

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 10, 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.

## 4.3 Refugees, Agglomeration and Local Economic Development

So far I have shown that the regional inflow of refugees was associated with an expansion of the local manufacturing sector. Moreover, the results in Table 7 show a positive correlation between the share of refugees and the entry of new plants. These patterns suggest that the settlement of refugees and their absorption in the manufacturing sector could have induced an increase in local manufacturing productivity. To focus more directly on such agglomeration economies, I now study the cross-sectional relationship between the share of refugees and 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 what follows I will for simplicity refer to this measure as "GDP per capita". My main empirical specification takes the now familiar form

$$\ln y_r^{1961} = \delta_s + \beta \times \mu_r^{50} + \alpha \times \pi_{M,r}^{39} + \phi \times \ln \mathrm{pd}_r^{39} + \varphi \times \mathrm{wd}_r + \tau \times \mathrm{exp\_dist}_r + \chi \times h_r + \zeta \times C_r + v_r, \qquad (7)$$

where $y_{cs}^{1961}$ denotes GDP per capita in region $r$ in 1961 and all other variables are defined as above. Unfortunately, I do not have information on measures of income per capita prior to the war. Hence, in contrast to the results for the manufacturing sector above, I cannot estimate equation 7 for the growth rate of GDP but only in levels. The results are contained in Table 11.

In the first column I report the simple cross-sectional relationship between GDP per capita in 1961 and the share of refugees. This relationship is strongly negative as refugees were allocated to less developed, rural areas farther to the East. This is directly seen in column 2, where I show that places with a higher population density in 1939 are richer 20 years later. Once this observable pre-war heterogeneity and the controls for the housing stock are controlled for, the share of refugees shows a positive correlation with GDP per capita. Columns 4 and 5 show that this effect of refugees on local economic development is robust to controlling for richer controls in the pre-war sectoral structure, the population density in 1933, a regions geographical location and the region's pre-war level of urbanization. Quantitatively, the results in Table 11 suggest a sizable effect of refugee inflows on income per capita: an increase in the share of refugees by 10 percentage points in 1950 increases income per capita by roughly 6% ten years later.

As in Table 8, I also relate income per capita to the contemporaneous share of refugees. Expectedly, the point estimate increases, which presumably reflects the higher extent of directed spatial mobility of refugees in the 1950s. Finally, columns 8 and 9 contain the IV specification. While the coefficients is still positive and significant, the standard errors also increase substantially, rendering the economic magnitude of coefficient had to interpret.

In order to interpret these results as driven by changes in local productivity (e.g. driven by the entry of new plants), the simple model, in particular (4), highlights the importance of relative skill supplies. If wages reflect marginal products, innate differences in skills will be reflected in relative earnings. In terms of the simple model above, relative earning between refugees and natives are exactly given by $\frac{e_R}{e_N} = \frac{wQ}{w} = Q$. According to the theory, the positive effect of refugees on income per capita requires some forms of agglomeration economies if $\lambda < 1$. In Table 12 I show that this seems to indeed be case. In particular, I use the EVS microdata to run regressions of the form

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

where $\overline{e}_i$ denotes earning of individual $i$, $Refugee$ indicates the refugee status, $x_i$ controls for demographic characteristics and $\delta_s, \delta_{Ind}, \delta_{City}$ and $\delta_{Occ}$ are state, industry, city size and occupation fixed effects. The results in Table 12 show that 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. Taken together, Tables 11 and 12 suggest a positive relationship between refugee inflows and local productivity. In the next section I will calibrate a simple general equilibrium model of trade to quantify the strength of such relationship.

| | OLS | | | | | | IV | |
|---|---|---|---|---|---|---|---|---|
| Share of refugees in 1950 | -0.705*** | | 0.671*** | 0.640*** | 0.625*** | | 1.702** | 1.453** |
| | (0.169) | | (0.196) | (0.180) | (0.186) | | (0.748) | (0.696) |
| Share of refugees in 1961 | | | | | | 0.871*** | | |
| | | | | | | (0.250) | | |
| lnExpulsionDistance | | 0.288*** | 0.389*** | 0.478*** | 0.433*** | 0.420*** | 0.508*** | 0.568*** |
| | | (0.080) | (0.086) | (0.076) | (0.114) | (0.116) | (0.121) | (0.153) |
| ln pop dens. 1939 | | 0.093*** | 0.088*** | 0.051*** | 0.061** | 0.060** | 0.088*** | 0.070*** |
| | | (0.007) | (0.012) | (0.012) | (0.024) | (0.026) | (0.012) | (0.025) |
| Share of housing stock damaged | | | 0.014 | 0.068 | 0.041 | 0.032 | 0.141 | 0.137 |
| | | | (0.067) | (0.062) | (0.066) | (0.069) | (0.114) | (0.110) |
| Avg num of rooms (rooms/apt) | | | -0.061** | 0.020 | 0.031 | 0.028 | -0.098*** | -0.002 |
| | | | (0.024) | (0.024) | (0.025) | (0.027) | (0.035) | (0.034) |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Distance | No | No | No | No | Yes | Yes | No | Yes |
| Border FE | No | No | No | No | Yes | Yes | No | Yes |
| Manufacturing share 1939 | No | No | No | Yes | Yes | Yes | No | Yes |
| 1933 Controls | No | No | No | Yes | Yes | Yes | No | Yes |
| Urbanization 1939 | No | No | No | No | Yes | Yes | No | Yes |
| Housing share post 1945 | No | No | No | No | Yes | Yes | No | Yes |
| Observations | 499 | 499 | 499 | 497 | 497 | 444 | 499 | 497 |
| $R^2$ | 0.237 | 0.465 | 0.485 | 0.583 | 0.606 | 0.614 | 0.448 | 0.583 |

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

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 (1)).

Table 11: Refugees and GDP per capita

| | 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 12: Refugees vs Natives: Earnings in 1962

# 5 Quantitative Framework

So far I have shown direct empirical evidence of a positive relationship between refugee inflows and manufacturing employment, the entry of manufacturing plants and local income per capita. Through the lens of the simple theoretical framework introduced above, these empirical patterns are suggestive of agglomeration forces whereby manufacturing productivity depends positively on the size of the local economy. However, the model is too stylized to use it to credibly quantify the strength of agglomeration forces. In this section I will therefore calibrate a fully specified canonical general equilibrium model of inter-regional trade that is rich enough to be taken to the data. Compared to the simple setup introduced above, the model has (i) many heterogeneous regions, (ii) determines the sectoral prices $p_A$ and $p_M$ in general equilibrium, (iii) allows for a less than perfectly elastic labor supply across sectors and (iv) explicitly takes the heterogeneity in refugees' and natives' skills into account.

I will calibrate the structural parameters of the model to match the cross-sectional relationships between manufacturing employment, the share of refugees, population density, war-damages and counties' distance to the expulsion regions. Hence, the model and calibration strategy explicitly recognizes the correlation between the observed refugee flows and the underlying unobserved fundamentals like productivity or land supply. The key empirical moment to discipline the endogenous relationship between productivity and market size is the cross-county relationship between refugee-inflows and manufacturing employment shown in Figure 7.

## 5.1 The Model

I will use the model mainly as a measurement-device. I therefore do not innovate on the model-setup but borrow most ingredients from the existing literature in spatial economics.[11]

[11] See Redding and Rossi-Hansberg (2016) for a recent survey.

**Preferences and Technology** I consider a multi-region economy, where each region $r = 1, .., R$ corresponds to a county of the empirical analysis above. 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 different local varieties stemming from the $R$ regions. Consumers also value housing services, which they acquire locally. Specifically, consumers' utility is given by

$$u = u(C_A, C_M, h) = \left(C_A^{\alpha} C_M^{1-\alpha}\right)^{\zeta} h^{1-\zeta} \text{ with } C_s = \left(\sum_r c_{s,r}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{8}$$

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

Production technologies are as before, i.e. the agricultural good is produced using labor and land and the manufacturing good is subject to increasing returns through local agglomeration forces

$$Y_{rt,A} = Z_{rt,A} \times H_{rt,A}^{1-\gamma} T_r^{\gamma}, \tag{9}$$

$$Y_{rt,M} = \tilde{Z}_{rt,M} \times H_{rt,M} \text{ where } \tilde{Z}_{rt,M} = Z_{rt,M} \times H_{rt,M}^{\vartheta}. \tag{10}$$

Again, I denote by $T_r$ the amount of land in county $r$, $Z_{rt,s}$ denotes the exogenous sectoral productivity in county $r$ and $H_{rt,s}$ is total amount of labor (again measured in efficiency units) used in the production of sector $s$ goods. Note that (in contrast to above) I assume that manufacturing productivity depends on the equilibrium amount of human capital employed in the manufacturing sector. This is mostly for aesthetic reasons because (10) can be explicitly derived from the basic Krugman (1980b) model - see Appendix. In that model, the source of agglomeration is the existence of demand externalities combined with free entry. Hence, manufacturing productivity is endogenous because an increase in supply triggers the entry of new plants. This is in line with the empirical results presented above.

**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. 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 $\theta$ and location $(Q_A^g, Q_M^g)$. I normalize the efficiency of natives in each sector to unity, i.e. $Q_A^N = Q_M^R = 1$. As is well known, this structure is extremely tractable, because it implies convenient closed-form expressions for the aggregate supply of sector human-capital $H_r^S$, average (group-specific) earnings $\overline{w}_r^g$ and (group-specific) sectoral employment shares $\pi_{r,s}^g$. These are given by

$$H_{rt,s} = L_{rt}\nu \times \left[\mu_{rt}\left(\pi_{rt,s}^N\right)^{\frac{\theta-1}{\theta}} + (1-\mu_r)\left(\pi_{rt,s}^R\right)^{\frac{\theta-1}{\theta}}\left(Q_s^R\right)^{\frac{1}{\theta}}\right], \tag{11}$$

$$\overline{w}_{rt}^g = \nu \times \left(Q_A^g w_{rt,A}^{\theta} + Q_M^g w_{rt,M}^{\theta}\right)^{1/\theta} \tag{12}$$

$$\pi_{rt,M}^g = \frac{\frac{Q_M^g}{Q_A}\left(\frac{w_{rt,M}}{w_{rt,A}}\right)^{\theta}}{1 + \frac{Q_M^g}{Q_A}\left(\frac{w_{rt,M}}{w_{rt,A}}\right)^{\theta}} = 1 - \pi_{rt,A}^g, \tag{13}$$

where $\nu = \Gamma\left(1 - \frac{1}{\theta}\right)$ is a constant related to the gamma function $\Gamma$. Hence, compared to the simple model analyzed above, this setup contains two difference. First of all, each sector faces an upward sloping supply curve. Secondly, refugees and natives now differ not only in absolute advantage but also in comparative advantage. This structure nests the simplified structure from Section 4 for the case of $Q_A = Q_M = Q$ and $\theta \to \infty$.

Above I showed that - empirically - refugees have higher employment shares in the manufacturing sector and lower earnings.[12] These two observations directly imply that $\frac{Q_M^R}{Q_A^R} > 1$ and $Q_A^R < 1$. I will calibrate both of these skill parameters to match the microdata on relative earnings and relative employment shares. While the theory is agnostic about the source of productivity differences between refugees and natives, I argued above that refugees' were plausibly subject to frictions in the land market, which prevented them from entering the agricultural sector. In the Appendix I present an explicit microfoundation with frictional land markets, which is isomorphic to the model above. In that microfoundation, refugees' level of skills $Q_s^R$ explicitly depend on a collateral requirements to buy farm land.[13]

**Equilibrium** I will analyze the data as if stemming from the equilibrium of this model. I will utilize two notions of equilibria. I will refer to a *trade equilibrium* as an equilibrium allocation for a given distribution of people across space. In contrast, I will call an allocation a *spatial equilibrium* if individuals are also mobile across space and the allocation of people is consistent with individuals' mobility decisions. Given the structure above, it is easy to verify that a trade equilibrium is fully characterized by the conditions

$$w_{M,r} H_{M,r} = H_{M,r}^{(\sigma-1)\vartheta} Z_{A,r}^{\sigma-1} w_{M,r}^{1-\sigma} P_M^{\sigma-1} \times (1-\alpha)\,\zeta \left(\sum_r Y_r\right) \tag{14}$$

$$\frac{1}{1-\gamma} w_{A,r} H_{A,r} = Z_{A,r}^{\sigma-1} \left(\frac{1}{1-\gamma}\left(\frac{H_{A,r}}{T_r}\right)^{\gamma} w_{A,r}\right)^{1-\sigma} \times \alpha\zeta \left(\sum_r Y_r\right) \tag{15}$$

$$Y_r = \frac{1}{\zeta}\left(w_{M,r} H_{M,r} + \frac{1}{1-\gamma} w_{A,r} H_{A,r}\right), \tag{16}$$

where $Y_r$ denotes region $r$'s total income in terms of the agricultural good (which I take to be the numeraire) and $P_M$ is the usual CES price index associated with (8). Importantly, $H_{A,r}$ and $H_{M,r}$ are the sectoral human capital supply functions, which are given in (11) and directly depend on the number of refugees allocated to a particular region.

While a trade equilibrium takes the distribution of people as given and determines regional factor prices $[w_{rt,M}, w_{rt,A}]$ and human capital supplies $[H_{rt,M}, H_{rt,M}]$ as the solution to (14)-(16), a spatial equilibrium treats individuals' location choices as endogenous. Given the specification of utility in (8), the expected utility of an individual of group $g$ locating in region $r$ prior to the realization of the skill vector $e^i$, is given by $\overline{u}_{rt}^g = \left(P_A^{\alpha\zeta} P_M^{(1-\alpha)\zeta}\right)^{-1} \times \frac{\overline{w}_{rt}^g}{b_{rt}^{1-\zeta}}$,where average earnings $\overline{w}_{rt}$ are given in (12) and $b_{rt}$ denotes the equilibrium rental rate for residential housing. Letting $B_r$ the supply of housing services, the Cobb-Douglas demand structure implies that $b_{rt} = (1-\zeta)\,Y_r$. Hence, $\overline{u}_{rt}^g$ is fully determined in any trade equilibrium.

To generate a well-behaved spatial labor supply function, I follow Redding (2016) and assume that consumers have idiosyncratic preferences for particular locations. In particular, suppose that the utility of residing in region $r$ is given by $V_r = \overline{u}_{rt}^g \times \delta_r^i$. Assuming that $\delta_r^i$ is iid across individuals and regions and drawn from a Frechet distribution with parameter $\varepsilon$, it follows directly that the share of people from group $g$ preferring region $r$over all other regions $j$ is given by $\rho_{rt}^g = \frac{\left(\overline{w}_{rt}^g / b_{rt}^{1-\varsigma}\right)^{\varepsilon}}{\sum_j \left(\overline{w}_{jt}^g / b_{jt}^{1-\varsigma}\right)^{\varepsilon}}$. A tupel $[w_{rt,s}, H_{rt,s}, L_{rt}^g]$, which satisfies the above conditions for a

[12] Note that average earnings are equalized across both sectors $s$ and depend on the level of skills. 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.

[13] See also Hsieh et al. (2013), who consider a model with exogenous frictions for some individuals to enter particular occupations.

trade equilibrium and the restriction $\frac{L^g_{rt}}{\sum L^g_{jt}} = \rho^g_{rt}$ constitutes a spatial equilibrium in this economy.

## 5.2 Empirical Strategy

To analyze the empirical results though the lens of this model, I adopt the following strategy. I focus on two cross-sections - 1939 and 1950. I assume that the allocation of people across space and across sectors of production within counties in 1939 is consistent with a spatial equilibrium. As no refugees have yet arrived in Western Germany, a region is fully characterized by the tuple

$$[Z_{1939,r,A}, Z_{1939,r,M}, T_r, B_r]_r .$$

The amount of land $T_r$ is directly observable. Similarly, I treat the supply of housing space $B_r$ as exogenous and measure it simply by the number of rooms per capita from the housing census. In the data I observe the distribution of people across space, $\frac{L_{r,1939}}{\sum_j L_{j,1939}}$, and the sectoral employment shares $\pi_{r1939,M}$. It is easy to verify that - given the structural parameters - the model above generates a unique mapping of these two data objects to the distribution of relative productivity $\left[\frac{Z_{1939,r,A}}{\sum_j Z_{1939,r,A}}, \frac{Z_{1939,r,M}}{\sum_j Z_{1939,j,M}}\right]_r$ and the sectoral expenditure share $\alpha$.[14] Hence, from the observed allocation in 1939, I can fully account for the distribution of sectoral productivity in 1939. Note that the observed data on regional land size ($T_r$) and housing supply ($B_r$) is solely a way to account for parts of the observed allocation with determinants other than regional "residual" productivity.

*Given* this regional productivity distribution I then turn to the cross-sectional data in 1950. I assume that the cross-sectional allocation in 1950 is characterized by two features:

1. I assume that the war has a direct effect on local manufacturing productivity. In particular, I assume that regional manufacturing productivity in 1950, $Z_{1950,r,M}$ is given by

$$\ln Z_{1950,r,M} = \ln Z_{1939,r,M} + \phi \times \text{exp\_dist}_r - \kappa \times ln\left(d_r\right), \tag{17}$$

   where Distance$_r$ is the regional distance to the expulsion regions in the East and $d_r$ is extent of war-related destruction. Note that the allocations in the model only depends on relative productivities. Hence, $\phi > 0$ and $\kappa > 0$ imply that regions in the East and regions with ample war-related damages see their manufacturing productivity decline *relative* to other regions.

   Allowing for a direct effect of a region's distance to Eastern Europe on its productivity after the war is a reduced-form way to capture the effect of changes in market access stressed by Redding and Sturm (2008) in my model, which - for simplicity - does not have any trade costs.[15] Because the allocation of refugees is correlated with the distance to the expulsion regions, the specification in (17) captures an important aspect of the endogeneity of the refugee allocation: to the extent that $\phi > 0$, the allocation of refugees will be negatively correlated with changes in manufacturing productivity between 1939 and 1950. Similarly, I allow for a *direct* effect of the Allied bombing campaign on manufacturing productivity. I do so, to explicitly account for an additional source of correlation between refugee flows and manufacturing productivity. In contrast to the case

[14] The reason why, given the data, there is a unique $\alpha$, which is consistent with the equilibrium allocation is the Cobb-Douglas assumption. As sectoral expenditure shares are fully determined from $\alpha$, the level of productivity is not identified from the data as the relative prices $P_M$ does not affect sectoral spending. In contrast, there is a unique $\alpha$, which is consistent with observed sectoral employment shares.

[15] Redding and Sturm (2008) exploit the German division after the Second World War as providing exogenous variation in market access. As locations close to the inner German border saw a larger decline in their market access, theory predicts that such regions should be effected more from the division. Redding and Sturm (2008) indeed find that cities closer to the border grew relatively less in the post-war period.

of geography, the negative correlation between the allocation of refugees and the spatial distribution of war damages implies a positive correlation between refugees and changes in manufacturing productivity.

2. In contrast to the 1939 cross-section, I posit that the data is generated by a trade equilibrium but *not* by a spatial equilibrium. In particular, I take the distribution of people - both natives and refugees - as exogenous directly from the data. Because the spatial distribution of people is highly persistent and productivity in 1939 is - by construction - correlated with population density in 1939, the model will generate a positive correlation between 1950 population density and relative productivity. However, this strategy does assume that there is no relationship between population changes and productivity changes in addition to the one captured by (17).

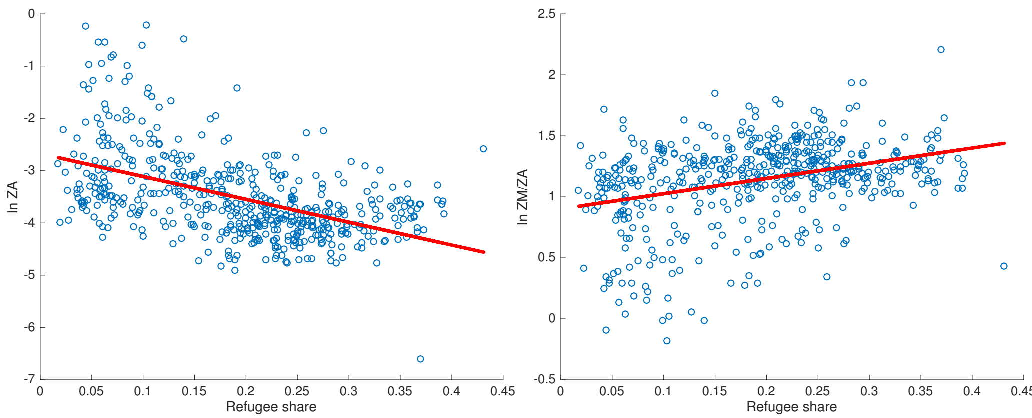
Given this structure, my strategy is as follows. The model is parameterized by 11 structural parameters. I will fix five parameters, the trade elasticity $\sigma$, the sectoral and spatial labor supply elasticities $\theta$ and $\varepsilon$, the extent of decreasing returns $\gamma$ and the expenditure share on housing $\zeta$ to values from the literature.[16] The remaining six parameters are calibrated within the model. The agricultural share $\alpha$ is directly calibrated to the match the average level of manufacturing employment in 1939 (recall the discussion in footnote 14 above). The levels of refugees' sectoral human capital $Q_A^R$ and $Q_M^R$ are calibrated to match the estimated earnings-differential of 9% and the 10% higher manufacturing employment share within regions. Finally, I calibrate the effect of war-time destruction ($\kappa$), changes in market access ($\phi$) and extent of agglomeration ($\vartheta$) through indirect inference. Specifically, I require the model to replicate the within-state cross-county elasticities of manufacturing employment with respect to the allocation of refugees, the extent of war-time destruction and distance to the expulsion regions holding the 1939 population density and the 1939 manufacturing employment share fixed. This is similar in spirit to Faber and Gaubert (2016) and Ahlfeldt et al. (2015), who also calibrate their structural parameters by requiring the model to match regression evidence. Formally, I consider the regression

$$\pi_{1950,r,M} = \delta_s + \beta \times \mu_r + \xi \times d_r + \varphi \times \text{Distance}_r + \psi_{pd} \times \text{pop dens}_r^{39} + \psi_{\pi_M} \times \pi_{1939,r,M} + u_r \tag{18}$$

for the both the model-generated manufacturing shares $\left(\pi_{1950,,r,M}^{Model}\right)$ and the observed allocations and calibrate the structural parameters $(\kappa, \phi, \vartheta)$ to ensure the coefficients $(\beta, \xi, \varphi)$ to be the same. Note that - in the model - these parameters exactly correspond to the direct effects of refugee flows, war-time destruction and changes in market access on manufacturing employment, as controlling for 1939 population density and the 1939 manufacturing share is equivalent to controlling for the initial distribution of productivity $(Z_{39,A,r}, Z_{39,M,r})$.

**Discussion and Limitations** Before turning to the results, let me briefly discuss three limitations of my approach. The first concerns the spatial distribution of human capital. The current structure assumes that the distribution of human capital does not vary across space. This is potentially important for the implied productivity distribution. In particular, the model infers the extent of comparative advantage $Z_{M,r}/Z_{A,r}$ from the data on manufacturing employment shares holding the number of people and the amount of land fixed. If skills were to vary across space and some regions happened to be inhabited by individuals whose human capital is biased towards the manufacturing sector, the model would require less differences in innate relative productivity to account for the data. Given the data on employment shares and population in 1939, these two margins are not separately identified. Instead, one would need data on relative wages across space to distinguish human capital from spatial productivity to account for the data on quantitates. To the best of my knowledge data on relative wages across counties of the German Reich in 1939 is not available.

[16] Note that $\varepsilon$ and $\zeta$ are not required to calculate the 1950 trade equilibrium. However, given the 1939 data, $\varepsilon$and $\zeta$ are required to solve for the implied productivity distribution.




Notes: The figure shows the relationship between the allocation of refugees in 1950 and agricultural productivity $\left(\ln Z_{39,A,r}\right)$ (left panel) and comparative advantage in manufacturing $\left(\ln \frac{Z_{39,M,r}}{Z_{39,A,r}}\right)$ (right panel) in 1939. The distribution of regional productivity in 1939, $\left[Z_{39,Ar}, Z_{39,M,r}\right]_r$, is calibrated so that the model exactly replicates the joint distribution of population density and manufacturing employment shares in 1939 as a spatial equilibrium.

Figure 9: The Allocation of Refugees and Productivity in 1939

The second limitation concerns the absence of regional characteristics other then productivity, the supply of housing and the amount of land to determine the spatial allocation of people. If, for example, regions were to differ in amenities, the model would require less of a dispersion in productivity to account for the dispersion in population density in 1939. Again, without data on e.g. factor prices, one cannot tell these margins apart and hence I abstracted from them amenities for now. Finally, note that - for now - I only use data from 1939 and 1950 to estimate the model's parameters. In particular, I do not exploit the information on GDP per capita in 1961. The chief reason is refugees' mobility response. As I discussed above, the 1961 refugee share correlates strongly with both manufacturing employment and GDP per capita in 1961. This suggests that refugees were in fact quite mobile across space during the 1950s. Interpreting the data in 1961 through the lens of the model therefore requires a richer model of population mobility.

### 5.3 Results

Consider first Figure 9, where I depict the correlation between the share of refugees and agricultural productivity ($Z_{1939,r,A}$) in the left panel and the extent of comparative advantage in manufacturing ($Z_{1939,r,M}/Z_{1939,r,A}$) in the right panel. Through the lens of the calibrated model, refugees are allocated to regions, which are less productive (reflecting the negative correlation with population density) and have a slight comparative advantage in the manufacturing sector. The model also highlights two features of the empirical strategy exploited above: While the allocation of refugees is not orthogonal to the underlying productivity distribution, there is ample cross-sectional variation, which can be exploited.

In Table 13 I report the calibrated parameters. The calibrated model implies a sizable agglomeration force - the elasticity of productivity to manufacturing employment is given by $\vartheta = 0.21$. To put this into perspective, note that in the canonical Krugman (1980b) economy, we have that $\vartheta = \frac{1}{\rho-1}$, where $\rho$ is the elasticity of substitution *across* differentiated varieties in the manufacturing sector. Hence, Table 13 implies that $\rho \approx 6$. Compared to other estimates reported in the literature, my estimate is slightly larger. The analysis in Bryan and Morten (2015) and Faber and Gaubert (2016) for example suggests an elasticity of about 8%-10%. The model also implies that war-related destruction reduces manufacturing productivity and that manufacturing productivity declines more in

| | Parameter | Value | Target | Moment | |
|---|---|---|---|---|---|
| | | | | Data | Model |
| $\vartheta$ | Agglomeration | 0.21 | Refugee-Manufacturing-relationship (see (5)) | 0.262 | 0.261 |
| $\kappa$ | $\frac{\partial ln Z_{1950,M}}{\partial ln(d_r)}$ (see (17)) | 0.047 | Destruction-Manufacturing-relationship (see (5)) | 0.01 | 0.01 |
| $\phi$ | $\frac{\partial ln Z_{1950,M}}{\partial \text{Distance}}$ (see (17)) | 0.012 | Distance-Manufacturing-relationship (see (5)) | 0.021 | 0.019 |
| $Q_A^R$ | Refugees' skills in Ag. | 0.61 | Earnings differential (see Table 12) | 9% | 9.2% |
| $Q_M^R$ | Refugees' skills in Man. | 0.96 | Difference in manufac. employment | 10% | 10% |
| $\alpha$ | Agricultural share | 0.71 | Sectoral employment in 1939 | - | - |
| $\theta$ | Sectoral elasticity of labor supply | 4 | set exogenously | | |
| $\gamma$ | Land share in ag. production | 0.3 | set exogenously | | |
| $\zeta$ | Consumption share in utility | 0.75 | set exogenously | | |
| $\sigma$ | Demand elasticity | 5 | set exogenously | | |
| $\varepsilon$ | Spatial elasticity of labor supply | 3 | set exogenously | | |

Notes: The table contains the calibrated parameters, the main calibration target (even though the parameters are naturally calibrated jointly) and the implied moments, both in the data and in the model. The parameters in the lower panel are not calibrated but set exogenously.

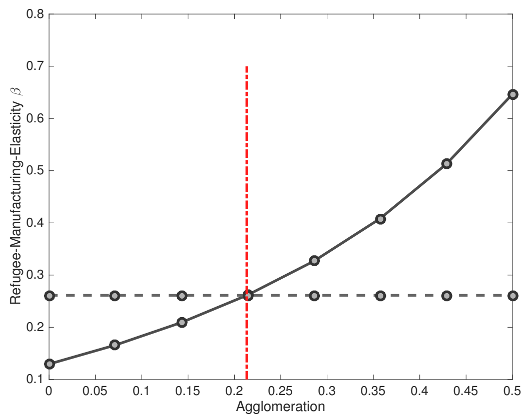
Table 13: Calibrated Parameters

regions close to the Border. Finally, the micro-data on refugees' lower relative earnings and their higher manufacturing employment shares implies that they are relatively less productive in the agricultural sector $\left(Q_A^R < 1\right)$, while almost equally productive in the manufacturing sector $\left(Q_M^R \approx 1\right)$.

The main source of identification for the extent of agglomeration economies $\vartheta$ is the cross-sectional relationship between refugees and manufacturing employment. In particular, the higher $\vartheta$, the more should an increase in relative manufacturing supply (at given relative wages) translate into equilibrium manufacturing employment as the presence of agglomeration forces prevent relative wages from declining. Another way to see this: differences in human capital determine differences in sectoral employment patters within regions, as refugees and natives face the same factor prices. Holding refugees' comparative advantage fixed, stronger agglomeration forces will make it easier for refugee-rich locations to expand their manufacturing sector. This is seen in Figure 10, where I report the implied cross-sectional semi-elasticity between the share of refugees and manufacturing employment, i.e. the coefficient $\beta$ in (18), for various levels of $\vartheta$. Importantly, for each choice of $\vartheta$, the model is re-calibrated to always match all the other moments reported in Table 13. As seen in Figure 10, a model without agglomeration forces is not able to quantitatively replicate the observed relationship between refugee flows and manufacturing employment. Note, however, that the relationship between refugees and manufacturing would still be positive. This is due to the fact that refugees do have a comparative advantage in the manufacturing sector and that refugee-areas have - by construction - more people and the agricultural sector is subject to decreasing returns.

## 6 Conclusion

In this paper I used a particular historical setting to empirically estimate wether local technology responds to changes in local factor supplies. I focused on the expulsion of the ethnic German population from their domiciles in Central and Eastern Europe and their subsequent resettlement in Western Germany. In the three years after the Second World War, roughly 8m people arrived in Western Germany. At the time, this amounted to an increase in population by about 20%. Furthermore, counties in Western Germany differed substantially in the extent to which




Notes: The figure plots the model-implied relationship between the extent of agglomeration $\vartheta$ and the partial correlation between the share of refugees $\mu_r$ and the manufacturing employment share $\pi_r^M$, i.e. the regression coefficient $\beta$ in (18). For each value of $\vartheta$ all other parameters of the model, including the distribution of productivity, are re-calibrated to perfectly match the moments reported in Table 13.

Figure 10: Agglomeration forces and the cross-sectional relationship between refugees and manufacturing

they participated in the incoming refugee flows. Using both the policies of the US and UK Military Government in Post-War Germany and the pre-war geographic distance from counties in Western Germany to the expulsion regions to isolate the exogenous component in refugee flows, I study the long-run economic consequences of such labor supply shocks across 500 counties in Western Germany.

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

To measure the strength of such agglomeration effects, I then calibrate a general equilibrium trade model to the data generated by this historical experiment. In particular, I discipline the parameters of the model through indirect inference, where I force the model to replicate the regression results from the historical experiment. The calibrated model implies an elasticity of manufacturing productivity with respect to sectoral labor supply of about 0.2.

# 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.

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

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

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

**Bryan, Gharad and Melanie Morten**, "Economic development and the spatial allocation of labor: Evidence from Indonesia," *Manuscript, London School of Economics and Stanford University*, 2015.

**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.

**Faber, Benjamin and Cecile Gaubert**, "Tourism and Economic Development: Evidence from Mexico's Coastline," Technical Report, National Bureau of Economic Research 2016.

**Fajgelbaum, Pablo and 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.

**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.

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

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

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

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

**Nellner, Werner**, "Grundlagen und Hauptergebnisse der Statistik," in Eugen Lember and Friedrich Edding, eds., *Die Vertriebenen in Westdeutschland: ihre Eінglierung 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.

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

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

**Redding, Stephen J**, "Goods trade, factor mobility and welfare," *Journal of International Economics*, 2016, *101*, 148–167.

_ **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**, "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

In Table 14 below I report the size of the ethnic German population for different countries in Eastern and Middle Europe in 1939. Naturally, there is a large German population in areas, which used to be part of the German Reich (see left column of Table 14). In addition, there were large population German population centers in other European countries, in particular the Sudetenland in Czechoslovakia.

| 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 14: The German Population in Central and Eastern Europe in 1939

## 7.2 The Simple Model

Consider the simple framework in Section 4. Equilibrium requires that the marginal product is equalized across sectors, i.e.

$$w = p_M Z_M = p_A (1-\gamma) H_A^{-\gamma} T^{\gamma},$$

Using (10), the equilibrium manufacturing share is determined by

$$p_M H^{\vartheta} = p_A (1-\gamma) H^{-\gamma} (1-\pi_M)^{-\gamma} T^{\gamma}.$$

Hence

$$ln(1-\pi_M) = -\frac{\gamma+\vartheta}{\gamma} ln(H) + ln\left(\frac{p_A}{p_M}(1-\gamma) T^{\gamma}\right),$$

so that

$$-\frac{d\pi_M}{1-\pi_M} = -\frac{\gamma+\vartheta}{\gamma} dln(H) = -\frac{\gamma+\vartheta}{\gamma}\frac{\lambda}{H} dL^R.$$

Rearranging terms yields (3). Income per capita is given by

$$y = \frac{wH_M + wH_A + RT}{L^R + L^N} = \frac{wH_M + \frac{1}{1-\gamma} wH_A}{H}\frac{H}{L^R+L^N} = w\left(\pi^M + \frac{1}{1-\gamma}\left(1-\pi^M\right)\right)\frac{H}{L}.$$

Hence

$$ln(y) = ln(w) + ln\left(\frac{1-\gamma\pi^M}{1-\gamma}\right) + ln\left(\frac{H}{L}\right).$$

Note that $dln(w) = \vartheta \frac{\lambda}{H} dL^R$ and

$$dln(H/L) = \frac{L}{H}\frac{\lambda L - H}{L^2} = \frac{1}{H}\frac{(\lambda-1)L^N}{L} = \frac{\lambda}{H}\frac{\lambda-1}{\lambda}(1-\mu),$$

we get that

$$\frac{dln\left(y\right)}{dL^{R}} \quad = \quad \vartheta\frac{\lambda}{H}-\frac{\gamma}{1-\gamma\pi^{M}}\frac{d\pi^{M}}{dL^{R}}+\frac{\lambda}{H}\frac{\lambda-1}{\lambda}\left(1-\mu\right).$$

Substituting for $\frac{d\pi^M}{dL^R}$ yields (4).

### 7.3 Characterization of the quantitative model in Section 5

In this section we characterize the equilibrium of the model laid out in section 5.

**The manufacturing sector: A Microfoundation** Consider first the manufacturing sector in region $r$. Let $w_{M,r}$ be the manufacturing wage in region $r$. As in Krugman (1980b), suppose that the local manufacturing good $Y_{r,M}$ is a composite of differentiated products of the manufacturing firms active in region $r$, which compete monopolistically. Specifically, assume that

$$Y_{r,M}=N_{r}^{\vartheta-\frac{1}{\rho-1}}\times\left(\int_{j=0}^{N_{r}}m_{j,r}^{\frac{\rho-1}{\rho}}dj\right)^{\frac{\rho}{\rho-1}}, \tag{19}$$

where $N_r$ denotes the number of active manufacturing firms, $m_{j,r}$ is firm $j$'s amount of manufacturing products and $\rho$ is the elasticity of substitution across firms' outputs. Importantly, $\vartheta$ parametrizes the extent of aggregate increasing returns - in a symmetric allocation where $m_{j,r} = M/N_r$, (19) 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 $\vartheta = \frac{1}{\rho-1}$. Because firms set a constant markup $\frac{\rho}{\rho-1}$ over their marginal costs, 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

$$m_{j,r}=m_{r}=\left(\rho-1\right)Z_{r}\zeta. \tag{20}$$

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

$$H_{r}^{M}=\int_{j=1}^{N_{j}}l_{j,r}+\zeta N_{j}=\rho\zeta N_{j}, \tag{21}$$

the number of varieties in equilibrium is proportional to the number of workers in the manufacturing sector. Furthermore, (19) and (20) imply that

$$Y_{M,r}=N_{j,r}^{\vartheta-\frac{1}{\rho-1}}\times m_{r}N^{\frac{\rho}{\rho-1}}=N_{j,r}^{\vartheta+1}\left(\rho-1\right)Z_{r}\zeta=\overline{\zeta}Z_{r}\left(H_{r}^{M}\right)^{\vartheta+1}, \tag{22}$$

where

$$\overline{\zeta}=\frac{\left(\rho-1\right)\zeta}{\left(\zeta\rho\right)^{\vartheta+1}}. \tag{23}$$

Hence, the degree of increasing returns is determined by $\vartheta$. (22) is the same equation as (10).

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

$$P_{M,r} \equiv \frac{X}{Y} = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}N_r^{-\vartheta} = \frac{\rho}{\rho-1}\frac{w_{M,r}}{Z_r}\left(\frac{1}{\rho\zeta}H_r^M\right)^{-\vartheta} = \frac{1}{\bar{\zeta}}\frac{w_{M,r}}{Z_r}\left(H_r^M\right)^{-\vartheta}. \tag{24}$$

**The agricultural sector** Given the production function in (9), the price of the agricultural variety in region $r$ is given by

$$p_{A,r} = \left(\frac{w_{A,r}}{\gamma}\right)^{\gamma}\left(\frac{R_r}{1-\gamma}\right)^{1-\gamma}, \tag{25}$$

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.

$$H_{A,r} = \frac{R_r}{w_{A,r}}\frac{\gamma}{1-\gamma}T_r. \tag{26}$$

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

$$w_{M,r}e_M^i \geq w_{A,r}e_A^i.$$

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

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

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

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

Defining

$$\delta \equiv \frac{Q^R_{A,r}\left(w_{A,r}\right)^{\theta} + Q^R_{M,r} w^{\theta}_{M,r}}{w^{\theta}_{M,r}} \tag{27}$$

we get that

$$\begin{aligned} \pi^R_{M,r} &= \frac{Q^R_{M,r}}{\delta} \int_{e_M} exp\left(-\delta e_M^{-\theta}\right) \delta\theta e_M^{-\theta-1} de_M = \frac{Q^R_{M,r}}{\delta} \\ &= \frac{Q^R_{M,r} w^{\theta}_{M,r}}{Q^R_{A,r}\left(w_{A,r}\right)^{\theta} + Q^R_{M,r} w^{\theta}_{M,r}}. \end{aligned} \tag{28}$$

Similarly, we get that

$$\pi^N_{M,r} = \frac{Q^N_{M,r} w^{\theta}_{M,r}}{Q^N_{A,r}\left(w_{A,r}\right)^{\theta} + Q^N_{M,r} w^{\theta}_{M,r}}. \tag{29}$$

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

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

Hence, the total labor supply in the manufacturing sector is[17]

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

where $\Gamma(.)$ is the gamma function. Using that $\delta = \frac{Q^R_{M,r}}{\pi^R_{M,r}}$ (see (27) and (28)) 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 $\nu \equiv \Gamma\left(1 - \frac{1}{\theta}\right)$.

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

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

$$\begin{aligned} H_{M,r} &= H_{M,r}^N + H_{M,r}^R \\ &= L_r^N \times \left(\pi_{M,r}^N\right)^{\frac{\theta-1}{\theta}} \times \nu \times \left(Q_{M,r}^N\right)^{\frac{1}{\theta}} + L_r^R \times \left(\pi_{M,r}^R\right)^{\frac{\theta-1}{\theta}} \times \nu \times \left(Q_{M,r}^R\right)^{\frac{1}{\theta}} \\ &= L_r^N \nu \left(\pi_{M,r}^N\right)^{\frac{\theta-1}{\theta}} \left(Q_{M,r}^N\right)^{\frac{1}{\theta}} \left\{1 + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\pi_{M,r}^R}{\pi_{M,r}^N}\right)^{\frac{\theta-1}{\theta}} \times \left(\frac{Q_{M,r}^R}{Q_{M,r}^N}\right)^{\frac{1}{\theta}}\right\}. \end{aligned} \tag{30}$$

Similarly, agricultural labor supply is given

$$H_{A,r} = L_r^N \nu \left(\pi_{A,r}^N\right)^{\frac{\theta-1}{\theta}} \left(Q_{A,r}^N\right)^{\frac{1}{\theta}} \left\{1 + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\pi_{A,r}^R}{\pi_{A,r}^N}\right)^{\frac{\theta-1}{\theta}} \times \left(\frac{Q_{A,r}^R}{Q_{A,r}^N}\right)^{\frac{1}{\theta}}\right\}. \tag{31}$$

**Goods demand** Given consumers' preferences in (8), the goods market clearing conditions are

$$Y_{M,r} = \bar{\zeta} Z_r \left(H_r^M\right)^{\vartheta+1} = \frac{1}{p_{M,r}} \left(\frac{p_{M,r}}{P_M}\right)^{1-\sigma} (1-\alpha) \times \zeta Y \tag{32}$$

$$Y_{A,r} = H_{A,r}^{\gamma} T_r^{1-\gamma} = \frac{1}{p_{A,r}} \left(\frac{p_{A,r}}{P_A}\right)^{1-\sigma} \alpha \times \zeta Y \tag{33}$$

where $Y$ denotes aggregate income, $p_{A,r}$ and $p_{M,r}$ are the regional prices given in (24) and (25) and $P_M$ and $P_A$ are the usual CES price indices

$$P_s = \left(\sum_{r=1}^{R} p_{s,r}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}. \tag{34}$$

(32) and (33) simply stem from the fact that consumers spend a fraction $\alpha$ $(1-\alpha)$ on manufacturing (agricultural) products and the usual demand relationship stemming from the CES structure of the Armington aggregators. Finally, aggregate income is given by

$$Y = \sum_{r=1}^{R} Y_r = \sum_{r=1}^{R} \left(T_r R_r + w_{M,r} H_{M,r} + w_{A,r} H_{A,r} + B_r b_r\right). \tag{35}$$

**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 (35), (34), (24) and (25). Hence, we have $5 \times R$ unknowns. The five corresponding equilibrium conditions are the $2 \times R$ equilibrium conditions (32) and (33), the $2 \times R$ labor market clearing conditions (30) and (31) and the optimality condition for agricultural inputs (26). 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 (35)), the goods market clearing conditions (32) and (33) are unaffected, labor supply is unaffected (as only relative wages matter) and so are the optimality conditions (26). Hence, we can impose the normalization (see (34))

$$P_A = \left(\sum_{r=1}^{R} p_{s,r}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \left(\sum_{r=1}^{R} \left(\left(\frac{w_{A,r}}{\gamma}\right)^{\gamma} \left(\frac{R_r}{1-\gamma}\right)^{1-\gamma}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = 1. \tag{36}$$

Because (26) 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

$$x_r \equiv \frac{R_r}{w_{A,r}} \text{ and } u_r = \frac{w_{A,r}}{w_{M,r}}. \tag{37}$$

Hence, we can write aggregate income $Y_r$ and the aggregate prices $P_M$ and $P_A$, i.e. (35), (24) and (25), as

$$Y = \sum_{r=1}^{R} \frac{1}{\zeta} \left(T_r R_r + w_{M,r} H_{M,r} + w_{A,r} H_{A,r}\right) = \frac{1}{\zeta} \sum_{r=1}^{R} w_{M,r} \left(H_{M,r} + \frac{1}{1-\gamma} T_r u_r x_r\right) \tag{38}$$

$$P_A = \left[\sum_r \left(\Gamma w_{M,r} u_r x_r^{1-\gamma}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{39}$$

$$P_M = \left(\sum_{r=1}^{R} \left(\frac{1}{\bar{\bar{\zeta}}} \frac{w_{M,r}}{Z_r} \left(H_r^M\right)^{-\vartheta}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}, \tag{40}$$

where $\Gamma \equiv \left(\frac{1}{\gamma}\right)^{\gamma} \left(\frac{1}{1-\gamma}\right)^{1-\gamma}$ and $\bar{\zeta}$ is given in (23). Using the demand equation (32) and the definition of the price index we get

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

so that

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

Hence, we can write the equilibrium system as (see (32), (33), (30) and (31)) as

$$\left(H_r^M\right)^{1-\vartheta(\sigma-1)} = \bar{\zeta}^{\sigma-1} Z_r^{\sigma-1} w_{M,r}^{-\sigma} P_M^{\sigma-1} \times \left(1-\alpha\right) \frac{1}{\zeta} Y \tag{41}$$

$$\left(\frac{\gamma}{1-\gamma}\right)^{\gamma} \times x_r^{\gamma} \times T_r = \left(\Gamma w_{M,r} u_r x_r^{1-\gamma}\right)^{-\sigma} \times \alpha \frac{1}{\zeta} Y$$

$$H_{M,r} = \nu L_r^N \times \Lambda_M\left(\mu_r, u_r\right)$$

$$x_r \frac{\gamma}{1-\gamma} T_r = \nu L_r^N \times \Lambda_A\left(\mu_r, u_r\right), \tag{42}$$

where

$$\Lambda_M\left(\mu_r, u_r\right) = \left(\frac{1}{1+\psi_r^N u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}} \left(Q_{M,r}^N\right)^{\frac{1}{\theta}} + \frac{\mu_r}{1-\mu_r} \times \left(\frac{1}{1+\psi_r^R u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}} \left(Q_{M,r}^R\right)^{\frac{1}{\theta}} \tag{43}$$

$$\Lambda_A\left(\mu_r, u_r\right) = \left(\frac{\psi_r^N u_r^{\theta}}{1+\psi_r^N u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}} \left(Q_{A,r}^N\right)^{\frac{1}{\theta}} + \frac{\mu_r}{1-\mu_r} \times \left(\frac{\psi_r^R u_r^{\theta}}{1+\psi_r^R u_r^{\theta}}\right)^{\frac{\theta-1}{\theta}} \left(Q_{A,r}^R\right)^{\frac{1}{\theta}}, \tag{44}$$

and $Y$ and $P_M$ are given in (38) and (40). To derive these equations we used that the agricultural good is the numeraire and that

$$\begin{aligned}
\pi_{M,r}^{g} &= \frac{Q_{M,r}^{g} w_{M,r}^{\theta}}{Q_{A,r}^{g}\left(w_{A,r}\left(1-\tau^{g}\right)\right)^{\theta}+Q_{M,r}^{g} w_{M,r}^{\theta}}=\frac{1}{1+\psi_{r}^{g} u_{r}^{\theta}} \\
\pi_{A,r}^{g} &= 1-\pi_{M,r}^{g}=\frac{\psi_{r}^{g} u_{r}^{\theta}}{1+\psi_{r}^{g} u_{r}^{\theta}} \\
\psi_{r}^{g} &\equiv \frac{Q_{A,r}^{g}}{Q_{M,r}^{g}}. \qquad (45)
\end{aligned}$$

### 7.4 Frictional land markets: A microfoundation for refugees' skills $\left(Q_A^R, Q_M^R\right)$

Suppose there is no agricultural labor market but agricultural goods are produced by self-employed farmers. Suppose all farmers have access to the technology

$$y = z^{1-\alpha} t^{\alpha},$$

where $z$ denotes the farmers' efficiency and $t$ denotes the amount of land. Let $p_A$ be the agricultural price and $R$ be the rental rate for a unit of land. It is easy to verify that land demands $t(z)$, production levels $y(z)$, revenue $s(z)$ and profits $\pi(z)$ are given by

$$\begin{aligned}
t(z) &= \left(\frac{\alpha p_A}{R}\right)^{\frac{1}{1-\alpha}} z \\
y(z) &= \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}} z \\
s(z) &= p_A y(z) = p_A \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}} z \\
\pi(z) &= (1-\alpha) s(z) = (1-\alpha) p_A \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}} z.
\end{aligned}$$

Suppose that refugees face constraints in the market for land. In particular, suppose they can only rent land up to a multiple $\lambda$ of the their sales, i.e.

$$Rt \leq \lambda \times p_A z^{1-\alpha} t^{\alpha}.$$

Then it is easy to see that the constraint is binding for all refugees if and only if $\lambda < \alpha$.[18] Assume that this is the case. Refugees' land demand therefore satisfies

$$Rt = \lambda \times p_A z^{1-\alpha} t^{\alpha},$$

so that

$$t(z,\lambda) = \left(\lambda \times \frac{p_A}{R}\right)^{\frac{1}{1-\alpha}} z = \left(\frac{\lambda}{\alpha}\right)^{\frac{1}{1-\alpha}} \times t(z).$$

[18] To see this, consider the unconstrained land choice. Then $Rt = \alpha p_A z^{1-\alpha} t^{\alpha}$. Hence, for $\lambda > \alpha$, the unconstrained choice is feasible. In contrast: for $\lambda < \alpha$ the constrained is always binding.

Hence,

$$
\begin{aligned}
t^{R}(z) &= \left(\frac{\lambda}{\alpha}\right)^{\frac{1}{1-\alpha}} \times t(z) \\
y^{R}(z) &= z^{1-\alpha}\left(t^{R}(z)\right)^{\alpha} = \left(\frac{\lambda}{\alpha}\right)^{\frac{\alpha}{1-\alpha}} \times y(z) \\
s^{R}(z) &= p_A y^{R}(z) = \left(\frac{\lambda}{\alpha}\right)^{\frac{\alpha}{1-\alpha}} \times s(z) \\
\pi^{R}(z) &= (1-\alpha)\, s^{R}(z) = \left(\frac{\lambda}{\alpha}\right)^{\frac{\alpha}{1-\alpha}} \times \pi(z).
\end{aligned}
$$

**Occupational Choice** Each individual draws a vector $(z, m)$ of efficiency units in each sector. Workers' optimal choice is therefore given by

$$z \times \tau_i \times \overline{\pi}(p_A, R) \geq mw,$$

where $w$ is the prevailing wage paid by manufacturing firms,

$$\tau_R = \left(\frac{\lambda}{\alpha}\right)^{\frac{\alpha}{1-\alpha}} < 1 = \tau_N,$$

and

$$\overline{\pi}(p_A, R) = (1-\alpha)\, p_A \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}}. \tag{46}$$

Let $(z, m)$ be drawn from a Frechet distribution with dispersion $\theta$ and location $Q_A$ and $Q_M$ respectively. Note that $Q_s$ does not vary between refugees and natives. Manufacturing employment shares are then given by

$$\omega_M^i = \frac{w_M^{\theta} Q_M}{w_M^{\theta} Q_M + (\tau_i)^{\theta}\, \overline{\pi}(p_A, R)^{\theta}\, Q_A^i}. \tag{47}$$

Hence, $\frac{\omega_M^R/\omega_A^R}{\omega_M^N/\omega_A^N} = \tau_R^{-\theta} > 1$. The total amount of efficiency units provided to the respective sector is given by $H_s^i = \left(\omega_M^i\right)^{\frac{\theta-1}{\theta}} (Q_s)^{\frac{1}{\theta}}$ and average earnings are

$$e^i = \left(w_M^{\theta} Q_M + (\tau_i)^{\theta}\, \overline{\pi}(p_A, R)^{\theta}\, Q_A^i\right)^{1/\theta}.$$

Hence, refugees have lower earnings (in *both* sectors).

**Sectoral Aggregates** Now consider the aggregation of the economy. Total labor payments in the manufacturing sector

$$wH_M = wL_r Q_M^{1/\theta} \times \left((1-\mu_r)\left(\omega_M^N\right)^{\frac{\theta-1}{\theta}} + \mu_r \left(\omega_M^R\right)^{\frac{\theta-1}{\theta}}\right). \tag{48}$$

Aggregate revenue in the agricultural sector is given by

$$
\begin{aligned}
p_A Y_A &= \sum_{g=N,R} \int_{z \in A^g} s(z)\, dz = p_A \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}} \sum_{g=N,R} \int_{z \in A^g} z dz \\
&= p_A \left(\frac{\alpha p_A}{R}\right)^{\frac{\alpha}{1-\alpha}} (Q_A)^{\frac{1}{\theta}} L_r \left[\left(\omega_A^N\right)^{\frac{\theta-1}{\theta}} (1-\mu_r) + \mu_r \tau_R \left(\omega_A^R\right)^{\frac{\theta-1}{\theta}}\right].
\end{aligned} \tag{49}
$$