Additional Results for "Market Size and Spatial Growth - Evidence from Germany's Post-War Population Expulsions"

OA-1 Theoretical Results

OA-1.1 Sectoral Labor Supply and Earnings

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

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

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

Then

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

Note that this object does not depend on i.

Proof. Suppose that $i = 1$ and let us derive the conditional distribution of x_1, conditional that x_1 is the highest $\{x_j\}_j$. The joint distribution of $\{x_j\}_j$ is given by

$$F(x_1, x_2, ..., x_3) = \prod_{j=1}^n F(x_i) \tag{OA-2}$$

because of independence. Hence, we get that

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

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

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

Using (OA-2), we get that

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

Hence,

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

Substituting this above yields

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

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

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

This implies (OA-1) $\square$

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

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

Hence, Lemma 3 implies that

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

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

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

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

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

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

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

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

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

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

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

Note that we can write

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

Hence,

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

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

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

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

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

OA-1.2 Deriving the Impulse Response Function $\Psi_d(p,\lambda)$

The change in the number of varieties N_{rt} in response to a change in the manufacturing workforce $\{d\ln H_{rPj}\}_{j=t_0}^{t_o+t}$ is given by

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

Assume that the initial shock at t_0 dies out slowly, i.e. $d\ln H_{rPj} = d\ln H_{rPt_0} \times p^{j-t_0}$where $p \leq 1$. This implies that

$$\begin{aligned} \sum_{j=t_0}^{t_0+d} \lambda^{-(j-(t_0+d))} d\ln H_{rPj} &= \sum_{j=t_0}^{t_0+d} \lambda^{-(j-(t_0+d))} d\ln H_{rPt_0} \times p^{j-t_0} = d\ln H_{rPt_0}\left(\sum_{j=t_0}^{t_0+d} \lambda^{-(j-(t_0+d))} p^{j-t_0}\right) \\ &= d\ln H_{rPt_0}\left(\lambda^{\tau}\sum_{i=0}^{d}\left(\frac{p}{\lambda}\right)^i\right) = d\ln H_{rPt_0}\frac{\lambda^{d+1}-p^{d+1}}{\lambda-p}. \end{aligned}$$

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

OA-1.3 Market Size and Variety Creation: Supply vs Demand

Equation (8) highlights that the growth of local varieties is fully determined by mass of local production workers H_{rPt}

$$\frac{N_{rt}}{N_{rt-1}} \quad = \quad \frac{1}{\rho-1}\frac{1}{f_E} H_{rPt} \times N_{rt-1}^{\lambda-1}.$$

Because H_{rPt} is an equilibrium outcome, both demand and supply forces are at play. To distinguish these forces, it is useful to consider a special case of the model where regions are sufficiently small to not affect the economy-wide aggregates. This implies that neither the size of the local population nor local factor prices affect the demand for tradable goods.

Using the two market clearing conditions for agricultural and manufacturing goods ((SM-2) and (SM-3)) and the equilibrium market size H_{rPt} relative to the manufacturing work force H_{rMt} yields

$$\begin{aligned} w_{rA}^{\sigma} H_{rA}^{1+(1-\gamma)(\sigma-1)} &= T_r^{(\sigma-1)(1-\gamma)} Q_{rt}^{\sigma-1}\gamma^{\sigma}\mathcal{D}_{rAt} \\ w_{rMt}^{\sigma}\left(\frac{\rho-1}{\rho}\right)^{1-\vartheta(\sigma-1)}\left(H_{rMt}+(1-\delta) f_E N_{rt-1}^{1-\lambda}\right)^{1-\vartheta(\sigma-1)} &= Q_{rt}^{\sigma-1} N_{rt-1}^{(\sigma-1)\lambda\vartheta}\varsigma^{1-\sigma}\beta\mathcal{D}_{rMt}, \end{aligned}$$

where $\mathcal{D}_{rMt} = (1-\alpha)\sum_j \frac{\tau_{rj}^{1-\sigma} Y_{jt}}{\sum_m \tau_{mj}^{1-\sigma} P_{mmM}^{1-\sigma}}$ and $\mathcal{D}_{rAt}$ similarly. Together with the labor supply functions

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

these four equations determine the four unknowns $(w_{rMt}, w_{rAt}, H_{rAt}, H_{rMt})$ as a function of local technologies $(T_r.Q_{rt}, N_{rt-1})$ and demand $(\mathcal{D}_{rMt}, \mathcal{D}_{rAt})$. The restriction that region r is small implies that we can treat $\mathcal{D}_{rjt}$ as exogenous from the point of view of region r.

To simply these expressions, assume first that labor supply is perfectly substitutable across sectors and abstract from type heterogeneity, that is $\theta \to \infty$ and $\phi_j^F = \phi_j^A = 1$. This implies that wages equalizes across sectors, $w_{rAt} = w_{rMt} = w_{rt}$, and that labor market clearing can be written as

$$H_{rAt} = (1 - s_{rMt}) L_{rt} \text{ and } H_{rMt} = s_{rMt} L_{rt}.$$

Furthermore, assume that the allocation is close to a steady-state. This implies that $H_{rP} = \frac{\rho-1}{\rho-1+\delta} H_{rMt}$. Under these assumptions, the local employment allocation s_{rMt} is determined from

$$\frac{(s_{rMt})^{1-\vartheta(\sigma-1)}}{(1-s_{rMt})^{1+(1-\gamma)(\sigma-1)}} = \xi \times L_{rt}^{(\vartheta+1-\gamma)(\sigma-1)} \times \frac{N_{rt-1}^{(\sigma-1)\lambda\vartheta}}{T_r^{(\sigma-1)(1-\gamma)}} \times \frac{\mathcal{D}_{r,M}}{\mathcal{D}_{r,A}},$$

where ξ is a constant. Note that the LHS is strictly increasing in s_{rMt}. Hence, as required in (SM-8),

$$s_{rMt} = h\left(L_{rt}^{(\vartheta+1-\gamma)(\sigma-1)} \times \frac{N_{rt-1}^{(\sigma-1)\lambda\vartheta}}{T_r^{(\sigma-1)(1-\gamma)}} \times \frac{\mathcal{D}_{r,M}}{\mathcal{D}_{r,A}} \right)$$

and $h(.)$ is increasing.

OA-1.4 The Model with Traded Intermediate Goods

In my baseline analysis I assumed that final goods are tradable, but intermediate goods are not. This modeling device is useful to highlight that productivity growth, which I here model as an expansion of varieties, is to a large extent local in nature. To see this more specifically, abstract for simplicity from the agricultural sector and suppose that intermediate goods were traded and also subject to a trade cost $\tau_{rj}^I \geq 1$, with $\tau_{rr}^I = 1$. Also assume that the manufacturing sector uses labor, i.e.

$$Y_M = \mathcal{Q}_{rt} H_{rP}^{\beta} X_{rt}^{1-\beta}.$$

The baseline model in the main text is the special case of $\beta = 0$. The intermediate good in location r then has a price of

$$P_{rXt} = \left(\sum_j \int_{i \in N_{jt}} p_{jrt}^{1-\rho} \right)^{\frac{1}{1-\rho}} = \left(\sum_j N_{jt} \left(\tau_{jr}^I\right)^{1-\rho} w_{jt}^{1-\rho} \right)^{\frac{1}{1-\rho}}.$$

The price of the final good from region r is then given by

$$\begin{aligned} P_{rt} &= \frac{1}{\mathcal{Q}_{rt}} \left(\frac{w_{rt}}{\beta}\right)^{\beta} \left(\frac{P_{rXt}}{1-\beta}\right)^{1-\beta} \\ &= \frac{1}{\mathcal{Q}_{rt}} \left(\frac{1}{\beta}\right)^{\beta} \left(\frac{1}{1-\beta}\right)^{1-\beta} \left(N_{rt} + \sum_{j \neq r} \left(\tau_{jr}^I\right)^{1-\rho} N_{jt} \left(\frac{w_{jt}}{w_{rt}}\right)^{1-\rho} \right)^{\frac{1-\beta}{1-\rho}} w_{rt} \end{aligned}$$

Hence, labor productivity in region r is given by

$$TFP_{rt} \propto \mathcal{Q}_{rt} \times \left(N_{rt} + \sum_{j \neq r} \left(\tau_{jr}^I\right)^{1-\rho} N_{jt} \left(\frac{w_{jt}}{w_{rt}}\right)^{1-\rho} \right)^{\frac{1-\beta}{\rho-1}}.$$

Total profits of a given intermediate producer in region r are given by

$$\pi_{rt} = \frac{1}{\rho} \frac{1-\beta}{\beta} w_{rt}^{1-\rho} \sum_j \frac{\left(\tau_{jr}^I\right)^{1-\rho}}{\sum_m N_{mt} \left(\tau_{mj}^I\right)^{1-\rho} w_{mt}^{1-\rho}} w_{jt} H_{jPt}$$

Hence, free entry requires that

$$w_{rt} f_E N_{rt-1}^{-\lambda} = \pi_{rt} = \frac{1}{\rho} \frac{1-\beta}{\beta} w_{rt}^{1-\rho} \sum_j \frac{\left(\tau_{jr}^I\right)^{1-\rho}}{\sum_m N_{mt} \left(\tau_{mj}^I\right)^{1-\rho} w_{mt}^{1-\rho}} w_{jt} H_{jPt}.$$

This implies that

$$w_{rt}^{\rho} f_E N_{rt-1}^{-\lambda} = \frac{1}{\rho} \frac{1-\beta}{\beta} \sum_j \frac{\left(\tau_{jr}^I\right)^{1-\rho}}{\sum_m N_{mt} \left(\tau_{mj}^I\right)^{1-\rho} w_{mt}^{1-\rho}} w_{jt} H_{jPt}. \tag{OA-3}$$

To solve for the mass of workers engaged in the production of intermediate goods, note that total labor payments in the intermediate sector are proportional to total profits

$$w_{rt} H_{rXt} = (\rho - 1) N_{rt} \pi_{rt}.$$

Hence,

$$H_{rXt} = (\rho - 1) N_{rt} \frac{\pi_{rt}}{w_{rt}} = (\rho - 1) N_{rt} f_E N_{rt-1}^{-\lambda}.$$

Finally, the number of entry workers are given by

$$H_{rEt} = f_E N_{rt-1}^{-\lambda} N_{rt} - (1-\delta) f_E N_{rt-1}^{1-\lambda}$$

Hence, labor market clearing requires

$$L_{rt} = H_{rPt} + \rho N_{rt} f_E N_{rt-1}^{-\lambda} - (1-\delta) f_E N_{rt-1}^{1-\lambda}. \tag{OA-4}$$

The market clearing for the traded final good is given by

$$w_{rt} H_{rPt} \quad = \quad \sum_{j=1}^{R} \left(\frac{\tau_{rj} P_{rMt}}{\overline{P}_{jMt}}\right)^{1-\sigma} \left(w_{jt} H_{jPt}\right), \tag{OA-5}$$

where $P_{rMt} = \frac{1}{\mathcal{Q}_{rt} \times \left(N_{rt} + \sum_{j \neq r} \left(\tau_{jr}^I\right)^{1-\rho} N_{jt} \left(\frac{w_{jt}}{w_{rt}}\right)^{1-\rho}\right)^{\frac{1-\beta}{\rho-1}}} w_{rt}$, $\overline{P}_{jMt} = \left(\sum_r \left(\tau_{rj} P_{rMt}\right)^{1-\sigma}\right)^{1/(1-\sigma)}$. Given the state variables N_{rt-1} and L_{rt}, the equilibrium is described by $R \times 3$ unknowns: $\{w_{rt}, N_{rt}, H_{rPt}\}_{r=1}^{R}$. These unknowns are determined from (OA-3), (OA-4) and (OA-5).

Prohibitive costs for intermediate goods: $\tau_{rj}^I = \infty$ To see that this setup is a strict generalization of the model the main text, suppose that intermediate goods are not traded, i.e. $\tau_{rj}^I = \infty$ and $\tau_{rr}^I = 1$. Equation (OA-3) then reduces to

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

which is the equation in the main text. Hence, in the absence of trade, local varieties only depends on local factor supply. Additionally, the labor market clearing condition in (OA-4) also simplifies to

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

Hence, these two equations directly determine N_{rt} and H_{rPt} as a function of local labor supply. Local wages $\{w_{rt}\}$ are then determined from the general equilibrium condition (OA-5).

Free trade for intermediate goods: $\tau_{rj}^{I}=1$ The other interesting special case is the one where intermediate goods are freely tradable. In that case, the free entry condition requires that

$$w_{rt}^{\rho}=\frac{1}{f_E}\frac{1-\beta}{\rho\beta}N_{rt-1}^{\lambda}\Phi_t, \tag{OA-6}$$

where $\Phi_t=\frac{\sum_j w_{jt}H_{jPt}}{\sum_j N_{jt}w_{jt}^{1-\rho}}$. Hence, local wages are independent of local factor supplies and directly tied to the existing number of varieties N_{rt-1}. Intuitively: If intermediate products are freely traded, profits are dissociated from local demand and are given by

$$\pi_{rt}=w_{rt}^{1-\rho}\times\Lambda_t,$$

where Λ_t is constant across space. The free entry condition requires that π_{rt} are tied to the entry costs, which are proportional to $w_{rt}N_{rt}^{-\lambda}$. Putting these together yields (OA-6). Given wages, the number of local varieties N_{rt} and the mass of production workers H_{rPt} are then determined from (OA-4) and (OA-5). However, note that local prices P_{rMt} are given by

$$P_{rMt}=\frac{1}{\mathcal{Q}_{rt}\left(\sum_{j\neq r}N_{jt}w_{jt}^{1-\rho}\right)^{\frac{1-\beta}{\rho-1}}}w_{rt}^{\beta},$$

i.e. relative prices across regions only depend in $\mathcal{Q}_{rt}$ and w_{rt} and not local market size. Even though local wages do not depend on the local population conditional on N_{rt-1} (see (OA-6)), a population shock does affect local wages through variety creation. In particular, an increase in L_{rt} will in general increase both H_{rPt} and N_{rt}.[45]

OA-1.5 The Model with Vertical Innovation

For the main analysis I used a model with horizontal innovation in the spirit of Romer (1990). To see that this focus on variety gains is a modeling device, in this section I consider a model with vertical innovations, i.e. improvements in quality, that delivers a very similar structure.

The setup is exactly as in the baseline model. The only difference is that the local intermediate input X_{rt} is given by

$$X_{rt}=\left(\int_0^1 z_{irt}^{\frac{1}{\rho_Z}}x_{irt}^{\frac{\rho_Z-1}{\rho_Z}}di\right)^{\rho_Z/(\rho_Z-1)},$$

where z_{irt} denotes the quality of input i at time t in region r and x_{irt} denotes the quantity. The number of firms in region r is normalized to unity and remains constant over time.

Firms have access to a technology that allows them to increase the quality of their products. Assume that to change the quality from z_{irt-1} to ϖz_{irt-1}, a firm needs to hire

$$h\left(\varpi;z_{irt-1},Z_{rt-1}\right)=f_Z z_{it-1}\varpi^{\zeta}Z_{rt-1}^{-\lambda_Z}$$

units of labor, where $\zeta>1$ and $\lambda_Z>0$. Three comments are in order. First, note the dependence of innovation costs on z_{it-1} and Z_{rt-1}. Innovation costs are rising in z_{it-1}, reflecting the fact that quality improvements are more expensive the higher the base quality. Also, innovation costs are decreasing in Z_{rt-1} (provided that $\lambda>0$), capturing the possibility of local innovation spillovers. Second, innovation costs are convex in ϖ. Third, I do not restrict $\varpi>1$. This is for analytical convenience. As an economic justification, it captures the idea that firms need to spend some resources to keep their technology usable.

[45]For example. in a steady state $N_{rt}=N_{rt-1}=N_r$, the labor market clearing condition (OA-4) implies that $d\ln H_{rPt}=d\ln L_{rt}$ and $d\ln N_{rt}=\frac{1}{1-\lambda}d\ln L_{rt}$.

This model has almost the same implications as the baseline model. The profit of firm i with quality z_{it} in region r is given by

$$\pi_{irt} = \frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{z_{irt}}{Z_{rt}} w_{rt} H_{rPt} \text{ where } Z_{rt} = \int z_{irt} di.$$

Hence, firms' optimal innovation choice is given by

$$\varpi_{rt}(z) \equiv \max_{\varpi} \left\{ \frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{\varpi z}{Z_{rt}} w_{rt} H_{rPt} - f_Z z \varpi^{\zeta} Z_{rt-1}^{-\lambda_Z} w_{rt} \right\}.$$

The solution $\varpi_{rt}(z)$ is thus defined by

$$\frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{1}{Z_{rt}} H_{rPt} = \zeta f_Z \varpi_{rt}(z)^{\zeta-1} Z_{rt-1}^{-\lambda_Z}.$$

Hence, $\varpi_{rt}(z) = \varpi_{rt}$ and $Z_{rt} = \varpi_{rt} Z_{rt-1}$. This implies that

$$\varpi_{rt} = \left(\frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{1}{\zeta f_Z} \right)^{1/\zeta} H_{rPt}^{1/\zeta} Z_{rt-1}^{-\frac{1-\lambda_Z}{\zeta}}$$

and

$$Z_{rt} = \varpi_{rt} Z_{rt-1} = \left(\frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{1}{\zeta f_Z} \right)^{1/\zeta} H_{rPt}^{1/\zeta} Z_{rt-1}^{1-\frac{1-\lambda_Z}{\zeta}}. \tag{OA-7}$$

The aggregate production function for the tradable good in region r is given by

$$Y_{rt} = \mathcal{Q}_{rt} \left(\frac{1-\beta_Z}{\beta_Z} \frac{\rho_Z - 1}{\rho_Z} \right)^{1-\beta_Z} Z_{rt}^{\frac{1-\beta_Z}{\rho_Z - 1}} H_{rPt}.$$

Substituting (OA-7) yields

$$Y_{rt} = \varsigma_Z \mathcal{Q}_{rt} Z_{rt-1}^{\left(1-\frac{1-\lambda_Z}{\zeta}\right)\frac{1-\beta_Z}{\rho_Z-1}} H_{rPt}^{1+\frac{1}{\zeta}\frac{1-\beta_Z}{\rho_Z-1}}, \tag{OA-8}$$

where ς_Z is an inconsequential constant.[46]

To see the similarity between this model and the model in the main text, define $\mathcal{Z}_{rt} = Z_{rt}^{\zeta}$. (OA-7) then implies that

$$\mathcal{Z}_{rt} = \left(\frac{1-\beta_Z}{\beta_Z \rho_Z} \frac{1}{\zeta f_Z} \right) H_{rPt} \mathcal{Z}_{rt-1}^{1-\frac{1-\lambda_Z}{\zeta}}.$$

Similarly, (OA-8) reads

$$Y_{rt} = \varsigma_Z \mathcal{Q}_{rt} \mathcal{Z}_{rt-1}^{\left(1-\frac{1-\lambda_Z}{\zeta}\right)\frac{1}{\zeta}\frac{1-\beta_Z}{\rho_Z-1}} H_{rPt}^{1+\frac{1}{\zeta}\frac{1-\beta_Z}{\rho_Z-1}}.$$

In the baseline model, the production function and the law of motion for N_{rt} are given by (see (9) and (8))

$$\begin{aligned} Y_{rMt} &= \varsigma \mathcal{Q}_{rt} N_{rt-1}^{\lambda \frac{1-\beta}{\rho-1}} H_{rPt}^{1+\frac{1-\beta}{\rho-1}} \\ N_{rt} &= \frac{1-\beta}{\rho \beta} \frac{1}{f_E} H_{rPt} N_{rt-1}^{\lambda}. \end{aligned}$$

Hence, the law of motion for productivity and the aggregate production function are isomorphic as long as

$$1 - \lambda_Z = (1-\lambda)\zeta \tag{OA-9}$$

$$\frac{1}{\zeta} \frac{1-\beta_Z}{\rho_Z - 1} = \frac{1-\beta}{\rho - 1}. \tag{OA-10}$$

[46] This constant is given by $\varsigma_Z = \left(\frac{1-\beta_Z}{\beta_Z} \frac{\rho-1}{\rho} \right)^{1-\beta_Z} \left(\frac{1-\beta_Z}{\beta_Z \rho} \frac{1}{\zeta f_Z} \right)^{\frac{1}{\zeta} \frac{1-\beta_Z}{\rho-1}}$.

In terms of equilibrium labor allocations, it is easy to verify that the share of production workers $\left(H_{rPt}/H_{rt}^{M}\right)$, innovation workers $\left(H_{rZt}/H_{rt}^{M}\right)$ and workers for intermediate production $\left(H_{rXt}/H_{rt}^{M}\right)$ satisfy

$$s_{rPt} = \frac{\beta_Z}{1-\frac{1-\beta_Z}{\rho_Z}\frac{\zeta-1}{\zeta}} \quad s_{rXt} = \frac{\beta_Z}{1-\frac{1-\beta_Z}{\rho_Z}\frac{\zeta-1}{\zeta}}\frac{\rho_Z-1}{\rho_Z}\left(\frac{1-\beta_Z}{\beta_Z}\right) \quad s_{rZt} = \frac{\beta_Z}{1-\frac{1-\beta_Z}{\rho_Z}\frac{\zeta-1}{\zeta}}\frac{1-\beta_Z}{\rho_Z\beta_Z}\frac{1}{\zeta}.$$

Hence, the allocation of workers across activities within the manufacturing sector is constant. This is the only implication, which is different in the baseline model, if varieties are long-lived. However, if $\delta = 1$ and varieties depreciate fully, the baseline model implies that

$$s_{rPt}^{Base} = \beta \quad s_{rXt}^{Base} = \beta\frac{\rho-1}{\rho}\left(\frac{1-\beta}{\beta}\right) \quad s_{rEt}^{Base} = \beta\frac{1-\beta}{\rho\beta}.$$

Note that the equilibrium allocation only depend on H_{rPt}. Hence, the equilibrium path is exactly the same in both models as long

$$\beta = \frac{\beta_Z}{1-\frac{1-\beta_Z}{\rho_Z}\frac{\zeta-1}{\zeta}}. \tag{OA-11}$$

For given parameters in the baseline model, we can hence find λ_Z, ρ_Z and β_Z to satisfy the restrictions (OA-9), (OA-10) and (OA-11) while still ensuring that $\zeta > 1$, $0 < \beta < 1$ and $\rho > 1$.

OA-2 Empirical Results

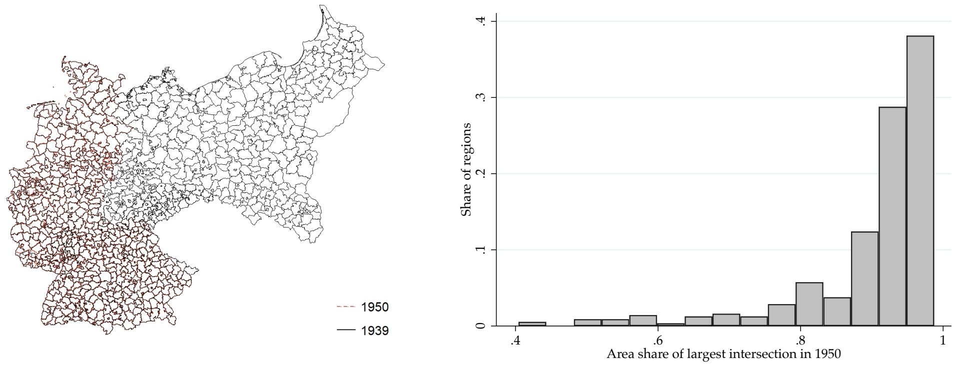
OA-2.1 Datasources

The data for the years 1933 and 1939 is published in Statistisches Reichsamt (1936) and Statistisches Reichsamt (1939). For the post-war data I had to rely on numerous publications for the individual states. For the state of North Rhine Westphalia (*Nordrhein-Westfalen*) the data is taken from Statistisches Landesamt Nordrhein-Westfalen (1952) and Statistisches Landesamt Nordrhein-Westfalen (1964). For the state of Bavaria (*Bayern*) the data is taken from Bayerisches Statistisches Landesamt (1953) and Bayerisches Statistisches Landesamt (1963b). For the state of Rhineland Palatinate (*Rheinland-Pfalz*) the data is taken from Statistisches Landesamt Rheinland-Pfalz (1950) and Statistisches Landesamt Rheinland-Pfalz (1961). For the state of Lower Saxony (*Niedersachsen*) the data is taken from Niedersächsisches Amt für Landesplanung und Statistik (1952) and Niedersächsisches Amt für Landesplanung und Statistik (1964). For the state of Hesse (*Hessen*) the data is taken from Hessisches Statistisches Landesamt (1952) and Hessisches Statistisches Landesamt (1962). The share of refugees in 1946 is taken from Ausschuss der Deutschen Statistiker für die Volks- und Berufszählung 1946 (1949). This data was only available for a subset of states.

OA-2.2 Construction of Time-invariant Boundaries

To perform my analysis, I construct a map of Western Germany with time-invariant district boundaries since the 1930s. To construct this crosswalk, I use GIS to perform a partition of the respective maps. Given this partition, I then aggregate all variables to the level of 1950 for the main analysis and to the level of 1975 for the analysis of the long-run response of income per capita. GDP response. I aggregate the respective data by using the size of the area as weights.

In Figures OA-1-OA-4 I depict for the respective episodes the overlay of the GIS maps with the district boundaries (left panel) and the distribution of the largest share of the intersections of the districts to be aggregated (right panel). More specifically, consider an "origin" map M^O that we want to aggregate to a "destination" map M^D. For example, to calculate the population in 1939 in the borders in 1950, the 1939 map would be the origin map and the map in 1950 would be the destination map. Let $\mathcal{R}^O$ and $\mathcal{R}^D$ denote the set of regions in the origin and destination map and let $\mathcal{R}^I$ the set of intersections between M^O and M^D. Then consider a region o in the origin map, $o \in \mathcal{R}^O$. Let for each region $o \in \mathcal{R}^O$ the set of intersection be $k(o)$. If o gets mapped to a single region $d \in \mathcal{R}^D$, then $k(o) = d$. If o gets split and is now partly contained in the three



(a) Map (b) Distribution of $\xi(o)$

Notes: The left panel shows the districts of the map of West Germany in 1950 (red boundaries) and the map of the German Reich in 1939 (black boundaries). The right panel shows the distribution of largest share of area of a county in 1950 within a district in 1939. Formally, it displays the distribution of $\xi(o) = \max_{i \in k(o)} \{a_i/a_o\}$, where o refers to a county in 1950, $k(o)$ the set of intersections between o and the map in 1939 and a_i (a_o) is the geographical area of a given intersection (a given origin county).

Figure OA-1: Construction of Time-invariant Boundaries: 1939 - 1950

countries d_1 d_2 and d_3, then $k(o) = \{d_1, d_2, d_3\}$.

Let a_i denote the geographical area of a given intersection and a_0 the area of the origin county o. Then consider the statistic $\xi(o) = \max_{i \in k(o)} \left\{ \frac{a_i}{a_o} \right\}$, i.e. the maximum area share of an origin locations that is left intact. If $\xi(o) = 1$,the location is fully contained in the new destination region. The smaller $\xi(o)$ the more severe the aggregation bias of constructing the crosswalk.

In Figure OA-1 I show the intersection of the maps in 1939 and 1950. I depict the borders of 1939 in black and the one in 1950 in red. The 1950 borders are only drawn for the area of Western Germany, which is the area of my analysis. Both the map (left panel) and the histogram (right panel) show that the majority of counties in 1950 were already part of a single county in 1933. For almost 90% of counties, more than 80% of the county's area remained intact.

My main empirical analysis focuses on the time period between 1950 and 1961. This analysis did not require any crosswalk because there were no border changes between 1950 and 1961. This is seen in Figure OA-2, which shows the intersection of the maps in 1950 and 1961.

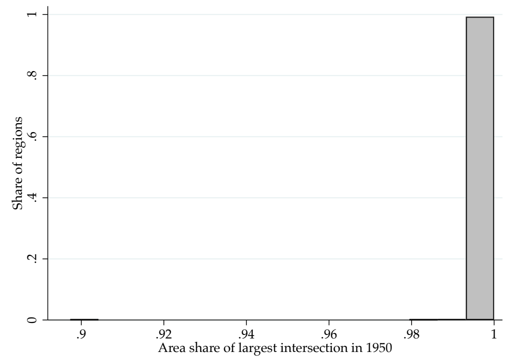
Finally, the analysis of the effect on income per capita in the 70s, 80s and 90s, required substantial aggregation of the data because the 1970s saw a variety of reforms where district boundaries where changed. In Figures OA-3 and OA-4 I show the intersection of the maps in 1950 and 1975 and between 1975 and 1988. Figure OA-3 shows that the number of counties shrank considerably between 1950 and 1975. At the same time, some borders were changed in a way so that a given county in 1950 was split. Still, for the vast majority of counties, the largest share of their area was still contained in a single district in 1975. By contrast, Figure OA-4 shows that the border changes between 1975 and 1988 were minor.

OA-2.3 Additional results: Instrumental Variable Estimates

In Table OA-1 I present the reduced form estimates using the same set of controls as for the IV estimates reported in Table 7. Given these measurement concern about the role of the expulsion distance in Bavaria, I allow for the reduced form effect to differ between Bavaria and the rest of Germany (see columns 2 and 6 in Table SM-4).[47] Focusing on the role of the expulsion distance outside of Bavaria, regions with a larger distance

[47] For brevity I restrict the reduced form effects among states other than Bavaria to be identical. In the IV estimates, I instrument the share of refugees with the interaction between the distance to the expulsion region and a state fixed effect.

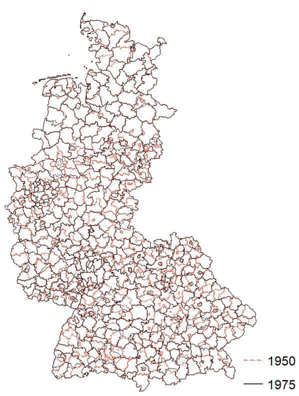
(a) Map



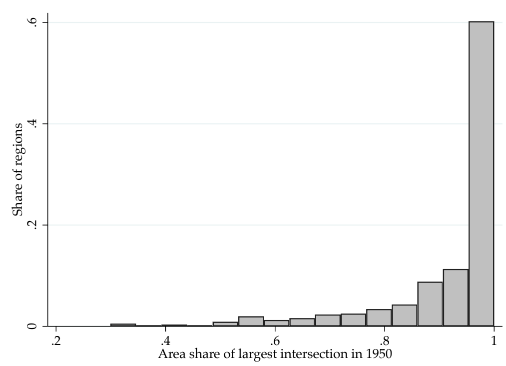
(b) Distribution of $\xi(o)$

Notes: The left panel shows the districts of the map of West Germany in 1950 (red boundaries) and 1961 (black boundaries). The right panel shows the distribution of largest share of area of a county in 1950 within a district in 1961. Formally, it displays the distribution of $\xi(o) = \max_{i \in k(o)} \{a_i/a_o\}$, where o refers to a county in 1950, $k(o)$ the set of intersections between o and the map in 1961 and a_i (a_o) is the geographical area of a given intersection (a given origin county).

Figure OA-2: Construction of Time-invariant Boundaries: 1950 - 1961



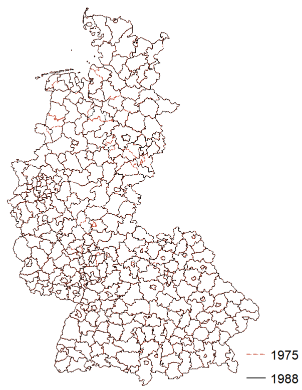
(a) Map



(b) Distribution of $\xi(o)$

Notes: The left panel shows the districts of the map of West Germany in 1950 (red boundaries) and 1975 (black boundaries). The right panel shows the distribution of largest share of area of a county in 1950 within a district in 1975. Formally, it displays the distribution of $\xi(o) = \max_{i \in k(o)} \{a_i/a_o\}$, where o refers to a county in 1950, $k(o)$ the set of intersections between o and the map in 1975 and a_i (a_o) is the geographical area of a given intersection (a given origin county).

Figure OA-3: Construction of Time-invariant Boundaries: 1950 - 1975



(a) Map

(b) Distribution of $\xi(o)$

Notes: The left panel shows the districts of the map of West Germany in 1975 (red boundaries) and 1988 (black boundaries). The right panel shows the distribution of largest share of area of a county in 1975 within a district in 1988. Formally, it displays the distribution of $\xi(o) = \max_{i \in k(o)} \{a_i/a_o\}$, where o refers to a county in 1975, $k(o)$ the set of intersections between o and the map in 1988 and a_i (a_o) is the geographical area of a given intersection (a given origin county).

Figure OA-4: Construction of Time-invariant Boundaries: 1975 - 1988

to the expulsion regions experience lower population growth, slower growth in manufacturing employment, less declines in agricultural employment and slower growth in industrial plants. There is no statistically significant impact on income growth and service employment growth. Moreover, the effect on manufacturing and agricultural employment is also only significant in some specifications.

In Table OA-2 I replicate the results from Table 7 when I weigh each observation with its local population in 1939. These results, which are very similar to the baseline estimates in Table 7, ease the concern that only small counties account for a large part of the variation. In Table OA-3 I report the results from Table 7 when I use robust standard errors instead of clustered standard errors. The main differences of this specification are that the relationship with the service share is now statistically positive in some specifications and that the effect on long-run GDP is just shy of being significant at the 10% level.

OA-2.4 The Experience of Different Cohorts: Additional Results From Micro Data

In Section 3.3 I documented pronounced differences in the way how refugees and natives changed their sectoral of employment along the life-cycle. In this section I provide additional evidence along these lines.

Consider first Table OA-4, which contains additional results on the occupational and sectoral employment shares between refugees and natives. The first two columns contain the results 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, individuals living in West 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 West Germany. In particular, their employment share in agriculture declines by more than 50%. At the same time, manufacturing employment among refugees increases dramatically, exceeds 50% and is now higher than for natives. The occupational data in the lower panel has additional information on these reallocation patterns: the decline in agricultural employment is essentially accounted for by a decline in self-employed farmers, i.e. famers who lost their land when being expelled. After the expulsion, these individuals take unskilled jobs, which are mostly in the manufacturing sector. The remaining columns in Table OA-4 show that these

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Population growth: $\ln L_{rt} - \ln L_{r1939}$								
	1939-1950				1939-1961			
ln ED × Rest	-0.657***	-0.386	-0.447**	-0.481**	-0.599*	-0.488	-0.660*	-0.553**
	(0.211)	(0.231)	(0.196)	(0.200)	(0.320)	(0.384)	(0.338)	(0.257)
ln ED × Bavaria	0.150	0.195	0.163	0.170	0.446***	0.457**	0.372*	0.404**
	(0.092)	(0.164)	(0.127)	(0.114)	(0.120)	(0.222)	(0.187)	(0.148)
Panel B: Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$								
	1939-1950				1939-1961			
ln ED × Rest	-0.031	-0.117	-0.101	-0.198***	-0.103**	-0.016	0.015	-0.085
	(0.045)	(0.096)	(0.097)	(0.071)	(0.043)	(0.106)	(0.102)	(0.079)
ln ED × Bavaria	0.031	-0.045	-0.015	-0.031	0.046	0.076	0.128	0.116*
	(0.048)	(0.080)	(0.080)	(0.061)	(0.057)	(0.082)	(0.077)	(0.062)
Panel C: Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$								
	1933-1950				1933-1961			
ln ED × Rest	0.099	0.101	0.237**	0.287***	0.058	-0.051	0.088	0.142
	(0.060)	(0.137)	(0.105)	(0.103)	(0.083)	(0.139)	(0.115)	(0.106)
ln ED × Bavaria	-0.051	-0.132	-0.063	-0.029	-0.077	-0.226**	-0.171*	-0.136
	(0.074)	(0.103)	(0.080)	(0.083)	(0.100)	(0.110)	(0.093)	(0.082)
Panel D: Service employment: $\pi^S_{rt} - \pi^S_{r1933}$								
	1933-1950				1933-1961			
ln ED × Rest	-0.045	-0.129	-0.068	-0.060	0.025	-0.105	-0.037	-0.026
	(0.038)	(0.093)	(0.104)	(0.081)	(0.038)	(0.107)	(0.123)	(0.096)
ln ED × Bavaria	0.105**	0.078	0.083	0.069	0.112*	0.049	0.061	0.044
	(0.048)	(0.047)	(0.052)	(0.043)	(0.055)	(0.050)	(0.058)	(0.046)
Panel E: GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$								
	1935-1950				1935-1961			
ln ED × Rest	0.201	-0.031	-0.155	-0.056	0.274	0.417	0.152	0.180
	(0.224)	(0.420)	(0.371)	(0.340)	(0.252)	(0.298)	(0.235)	(0.232)
ln ED × Bavaria	-0.849***	0.105	0.083	0.022	-0.907*	0.753**	0.681**	0.698**
	(0.257)	(0.334)	(0.362)	(0.292)	(0.454)	(0.294)	(0.267)	(0.256)
Panel F: Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$								
	1933-1950				1933-1956			
ln ED × Rest	0.139	-1.264***	-1.180***	-1.229***	-0.778	-0.459	-1.201**	-1.588***
	(0.276)	(0.409)	(0.305)	(0.278)	(0.528)	(0.862)	(0.453)	(0.394)
ln ED × Bavaria	0.223	-0.486	-0.536**	-0.518**	-0.041	0.362	-0.193	-0.345
	(0.276)	(0.289)	(0.200)	(0.194)	(0.604)	(0.522)	(0.327)	(0.252)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Pop. density (1939)		✓	✓	✓		✓	✓	✓
Wartime destr.		✓	✓	✓		✓	✓	✓
Geography		✓	✓	✓		✓	✓	✓
Levels of dep. variable		✓	✓	✓		✓	✓	✓
Pre-war controls			✓	✓			✓	✓
Addtl. pre-war controls				✓				✓

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

Table OA-1: The Effects of Refugee Inflows on the Local Economy: Reduced Form

Panel A: Population growth: $\ln L_{rt} - \ln L_{r1939}$								
	1939-1950				1939-1961			
Share of refugees in 1950	2.021***	1.718***	1.740***	1.700***	0.999***	1.540***	1.628***	1.591***
	(0.312)	(0.204)	(0.227)	(0.204)	(0.215)	(0.299)	(0.267)	(0.256)
N	526	526	509	463	526	526	509	463
F-Stat	23.735	14.228	12.466	14.460	18.174	17.052	14.866	17.578
Panel B: Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$								
	1939-1950				1939-1961			
Share of refugees in 1950	0.041	0.308**	0.344***	0.407***	0.245**	0.246**	0.262**	0.316***
	(0.147)	(0.130)	(0.125)	(0.082)	(0.106)	(0.124)	(0.119)	(0.093)
N	535	535	519	472	535	535	519	472
F-Stat	18.097	17.807	16.098	17.503	18.097	17.807	16.098	17.503
Panel C: Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$								
	1933-1950				1933-1961			
Share of refugees in 1950	-0.388**	-0.346**	-0.484***	-0.520***	-0.360	-0.213	-0.327**	-0.352**
	(0.184)	(0.143)	(0.131)	(0.153)	(0.325)	(0.159)	(0.138)	(0.156)
N	523	523	519	472	523	523	519	472
F-Stat	18.286	17.912	16.098	17.503	18.286	17.912	16.098	17.503
Panel D: Service employment: $\pi^S_{rt} - \pi^S_{r1933}$								
	1933-1950				1933-1961			
Share of refugees in 1950	0.153	0.191	0.193	0.182	-0.054	0.213	0.184	0.160
	(0.137)	(0.212)	(0.183)	(0.205)	(0.115)	(0.240)	(0.199)	(0.216)
N	523	523	519	472	523	523	519	472
F-Stat	18.286	19.515	16.098	17.503	18.286	19.515	16.098	17.503
Panel E: GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$								
	1935-1950				1935-1961			
Share of refugees in 1950	-0.707	0.934	0.364	-0.129	-0.506	0.905**	0.589***	0.483**
	(0.881)	(0.970)	(0.950)	(0.572)	(0.922)	(0.408)	(0.218)	(0.193)
N	523	523	519	472	519	519	515	468
F-Stat	18.286	19.464	16.098	17.503	17.663	19.148	16.188	17.101
Panel F: Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$								
	1933-1950				1933-1956			
Share of refugees in 1950	-0.246	2.562***	2.476***	2.481***	1.517	1.459	1.638*	1.632***
	(0.850)	(0.810)	(0.742)	(0.547)	(1.008)	(1.050)	(0.859)	(0.624)
N	520	520	519	472	520	520	519	472
F-Stat	17.224	16.952	16.098	17.503	17.224	16.952	16.098	17.503
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Pop. density (1939)		✓	✓	✓		✓	✓	✓
Wartime destr.		✓	✓	✓		✓	✓	✓
Geography		✓	✓	✓		✓	✓	✓
Levels of dep. variable		✓	✓	✓		✓	✓	✓
Pre-war controls			✓	✓			✓	✓
Addtl. pre-war controls				✓				✓

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

Table OA-2: IV Results (Table 7): Estimates with Population Weights

	Panel A: Population growth: $\ln L_{rt} - \ln L_{r1939}$							
	1939-1950				1939-1961			
Share of refugees in 1950	1.897***	1.459***	1.563***	1.614***	1.018***	1.227***	1.450***	1.501***
	(0.132)	(0.133)	(0.146)	(0.139)	(0.184)	(0.216)	(0.225)	(0.216)
N	526	526	509	463	526	526	509	463
F-Stat	47.905	26.685	22.276	27.334	47.109	26.326	21.728	26.908
	Panel B: Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$							
	1939-1950				1939-1961			
Share of refugees in 1950	0.124**	0.271***	0.297***	0.406***	0.279***	0.199*	0.222*	0.333***
	(0.062)	(0.077)	(0.085)	(0.067)	(0.084)	(0.106)	(0.114)	(0.099)
N	535	535	519	472	535	535	519	472
F-Stat	47.538	26.231	23.145	27.484	47.538	26.231	23.145	27.484
	Panel C: Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$							
	1933-1950				1933-1961			
Share of refugees in 1950	-0.337**	-0.441***	-0.573***	-0.607***	-0.261	-0.294**	-0.449***	-0.472***
	(0.158)	(0.154)	(0.128)	(0.103)	(0.181)	(0.145)	(0.119)	(0.101)
N	523	523	519	472	523	523	519	472
F-Stat	45.863	24.747	23.145	27.484	45.863	24.747	23.145	27.484
	Panel D: Service employment: $\pi^S_{rt} - \pi^S_{r1933}$							
	1933-1950				1933-1961			
Share of refugees in 1950	0.146	0.307**	0.228**	0.188**	-0.007	0.271**	0.202*	0.143
	(0.091)	(0.121)	(0.115)	(0.094)	(0.087)	(0.129)	(0.121)	(0.111)
N	523	523	519	472	523	523	519	472
F-Stat	45.863	25.204	23.145	27.484	45.863	25.204	23.145	27.484
	Panel E: GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$							
	1935-1950				1935-1961			
Share of refugees in 1950	-0.400	0.221	0.341	-0.003	-0.671	0.208	0.517	0.471
	(0.608)	(0.621)	(0.621)	(0.518)	(0.616)	(0.349)	(0.331)	(0.299)
N	523	523	519	472	519	519	515	468
F-Stat	45.863	23.799	23.145	27.484	45.897	24.072	23.293	27.898
	Panel F: Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$							
	1933-1950				1933-1956			
Share of refugees in 1950	-0.290	1.675**	1.553**	1.851***	1.449	1.583	2.097**	2.567***
	(0.689)	(0.726)	(0.774)	(0.409)	(1.068)	(1.080)	(1.012)	(0.705)
N	520	520	519	472	520	520	519	472
F-Stat	45.377	25.658	23.145	27.484	45.377	25.658	23.145	27.484
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Pop. density (1939)		✓	✓	✓		✓	✓	✓
Wartime destr.		✓	✓	✓		✓	✓	✓
Geography		✓	✓	✓		✓	✓	✓
Levels of dep. variable		✓	✓	✓		✓	✓	✓
Pre-war controls			✓	✓			✓	✓
Addtl. pre-war controls				✓				✓

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

Table OA-3: IV Results (Table 7): Estimates with Robust Standard Errors

	Pre expulsion 1939		Post expulsion 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.

Table OA-4: Sectoral and Occupational Mobility from 1939 to 1971

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 the occupational employment shares within sectors for refugees and natives in 1950.

Table OA-5: Occupational Distribution within Sectors: Natives vs Refugees

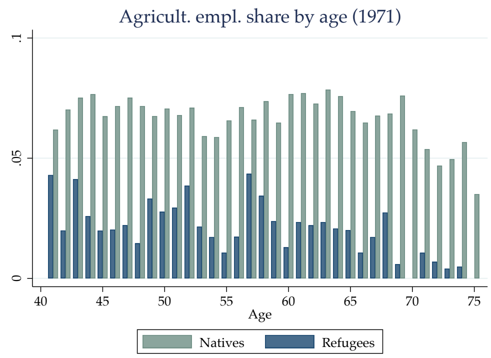
reallocation patterns in 1950 are not transitory but they persist throughout the 1950s and 1960s.

To see this even more directly, consider Table OA-5, 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, 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. While it also the case that natives were more likely to be self-employed in the manufacturing sector, the difference between natives and refugees was much less severe.

In Figure OA-5, which is also taken from the supplement of the 1971 Census, I show the agricultural employment share in 1971 for different age groups. I focus on all cohorts older than 40, i.e. born before 1930. Hence, these are cohorts who completed most of their educational experience prior to the expulsion. The figure shows vividly that - despite living in agriculturally specialized countries in West Germany - refugees are much less likely to work in the agricultural sector.

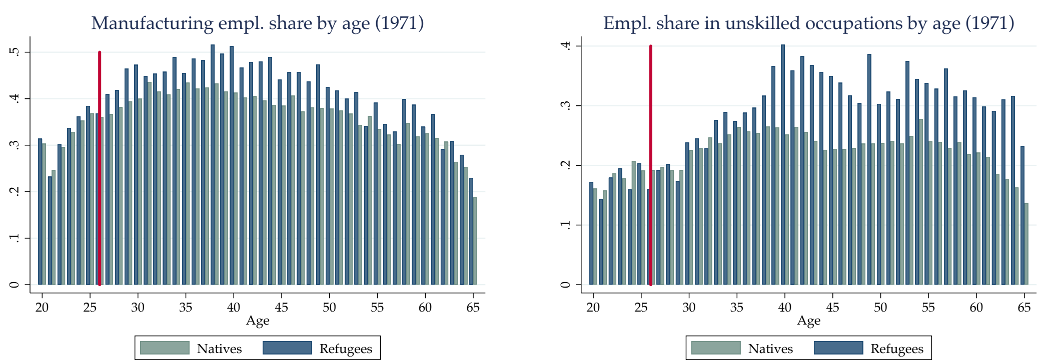
How then did the younger generation among the refugees react to the expulsion experience? In Figure OA-6 I display the age-profile of manufacturing employment shares (left panel) and employment shares in unskilled occupations (right panel). I also superimpose a red line at age 26, which is the cohort of individuals born in 1945, i.e. after the expulsion. Figure OA-6 suggests that cohort effects were very important for the economic integration of refugees. While older refugees are much more likely to work in the manufacturing sector and in unskilled occupations, this is not true for the younger cohorts. Among the individuals who grew up in West Germany, the differences between natives and refugees are hardly noticeable.

This pattern of a relatively successful integration of younger cohorts is also apparent in self-reported assessments about individuals' social mobility. More specifically, the supplement of the 1971 Census asked



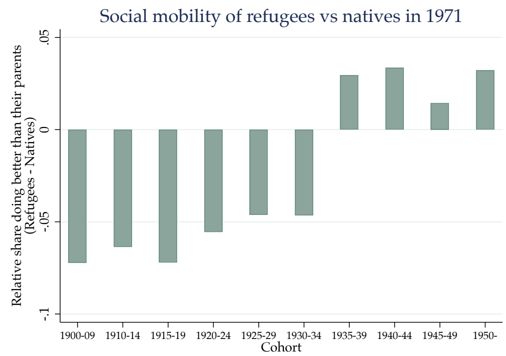
Notes: The figure shows the agricultural employment share in 1971 for different age groups for both natives and refugees

Figure OA-5: Agricultural Employment Shares by Age



Notes: The figure shows the employment share in manufacturing (left panel) and in unskilled occupations (right panel) in 1971 for different age groups for both natives and refugees.

Figure OA-6: Employment Shares in Manufacturing and Unskilled Occupations



Notes: The figure shows the relative social mobility of refugees relative to natives. Let p_a^{Ref} and p_a^{Nat} the share of refugees and natives of age a who report doing economically better than their parents. In the figure I report $\Delta_a = p_a^{Ref} - p_a^{Nat}$ for different cohorts.

Figure OA-7: Self-reported Social Mobility: Refugees vs Natives

	log income		
Refugee	-0.075***	-0.083***	-0.081***
	(0.007)	(0.008)	(0.008)
Refugee × Young		0.067***	0.052***
		(0.019)	(0.020)
State FE	✓	✓	✓
Demographics	✓	✓	✓
Cohort FE			✓
N	23831	23831	23831
R^2	0.111	0.111	0.112

Note: Robust standard errors in parentheses. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. "Young" is an indictor for whether the individual is below 30 years old in 1961. All specifications control for age, the squared of age and male dummy ("Demographics") and state fixed effects. Column 4 controls for fixed effects for 4 cohorts (below 30 years, 30-40, 40-50 and above 50 years).

Table OA-6: Relative Earnings by Cohort: Refugees vs Natives

individuals if they felt they were doing better than their fathers. In Figure OA-7 I report this measure of social mobility of refugees relative to natives. More specifically, let p_c^{Ref} and p_c^{Nat} be the share of refugees and natives among cohort c who consider themselves doing economically better than their parents. In Figure OA-7 I report $\Delta_a = p_a^{Ref} - p_a^{Nat}$ for different cohorts. The pattern of catch-up is clearly apparent. For all cohorts born prior to 1930, refugees experience *less* social mobility relative to natives, i.e. the share of people who consider themselves doing worse than their parents is higher among refugees. For the later cohorts this pattern flips. Precisely because the older generation of refugees experienced a stark break in their professional careers, their children see themselves doing relatively better. This finding is qualitatively consistent with the results reported in Becker et al. (2020), who, using data from Poland, argue that families with an experience of forced migration invest more in human capital.

Finally, in Table OA-6 I show that this pattern of perceived social mobility is also apparent in the micro data on earnings taken from the EVS 62 data. More specifically I regress log earnings on let the dummy on the refugee status of the individual differ by age. I define "young" individual as being younger than 30 in 1962, i.e. born after 1932. In column 1 I replicate for comparison the baseline earnings difference between refugees and natives that I target in the quantitative model. Columns 2 and 3 show that this "refugee discount" is much lower (and in fact not statistically significant) among younger refugees.

	Population growth 39-50	Change in ... share Manufac 39-50	Agric. 33-50	Service 33-50	GDP pc growth 35-50	Plant growth 33-50
Share of refugees in 1950	1.286***	0.322***	-0.228***	-0.059	-0.017	0.653**
	(0.137)	(0.075)	(0.062)	(0.055)	(0.369)	(0.283)
ln pop dens. 1939	-0.132***	0.001	-0.087***	0.044***	0.433***	0.745***
	(0.038)	(0.005)	(0.017)	(0.013)	(0.093)	(0.054)
ln pop dens. 1933	0.088**	0.007	0.008	-0.005	-0.219**	-0.734***
	(0.034)	(0.005)	(0.020)	(0.013)	(0.089)	(0.061)
Share of housing stock damaged	-0.235***	0.026	0.027	-0.079***	0.001	0.165
	(0.031)	(0.019)	(0.030)	(0.025)	(0.148)	(0.112)
Manufacturing share in 1939	0.208**	-0.177**	-0.486***	-0.169**	0.108	-0.207
	(0.081)	(0.078)	(0.085)	(0.068)	(0.325)	(0.151)
Agricultural share in 1933	-0.037	-0.079	-0.872***	0.022	-0.007	-0.361**
	(0.069)	(0.051)	(0.079)	(0.062)	(0.294)	(0.142)
Service share in 1933	-0.112	-0.096	-0.271	-0.534**	-0.671	-0.460
	(0.162)	(0.216)	(0.165)	(0.203)	(0.688)	(0.305)
ln y_{1935}	0.025	-0.013***	-0.023***	0.023***	-0.583***	-0.011
	(0.015)	(0.004)	(0.006)	(0.005)	(0.074)	(0.035)
ln num of manufac. plants (1933)	-0.032***	-0.005	0.009	-0.008*	0.070	-0.182***
	(0.010)	(0.003)	(0.006)	(0.004)	(0.047)	(0.045)
lndistance	-0.007	0.018	-0.018	0.011	-0.000	0.098
	(0.031)	(0.018)	(0.035)	(0.035)	(0.098)	(0.095)
Border with East Germany	-0.095	-0.005	0.013	-0.006	0.099	-0.136**
	(0.057)	(0.008)	(0.015)	(0.013)	(0.083)	(0.051)
State FE	✓	✓	✓	✓	✓	✓
N	519	519	519	519	519	519
R^2	0.698	0.421	0.776	0.442	0.540	0.664

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke*. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The results correspond to specification 3 in Table OA-7.

Table OA-7: OLS Results in 1950: Coefficients on Remaining Covariates

OA-2.5 OLS Estimates: Additional Results

Table 6 in the main text only reported the coefficient on the share of refugees in 1950, which is the main coefficient of interest. In Tables OA-7 and OA-8 I report the coefficients on the remaining covariates. I focus on specification 3 for 1950 (see Table OA-7) and specification 7 for 1961 (see Table OA-8). Additionally, in Table OA-9 I replicate the estimates of Table 6 with robust instead of clustered standard errors. This strengthens the results because the clustered standard errors turn out to be generally larger.

OA-2.6 Nonlinear effects of population density

The nature of the allocation rule implies that refugees were settled to relatively rural, low population density locations. Even though all my results reported in Tables 6 and 7 control for the population density in 1939, one might still be concerned that counties that differed in their pre-war population density were on different trends and that such differences are not captured by the controls used in Tables 6 and 7. In Table OA-10 I therefore address this concern by controlling for pre-war population density and pre-war urbanization in very flexible way. In particular, I include fixed effects for 20 quantiles of the distribution of population density in 1939, population density in 1933 and the county-level share of urbanization, i.e. the share of population living large cities. Hence, this specification includes 60 fixed effects to control for population density and the urban configuration of the county. For parsimony I focus on the three outcomes I also use for my structural estimation: population growth, changes in the manufacturing employment share and income per capita. All specifications also include the controls present in my main specification, i.e. state FE, the extent of war-time destruction and the initial level of respective dependent variable. For each outcome I report both the short-run and the long-run effect. Table OA-10 shows that even this level of saturation leaves the relationship between refugee inflows and subsequent growth unchanged - the estimates are statistically indistinguishable from my

	Population growth 39-61	Change in ... share Manufac 39-61	Agric. 33-61	Service 33-61	GDP pc growth 35-61	Plant growth 33-56
Share of refugees in 1950	1.235***	0.244***	-0.151**	-0.054	0.658***	0.830*
	(0.227)	(0.087)	(0.060)	(0.068)	(0.210)	(0.456)
ln pop dens. 1939	-0.165**	-0.023***	-0.051***	0.039***	0.084***	0.784***
	(0.066)	(0.008)	(0.015)	(0.012)	(0.028)	(0.066)
ln pop dens. 1933	0.131*	0.008	0.001	-0.003	-0.044	-0.809***
	(0.067)	(0.008)	(0.017)	(0.012)	(0.029)	(0.067)
Share of housing stock damaged	-0.018	0.003	0.022	-0.049*	0.120*	-0.104
	(0.068)	(0.022)	(0.033)	(0.029)	(0.068)	(0.180)
Manufacturing share in 1939	0.686***	-0.128	-0.499***	-0.195***	0.576***	2.318***
	(0.176)	(0.077)	(0.080)	(0.069)	(0.132)	(0.396)
Agricultural share in 1933	0.053	-0.035	-0.895***	0.014	-0.127	-0.250
	(0.171)	(0.049)	(0.067)	(0.062)	(0.095)	(0.279)
Service share in 1933	0.260	0.002	-0.287*	-0.559**	-0.466**	0.235
	(0.315)	(0.217)	(0.152)	(0.217)	(0.225)	(0.790)
ln y_{1935}	0.053	-0.013**	-0.017**	0.020***	-0.904***	0.108**
	(0.032)	(0.006)	(0.007)	(0.006)	(0.019)	(0.047)
ln num of manufac. plants (1933)	-0.111	-0.007	0.012	-0.010*	0.028	-0.240***
	(0.071)	(0.007)	(0.009)	(0.005)	(0.028)	(0.074)
lndistance	0.094*	-0.010	-0.028	0.037	0.166***	-0.085
	(0.047)	(0.024)	(0.032)	(0.038)	(0.044)	(0.122)
Border with East Germany	-0.016	-0.010	0.006	0.011	0.036	-0.099
	(0.129)	(0.010)	(0.011)	(0.014)	(0.040)	(0.109)
State FE	✓	✓	✓	✓	✓	✓
N	472	519	519	519	515	519
R^2	0.286	0.357	0.818	0.283	0.906	0.617
✓						

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke*. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. The results correspond to specification 7 in Table OA-7.

Table OA-8: OLS Results in 1961: Coefficients on Remaining Covariates

	1939-1950				1939-1961			
Panel A: Population growth: $\ln L_{rt} - \ln L_{r1939}$								
Share of refugees in 1950	2.032***	1.255***	1.286***	1.367***	1.246***	1.147***	1.235***	1.339***
	(0.106)	(0.113)	(0.116)	(0.105)	(0.240)	(0.198)	(0.205)	(0.198)
N	536	523	519	472	488	475	472	472
R^2	0.610	0.683	0.698	0.732	0.173	0.175	0.283	0.338
Panel B: Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$								
	1939-1950				1939-1961			
Share of refugees in 1950	0.203***	0.317***	0.322***	0.353***	0.451***	0.241***	0.244***	0.255***
	(0.041)	(0.048)	(0.048)	(0.042)	(0.048)	(0.060)	(0.061)	(0.058)
N	535	535	519	472	535	535	519	472
R^2	0.301	0.393	0.423	0.539	0.230	0.351	0.357	0.424
Panel C: Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$								
	1933-1950				1933-1961			
Share of refugees in 1950	-0.454***	-0.186**	-0.228***	-0.423***	-0.716***	-0.097	-0.151**	-0.326***
	(0.110)	(0.080)	(0.073)	(0.060)	(0.122)	(0.082)	(0.071)	(0.062)
N	523	523	519	472	523	523	519	472
R^2	0.091	0.701	0.776	0.842	0.122	0.761	0.817	0.858
Panel D: Service employment: $\pi^S_{rt} - \pi^S_{r1933}$								
	1933-1950				1933-1961			
Share of refugees in 1950	-0.089	0.014	-0.059	0.051	-0.098*	0.017	-0.054	0.057
	(0.058)	(0.056)	(0.055)	(0.046)	(0.056)	(0.059)	(0.058)	(0.054)
N	523	523	519	472	523	523	519	472
R^2	0.211	0.359	0.441	0.602	0.053	0.172	0.276	0.448
Panel E: GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$								
	1935-1950				1935-1961			
Share of refugees in 1950	-1.219***	-0.083	-0.017	-0.017	1.159***	0.502***	0.658***	0.746***
	(0.317)	(0.312)	(0.315)	(0.279)	(0.391)	(0.190)	(0.179)	(0.185)
N	523	523	519	472	519	519	515	468
R^2	0.110	0.511	0.540	0.582	0.101	0.889	0.905	0.903
Panel F: Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$								
	1933-1950				1933-1956			
Share of refugees in 1950	-0.450	0.726**	0.653***	0.817***	-0.819	0.697	0.830*	1.169***
	(0.306)	(0.296)	(0.198)	(0.200)	(0.547)	(0.542)	(0.423)	(0.416)
N	520	520	519	472	520	520	519	472
R^2	0.045	0.393	0.664	0.680	0.140	0.372	0.617	0.626
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Pop. density (1939)		✓	✓	✓		✓	✓	✓
Wartime destr.		✓	✓	✓		✓	✓	✓
Geography		✓	✓	✓		✓	✓	✓
Levels of dep. variable		✓	✓	✓		✓	✓	✓
Pre-war controls			✓	✓			✓	✓
Addtl. pre-war controls				✓				✓

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

Table OA-9: OLS Results (Table 6): Estimates with Robust Standard Errors

	Pop. growth		Manufac. growth		Income growth	
	1939-50	1939-61	1939-50	1939-61	1935-50	1935-61
Share of refugees in 1950	1.435***	1.175***	0.320***	0.247***	0.012	0.677***
	(0.103)	(0.180)	(0.072)	(0.089)	(0.380)	(0.180)
State FE	✓	✓	✓	✓	✓	✓
ln pop density 1939	✓	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓	✓
Lagged dep. variable	✓	✓	✓	✓	✓	✓
FE for 1939 pop density	✓	✓	✓	✓	✓	✓
FE for 1933 pop density	✓	✓	✓	✓	✓	✓
FE for 1939 urbanization	✓	✓	✓	✓	✓	✓
N	513	513	522	522	523	519
R^2	0.855	0.530	0.443	0.420	0.593	0.917

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke*. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. All specifications control for state fixed effects, population density in 1939, the share of wartime destruction, the distance to the inner-german border, a dummy for whether a county is directly at the border and 20 fixed effects for the respective 5% quantiles of population density in 1933, population density in 1939 and the share of the population living in cities with less than 2000 people.

Table OA-10: The Role of Pre-war Population Density and Urbanization

baseline results reported in Table 6.

OA-2.7 An Alternative IV strategy

This sections contains the details of an alternative IV strategy to estimate (2). Table 5 showed that the local housing supply as measured by the local population density and the extent or war-time destruction was an important determinant of the inflow of refugees. The share of refugees is also correlated with another aspect of the local housing supply: the average number of rooms. The reason why the number of rooms within natives' houses was 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 apartment, this was less the case if only 2 bedrooms were available. 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, I expect the *interaction* between the average house size and the distance to the expulsion regions to be a predictor of refugee flows.

In OA-11 I show this is the case. Columns 1 and 2 shows that the average number of rooms positively predicts the share of refugees. In column 3 I restrict the sample to counties, which are geographically close to the expulsion regions. Specifically, I consider all counties, whose expulsion distance is below the median of the observed distribution. For these counties, average house size remains an important predictor of refugees inflows. In contrast, column 4 shows that for counties far away from the expulsion regions (i.e. whose distance exceeds the median distance) the local housing stock is not a predictor of refugee inflows. Hence, the existence of marginal rooms is particularly important in counties, which are close to the expulsion regions and which therefore receive a large refugee inflow. In the last column I focus on this interaction effect directly: larger houses as measured by the average room size or the share of apartment with more than four rooms, are better predictors for refugee inflows the closer the county is to the expulsion regions. However, this differential trend is only significant just below the 10% level.

These patterns can be used as an alternative instrument variable strategy. 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 regions and average room size is a valid instrument for the

	Full Sample		low ED	large ED	Full Sample
Avg num of rooms (rooms/apt)	0.034***	0.039***	0.033***	0.010	1.019*
	(0.010)	(0.010)	(0.011)	(0.009)	(0.585)
Avg rooms * Exp. Dist.					-0.074
					(0.044)
State FE	✓	✓	✓	✓	✓
ln Expulsion Distance		✓	✓	✓	✓
ln pop dens 1939	✓	✓	✓	✓	✓
Wartime destr.	✓	✓	✓	✓	✓
Geography	✓	✓	✓	✓	✓
Prewar Industrial Structure	✓	✓	✓	✓	✓
N	522	522	267	255	522
R^2	0.797	0.805	0.705	0.832	0.809

Note: Standard errors are clustered at the level of 37 *Regierungsbezirke*. *, ** 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". Column 4 (5), indicated by "low ED" ("large ED") conditions on counties, whose expulsion distance is below (higher than) the median of the expulsion distance distribution. All specification control for state fixed effect, log population density in 1939, the extent of wartime destruction, the manufacturing employment share in 1939 and 1933, the log distance of the inner german border and a dummy if a county is at the inner german border.

Table OA-11: The Allocation of Refugees, the Expulsion Distance and the Size of Houses

allocation of refugees. As for the baseline specification I rely on the variation within states, i.e. I instrument the share of refugees with the interactions of average room size, the distance to the expulsion regions and state fixed effects. In Table OA-12 I report the results of this IV strategy. The table has the same structure as the baseline results in Table 7, i.e. I consider all six outcomes, both in the short- and the long-run for a variety of specifications. While the results are noticeably less precisely estimated, this alternative strategy delivers roughly the same qualitative patterns as local growth in population, manufacturing employment, GDPpc and the number of plants is increasing in the allocation of refugees. In contrast to the main results, this strategy points to a reallocation between manufacturing and services rather then agriculture.

OA-2.8 Spatial Sorting: Additional evidence

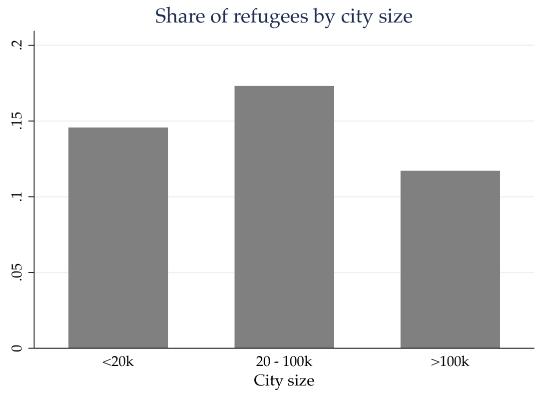
In Table 12 I presented regression evidence that suggested that the share of industrial workers among refugees exceeded the ones of natives in rural areas. The EVS micro data offer another way to look at the implications of spatial sorting. Recall that the EVS data stems from the year 1962. While this data does not contain identifiers for the specific county individuals live in, they do report the size of the respondent's city. In the left panel of Figure OA-8 I report the share of refugee by the size of the city. The spatial distribution of refugees is tilted towards cities in the range of 20-100.000 people. This pattern that refugees leave the very small cities is consistent with the analysis of the Bavarian village-level data, which is contained Section SM-2.2.3 in the Appendix. More interestingly, in the right panel I study differences in the sectoral employment share between natives and refugees in cities of different size. Specifically, I show $\pi^{Natives}_{sector,city} - \pi^{Refugee}_{sector,city}$ for the three sectors and the three city size categories. Refugees have a very large comparative advantage in the manufacturing sector in the smallest locations but sectoral employment shares are close to equalized in the larger locations. This pattern is consistent with the results in Table 12 that the initial refugee inflow was particularly biased towards the manufacturing in the most rural locations.

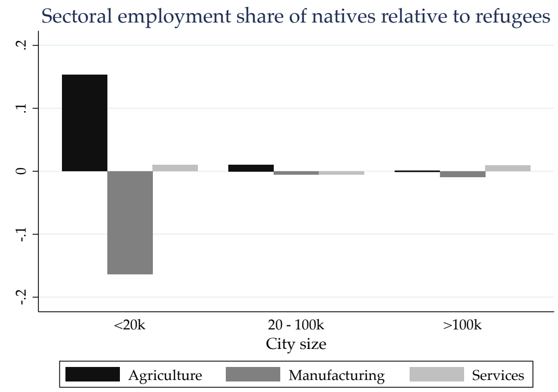
Similarly, one can look at the earnings data. If natives are indeed better spatially sorted than natives, the refugee "discount" should be higher in larger locations that have a comparative advantage in the non-agricultural sector. In Table OA-13 I report the results of a cross-sectional regression of log wages on a refugee dummy interacted with a dummy for the city size categories. The first column shows that refugees earn – on average – 7.5% less than natives. This is the overall native premium that is targeted in the main text. In column 3 I now let this effect vary by city size and also control for the main effect of city size. Refugees earn

	Panel A: Population growth: $\ln L_{rt} - \ln L_{r1939}$							
	1939-1950				1939-1961			
Share of refugees in 1950	1.900***	1.500***	1.195***	1.397***	1.059***	0.936**	0.382	1.074***
	(0.310)	(0.220)	(0.218)	(0.225)	(0.388)	(0.435)	(0.353)	(0.330)
N	526	526	509	463	526	526	509	463
F-Stat	314.434	18.190	13.102	10.811	291.597	16.855	12.593	10.560
	Panel B: Manufacturing employment: $\pi^M_{rt} - \pi^M_{r1939}$							
	1939-1950				1939-1961			
Share of refugees in 1950	0.208	0.398**	0.435***	0.459***	0.112	0.119	0.149	0.108
	(0.163)	(0.180)	(0.160)	(0.116)	(0.140)	(0.189)	(0.150)	(0.124)
N	535	535	519	472	535	535	519	472
F-Stat	300.672	20.505	16.034	11.409	300.672	20.505	16.034	11.409
	Panel C: Agricultural employment: $\pi^A_{rt} - \pi^A_{r1933}$							
	1933-1950				1933-1961			
Share of refugees in 1950	-0.243	0.011	0.432	-0.116	-0.352	0.126	0.596*	0.062
	(0.353)	(0.349)	(0.343)	(0.313)	(0.443)	(0.338)	(0.340)	(0.312)
N	523	523	519	472	523	523	519	472
F-Stat	292.630	28.035	16.034	11.409	292.630	28.035	16.034	11.409
	Panel D: Service employment: $\pi^S_{rt} - \pi^S_{r1933}$							
	1933-1950				1933-1961			
Share of refugees in 1950	-0.258	-0.687***	-0.690***	-0.409*	0.036	-0.377	-0.432*	-0.073
	(0.247)	(0.231)	(0.226)	(0.224)	(0.326)	(0.291)	(0.256)	(0.234)
N	523	523	519	472	523	523	519	472
F-Stat	292.630	26.116	16.034	11.409	292.630	26.116	16.034	11.409
	Panel E: GDP per capita growth: $\ln y_{rt} - \ln y_{r1935}$							
	1935-1950				1935-1961			
Share of refugees in 1950	-1.214	-0.185	-1.149	-0.366	0.286	1.490*	0.746	1.678**
	(0.991)	(1.536)	(1.360)	(1.137)	(1.761)	(0.840)	(0.723)	(0.695)
N	523	523	519	472	519	519	515	468
F-Stat	292.630	21.928	16.034	11.409	332.517	22.248	16.455	12.058
	Panel F: Growth of industrial plants: $\ln N_{rt} - \ln N_{r1933}$							
	1933-1950				1933-1956			
Share of refugees in 1950	0.539	1.090	1.679*	2.001**	2.504	5.255**	3.679**	4.469***
	(1.201)	(0.737)	(1.007)	(0.955)	(2.704)	(2.335)	(1.656)	(1.503)
N	520	520	519	472	520	520	519	472
F-Stat	289.260	21.302	16.034	11.409	289.260	21.302	16.034	11.409
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Pop. density (1939)		✓	✓	✓		✓	✓	✓
Wartime destr.		✓	✓	✓		✓	✓	✓
Geography		✓	✓	✓		✓	✓	✓
Levels of dep. variable		✓	✓	✓		✓	✓	✓
Pre-war controls			✓	✓			✓	✓
Addtl. pre-war controls				✓				✓

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

Table OA-12: The Effects of Refugee Inflows on the Local Economy: An Alternative IV Strategy





Note: The left panel shows the share of refugees by three city size categories in 1962. The right panel shows the relative sectoral employment share $\pi_{rst}^{Ref} - \pi_{rst}^{Nat}$.

Figure OA-8: Refugees' relative sectoral employment shares by city size

relatively less in larger cities but the effect is noisy. In columns 4 and 5 I show the same specification when I control for a set of six occupation fixed effects. This not only reduces the refugee effect but within occupations, such lower relative earnings accrue entirely in larger cities. Hence, at least qualitatively, the results in Table OA-13 and Figure OA-8 are consistent with the mechanism of spatial sorting highlighted in my theory.

OA-2.9 Persistent Effects of the Historical Allocation Rule: Additional Results

In Section 5.5 I showed that the policy of sending the refugees to predominantly rural counties played an important role for the industrialization process of these communities. If the initial refugee share had been equalized, the expansion of the manufacturing sector in rural communities had been much less pronounced. In this section I offer a more detailed analysis of these two different policies.

Consider first Figure OA-9, where I show relative population growth (top row) and income growth (bottom row) as a function of pre-war population density for both the actual allocation (orange) and a counterfactual policy which equalized the local refugee share (blue). All outcomes are relative to a situation where no refugees arrived but the sequence of productivity shocks $\mathcal{Q}_{rt}$ is the same as with the refugee "treatment".

The top row of Figure OA-9 shows that the allocation of refugees was an important source of population growth for rural labor markets, both in 1961 and 2000. If the share of refugees had been equalized initially, population growth had been much more balanced. The bottom row of Figure OA-9 shows the relative growth in income per capita. Interestingly, and consistent with the findings for local industrialization shown in Figure 7 in the main text, the inflow of refugees would have been pro-rural even in the absence of the pro-rural initial allocation policy. However, this pattern is even more pronounced under the actual allocation. Also note that under both allocation rules some regions with relatively high density (in the range around 7) were large beneficiaries of the refugee inflow, in particular under the actual allocation rule. These regions are - on average - close to the german border and hence received a large inflow despite their relatively large density. Without the refugee inflow they would have been adversely affected from the loss in market access.

OA-3 Computational Implementation

OA-3.1 Computational Algorithm

In this section I gather all equations and parameters to implement the theory computationally.

Figures/Allocation_cf_dpop.eps

Figures/Allocation_cf_dpop_lr.eps

Figures/Allocation_cf_dlny.eps

Figures/Allocation_cf_dlny_lr.eps

Note: The figure shows local population growth (top row) and local income pc growth (bottom row), relative to an allocation without refugee inflows, for the historical allocation (orange) and a counterfactual allocation policy that equalizes the initial share of refugees across counties.

Figure OA-9: Persistent Effects of the Refugee Settlement: Rural Income and Population Growth

	ln income				
Refugee	-0.075***	-0.082***	-0.073***	-0.030***	-0.004
	(0.007)	(0.007)	(0.011)	(0.007)	(0.010)
Refugee × Medium city			-0.020		-0.045**
			(0.020)		(0.019)
Refugee × Large city			-0.013		-0.047***
			(0.016)		(0.015)
Demographics	✓	✓	✓	✓	✓
State FE	✓	✓	✓	✓	✓
Occupation FE				✓	✓
Citytype FE		✓	✓	✓	✓
N	23831	23822	23822	23822	23822
R^2	0.111	0.124	0.124	0.251	0.251

Note: Robust standard errors in parenthesis. *, ** and *** denote statistical significance at the 10%, 5% and 1% level respectively. All specifications control for state fixed effect and demographics (age, age squared and a male dummy). Specifications with city type fixed effects control for 5 fixed effects of different size categories (<2k, 2-20k, 20-100k, 100-500k, >500k) and 5 fixed effects the type of the city (urban center, urban fringe, industrial zone, agricultural zone and mixed). Specifications with occupation fixed effects control for 10 occupation fixed effects.

Table OA-13: Refugee earnings by city size

1. **Exogenous regional fundamentals**: In our theory, locations are characterized by land supplies, amenities and innate productivities
$$\{T_r, V_r, Q_{rt}\}_{r=1}^{R}. \tag{OA-12}$$
While land supplies and amenities are assumed to be constant, innate productivity evolves according to
$$\ln Q_{rt} = (1-\varrho)\ln Q_r + \varrho \ln Q_{rt-1} + \varpi u_{rt}, \tag{OA-13}$$
where $u_{rt} \sim \mathcal{N}(0,1)$.

2. **Spatial labor supply:** Let L_{rt-1}^{Nv} and L_{rt-1}^{Rv} denote the number of natives and refugees of type $v = I, R$. The total supply of workers of type v in region r at time t is given as
$$L_{rt}^{\nu} = (1-\psi) L_{rt-1}^{\nu} + \psi \sum_{j=1}^{R} L_{jt-1}^{v} m_{jrt}^{v}, \tag{OA-14}$$
where $L_{rt}^{\nu} = L_{rt}^{Nv} + L_{rt}^{Rv}$ and
$$m_{jrt}^{v} = \frac{\left(V_r \mu_{jr} \overline{u}_{rt}^{\nu}\right)^{\varepsilon}}{\sum_k \left(V_k \mu_{jk} \overline{u}_{rt}^{\nu}\right)^{\varepsilon}}. \tag{OA-15}$$

3. **Sectoral labor supply:** Given, L_{rt}^{ν}, total labor supply in sectors A and M are given by
$$H_{rAt} = \Gamma_\theta \sum_{v=I,R} L_{rt}^{\nu} \phi_A^v \left(\frac{w_{rAt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1} \tag{OA-16}$$
$$H_{rMt} = \Gamma_\theta \sum_{v=I,R} L_{rt}^{\nu} \phi_M^v \left(\frac{w_{rMt}}{\overline{w}_{rt}^{\nu}}\right)^{\theta-1}, \tag{OA-17}$$
where $\overline{w}_{rt}^{\nu}$ is given as
$$\overline{w}_{rt}^{\nu} = \left(\phi_A^v w_{rAt}^{\theta} + \phi_M^v w_{rMt}^{\theta}\right)^{1/\theta}. \tag{OA-18}$$

and $\Gamma_\theta = \Gamma(1 - 1/\theta)$ and $\Gamma(.)$ is the gamma function.

4. **Goods market clearing:** The two goods market clearing conditions are given as

$$\frac{\rho}{\rho-1} w_{rMt} H_{rPt} = (1-\alpha) \sum_{j=1}^{R} \left(\frac{P_{rjM}}{P_{jM}}\right)^{1-\sigma} \left(\frac{w_{jAt} H_{jAt}}{\gamma} + \frac{\rho}{\rho-1} w_{jMt} H_{jPt}\right) \quad \text{(OA-19)}$$

$$\frac{w_{rAt} H_{rAt}}{\gamma} = \alpha \sum_{j=1}^{R} \left(\frac{P_{rjA}}{P_{jA}}\right)^{1-\sigma} \left(\frac{w_{jAt} H_{jAt}}{\gamma} + \frac{\rho}{\rho-1} w_{jMt} H_{jPt}\right), \quad \text{(OA-20)}$$

where $P_{rjM} = \frac{\tau_{rj} w_{rMt}}{Q_{rt} N_{rt-1}^{\lambda\vartheta} H_{rPt}^{\vartheta}}$, $P_{rjA} = \frac{\tau_{rj} w_{rAt}}{Q_{rt}} \frac{H_{rAt}}{T_r}$ and $P_{jM} = \left(\sum_r P_{rjM}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ and $P_{jA} = \left(\sum_r P_{rjA}^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ and the short-run scale elasticity is $\vartheta = \frac{1}{\rho-1}$.

5. **Labor market clearing:** The labor market clearing condition for the manufacturing sector is given by

$$H_{rMt} = \frac{\rho}{\rho-1} H_{rPt} - (1-\delta) f_E N_{rt-1}^{1-\lambda}. \quad \text{(OA-21)}$$

6. **Dynamic evolution of state variable:** The dynamic state variable, the number of firms N_{rt}, evolves according to

$$N_{rt} = \frac{1}{\rho-1} \frac{1}{f_E} H_{rPt} \times N_{rt-1}^{\lambda}. \quad \text{(OA-22)}$$

7. **Initial conditions:** The initial conditions of the endogenous state variables are given by the number of plants N_{r0} and the labor supply by natives and refugees from either type

$$\left\{N_{r0}, L_{r0}^{NI}, L_{r0}^{NR}, L_{r0}^{RI}, L_{r0}^{RR}\right\}_{r=1}^{R}. \quad \text{(OA-23)}$$

These equations fully characterize the dynamic equilibrium. To see this, let the initial conditions in (OA-23) and the exogenous fundamentals in (OA-12) and (OA-13) be given. Let a vector of wages $[w_{rA1}, w_{rM1}]_r$ be given. Note that the system is homogenous of degree zero in the level of wages, because we did not pick a numeraire yet. Let us use the numeraire

$$w_{1M1} = 1$$

for an arbitrarily selected region 1. From (OA-14) and (OA-15) we can calculate $\left[L_{r1}^{I}, L_{r1}^{NI}\right]$. From (OA-16), (OA-17) and (OA-18) we can calculate $[H_{rAt}, H_{rMt}]$. From (OA-21) we can calculate $[H_{rP1}]$. Then we still have (OA-19) and (OA-20), which are $2 \times R$ equations, to solve for $[w_{rA1}, w_{rM1}]_r$. Given the equilibrium values for $[w_{rA1}, w_{rM1}]_r$, we can then use (OA-22) to calculate $[N_{r1}]$. This determines the set of new state variables $\left\{N_{r1}, L_{r1}^{I}, L_{r1}^{R}\right\}_{r=1}^{R}$ and we can solve for the equilibrium in the next period. Note that the equilibrium requires that the free entry condition is satisfied, i.e. that net investment in firms is positive.

$$N_{rt} > (1-\delta) N_{rt-1},$$

Using (OA-22) this implies $\frac{1}{(\rho-1)f_E} H_{rPt} > (1-\delta) N_{rt-1}^{1-\lambda}$.

OA-3.2 Inferring Fundamentals

Step 1: Infer equilibrium wages $[w_{rA}, w_{rM}]_r$ and prices $[P_{rrA}, P_{rrM}]_r$ Let the structural parameters be given. We can then infer $[w_{rA}, w_{rM}]_r$ from the following iterative procedure

1. Guess average for each type in region r, $\left[\overline{w}_r^F, \overline{w}_r^I\right]_r$ and the aggregate price index for each r, $\Psi_r \equiv P_{rA}^{\alpha} P_{rM}^{1-\alpha}$. Note that we guess directly Ψ_r, not the individual P_{rA} and P_{rM}. Set $\Psi_1 = 1$.

2. Given $\left[\overline{w}_r^F, \overline{w}_r^I, \Psi_r\right]_r$ solve for V_r and the share of high types across space ω_r^I. To do so we can use the following procedure

   (a) Guess $\left[\omega_r^I\right]_r$, i.e. the share of high types in region r

(b) Let $l_r = \frac{L_r}{L}$ be the observed population share. Then calculate the number of industrialists $L_r^I = \omega_r^I l_r$.

(c) Given L_r^I solve the amenities from the steady state equations

$$L_r^I = \sum_k \frac{\left(\mu_{kr} V_r \frac{\overline{w}_{rt}^I}{\Psi_r}\right)^\varepsilon}{\sum_j \left(\mu_{kj} V_j \frac{\overline{w}_{jt}^I}{\Psi_j}\right)^\varepsilon} L_k^I. \tag{OA-24}$$

To do so,

i. Guess $[V_r]$ and normalize $\left(\sum V_r^\varepsilon\right)^{1/\varepsilon} = 1$

ii. Then use (OA-24) to solve for V_r as

$$V_r = \left(\frac{L_r^I}{\sum_k \frac{\left(\mu_{kr} \frac{\overline{w}_{rt}^\nu}{\Psi_r}\right)^\varepsilon}{\sum_j \left(\mu_{kj} V_j \frac{\overline{w}_{jt}^\nu}{\Psi_j}\right)^\varepsilon} L_k^I}\right)^{1/\varepsilon} \tag{OA-25}$$

Hence, given a guess $\left[V_r^0\right]$, we can use (OA-25) to update $\left[V_r^1\right]_r$ and then iterate on (OA-25) until $\left[V_r^{n+1}\right]_r \approx \left[V_r^n\right]_r$. Note that whenever we use (OA-25) to update $\left[V_r^1\right]_r$ we need to impose $\left(\sum V_r^\varepsilon\right)^{1/\varepsilon} = 1$.

(d) Once we found a converged $[V_r]$, we can use the spatial labor supply equation for the famers to get an update for our guess ω_r^I. In particular, Given our guess ω_r^I we can calculate the number of famers in location r as

$$L_r^F = \sum_k \frac{\left(\mu_{kr} V_r \frac{\overline{w}_{rt}^F}{\Psi_r}\right)^\varepsilon}{\sum_j \left(\mu_{kj} V_j \frac{\overline{w}_{jt}^F}{\Psi_j}\right)^\varepsilon} \left(1 - \omega_k^I\right) l_k$$

Given L_r^F we can update our guess for ω_r^I as $\tilde{\omega}_r^I = \max\left\{1 - \frac{L_r^F}{l_r}, 0\right\}$Note that we should impose the max operator as there is no guarantee that $L_r^F < l_r$ for a given V_r.

(e) If $\tilde{\omega}_r^I \approx \omega_r^I$, stop as we have converged. Otherwise go back to (a) with $\tilde{\omega}_r^I$.

3. Given $\left[\overline{w}_r^F, \overline{w}_r^I, \Psi_r\right]_r$ and $\left[\omega_r^I\right]$ we can calculate the sectoral skill prices w_{rA} and w_{rM} from the same equations as in our current procedure. In particular,

(a) calculate w_{rA} from agricultural employment share s_{rA} as

$$w_{rA} = \left(\frac{s_{A,r}}{\sum_\nu \omega_r^\nu \phi_A^\nu \left(\frac{1}{\overline{w}_r^\nu}\right)^\theta}\right)^{1/\theta}, \tag{OA-26}$$

(b) Calculate w_{rM} from

$$py_r = \Gamma_\theta \left(\frac{1}{1 - \frac{(1-\delta)(1-\beta)}{\rho}} \sum_{\nu=I,F} \omega_r^\nu \left(\phi_A^\nu w_{rA}^\theta + \phi_M^\nu w_{rM}^\theta\right)^{1/\theta} \left(1 - \pi_{rA}^\nu\right) + \frac{1}{\gamma} \sum_{v=I,F} \omega_r^\nu \left(\phi_A^\nu w_{rAt}^\theta + \phi_M^\nu w_{rM}^\theta\right)^{1/\theta} \pi_{rA}^\nu\right),$$

where π_{Ar}^ν is the employment share in agriculture of type ν and ω_r^ν is the share of ν-types and py_r denotes income per capita in region r

4. Given (w_{rM}, w_{rA}) we can calculate the implied average income by type $\overline{w}_r^{\nu,impl} = \left(\phi_A^v w_{rA}^\theta + \phi_M^v w_{rM}^\theta\right)^{1/\theta}$ and calculate an updated guess for $\left[\overline{w}_r^R, \overline{w}_r^I\right]_r$ as $\overline{w}_r^{\nu,new} = \zeta \overline{w}_r^\nu + (1-\zeta)\, \overline{w}_r^{\nu,impl}$, where $\zeta \in (0,1)$ is a weight to update our wage guess and iterate on $\overline{w}_r^\nu$ until we find a fixed point.

5. Given (w_{rM}, w_{rA}) we can calculate H_{rM} and H_{rA}. Given H_{rM} we can calculate the number of production workers as

$$H_{rP} = \frac{\rho-1}{\rho-1+\delta} H_{rM}.$$

6. Then we can calculate the parameter α, which is consistent with the data as

$$\alpha = \frac{\sum_r \frac{w_{rA} H_{rA}}{\gamma}}{\sum_j \left(\frac{w_{jA} H_{jA}}{\gamma} + \frac{\rho}{\rho-1} w_{jM} H_{jP}\right)}$$

Note that the RHS is known.

7. Then we can determine the regional prices P_{rrM} and P_{rrA}, i.e. the price of region r goods in region r. In particular, market clearing for the manufacturing sector requires that

$$P_{rrM} = \left(\frac{\rho-1}{\rho} \frac{1}{w_{rM} H_{rP}}\right)^{\frac{1}{\sigma-1}} \left(\sum_j \frac{\tau_{rj}^{1-\sigma} (1-\alpha) \left(\frac{w_{jA} H_{jA}}{\gamma} + \frac{\rho}{\rho-1} w_{jM} H_{jP}\right)}{\sum_m \tau_{mj}^{1-\sigma} P_{mmM}^{1-\sigma}}\right)^{\frac{1}{\sigma-1}} \tag{OA-27}$$

Note again that P_{rrM} are only identified up to scale. Hence, we can solve for P_{rrM} by the following procedure

   (a) Guess $\left[P_{rrM}^0\right]_{r=1}^R$ and impose that $P_{11M}^0 = 1$

   (b) Plug $\left[P_{rrM}^0\right]_{r=1}^R$ in the RHS of (OA-27) to get P_{rrM}^1 from the LHS

   (c) Check if $P_{rrM}^1 \approx P_{rrM}^0$. If yes, stop. If no, take $\left[P_{rrM}^1\right]_{r=1}^R$ as the new guess and impose that $P_{11M}^1 = 1$ (i.e. divide P_{rrM}^1 by P_{11M}^1) and go back to (b).

We can use the exact same procedure to solve for $[P_{rrA}]_{r=1}^R$ from the equation, which is implied by (SM-3)

$$P_{rrA} = \left(\frac{\gamma}{w_{rA} H_{rA}} \sum_j \frac{\tau_{rj}^{1-\sigma} \alpha \left(\frac{w_{jA} H_{jA}}{\gamma} + \frac{\rho}{\rho-1} w_{jM} H_{jP}\right)}{\sum_m \tau_{mj}^{1-\sigma} P_{mmA}^{1-\sigma}}\right)^{\frac{1}{\sigma-1}}$$

8. Given $[P_{rrA}]_{r=1}^R$ and $[P_{rrM}]_{r=1}^R$ we can calculate the implied $\tilde{\Psi}_r = P_{rA}^\alpha P_{rM}^{1-\alpha}$. Note that $\tilde{\Psi}_1 = 1$ by construction. If $\tilde{\Psi}_r \approx \Psi_r$ we are done and found a solution. If $\tilde{\Psi}_r \neq \Psi_r$ we can update $\Psi_r^{new} = \zeta \Psi_r + (1-\zeta)\, \tilde{\Psi}_r$for some ζ

Step 2: Infer regional fundamentals Given $[P_{rrA}, P_{rrM}]_r$ inferring the actual fundamentals is trivial. For the manufacturing sector we get that $P_{rMt} = \frac{\rho}{\rho-1} \frac{1}{Q_{rt}} w_{rMt} N^{-\frac{1}{\rho-1}}$ and $N_{rt} = \left(\frac{1}{\rho-1} \frac{1}{f_E}\right)^{\frac{1}{1-\lambda}} H_{rPt}^{\frac{1}{1-\lambda}}$. Hence

$$Q_{rt} = \frac{\rho}{\rho-1} \frac{w_{rMt}}{P_{rMt}} \left(\frac{1}{\rho-1} \frac{1}{f_E}\right)^{-\frac{1}{\rho-1}\frac{1}{1-\lambda}} H_{rPt}^{-\frac{1}{\rho-1}\frac{1}{1-\lambda}}. \tag{OA-28}$$

For the agricultural sector we get that $P_{rAt} = \frac{w_{rAt}}{Q_{rt}} \left(\frac{H_{rAt}}{T_r}\right)^{1-\gamma}$, so that

$$T_r = \left(\frac{w_{rAt}}{Q_{rt} P_{rAt}}\right)^{\frac{1}{1-\gamma}} H_{rAt}. \tag{OA-29}$$

Figures/timing dynamics.pdf

Figure OA-10: Timing of Events

Hence, equations (OA-28) and (OA-29) contain expressions for T_r and $\mathcal{Q}_r$ in terms of observables.[48]

OA-4 A Dynamic Version of the Model

My analysis assumes that workers and firms behave myopically. This assumption simplifies the analysis, especially the structural estimation that hinges on the transitional dynamics of the system, considerably.

In this section I present a full dynamic version of my theory where I allow agents to be forward-looking. In Section OA-4 I characterize the dynamic equilibrium system. In Section OA-4.2 I discuss the main differences between the myopic model and the specification with forward-looking agents. I also characterize the steady state of the forward-looking model and show that the main theoretical implications are qualitatively similar to the model with myopic agents. In Section OA-4.3 I discuss how existing approaches in the literature address the problem of computational complexity with forward-looking agents and why these approaches are not applicable in my setting. Finally, in Section OA-4.4, I argue why my estimation strategy, which relies on indirect inference, might make the myopic approach less susceptible to the problem of model misspecification.

OA-4.1 The Model with Forward-Looking Agents

I now characterize the solution of the model when agents - both workers and firms - are forward-looking.

OA-4.1.1 Households

I follow Caliendo et al. (2019) and Desmet et al. (2018) and assume that households do not engage in intertemporal saving. Hence, the per-period expected consumption utility of an individual of type ν when residing in location r at time t before knowing the realization of her efficiency draws $z^i = \left(z_A^i, z_M^i\right)$ is given by

$$u_{rt}^{\nu} = \frac{\overline{w}_{rt}^{\nu}}{P_{rMt}^{1-\alpha} P_{rAt}^{\alpha}}, \tag{OA-30}$$

where, as before, $\overline{w}_{rt}^{\nu} = \left(\phi_A^{\nu} w_{rAt}^{\theta} + \phi_M^{\nu} w_{rMt}^{\theta}\right)^{1/\theta}$.

Given (OA-30), I can write the life-time utility of an individual of type ν residing in region r recursively. The timing of events is summarized in Figure OA-10. Let v_{rt}^{ν} be the value of type ν to reside in region r at

[48] Note also that a common level shift of GDPpc $[py_r]$ maps one-to-one in common level shifts of $[w_{rM}, w_{rA}]$. Suppose that $[w_{rM}, w_{rA}]$ solves the system above. Now suppose we scale the data on $[py_r]$ by a shifter κ. Scaling both $[w_{rM}, w_{rA}]$ by κ will also scale $\left[\overline{w}_r^I, \overline{w}_r^F\right]$ by κ. This will leave the inferred amenities $[V_r]$ and the sorting across space $[\omega_r^{\nu}]_r$ the same. It will also match the same agricultural shares $[s_{A,r}]$ as these only depend on $w_{rA}/\overline{w}_r^{\nu}$.

time t before the realization of her efficiency draw $z^i = \left(z_A^i, z_M^i\right)$. This value function can be written as

$$v_{rt}^{\nu} = u_{rt}^{\nu}\mathcal{A}_r + (1-\psi)\,\beta E_t\left[v_{rt+1}^{\nu}\right] + \psi\beta E_t\left[\int_{\bar{\xi}} \max_{j\in R}\left\{v_{jt+1}^{\nu} + \xi_j - \eta_{rj}\right\} dF\left(\bar{\xi}\right)\right], \tag{OA-31}$$

where - as before - ξ_j denotes the realization of the idiosyncratic shock for location j, $\bar{\xi}$ is the vector of shocks for all locations, η_{rj} denotes the moving costs from r to j and $\mathcal{A}_r$ denotes the amenities in region r. Note that v_{rt+1}^{ν} is a random variable as of time t because wages and prices are stochastic due to the presence of productivity shocks $\mathcal{Q}_{rt}$.

For simplicity I denominate both the moving costs and the idiosyncratic preference shocks in terms of per-period utility so that they enter additively. Under the assumption that ξ_j is distributed iid according to a Type-I Extreme value distribution with variance $1/\varepsilon$, we have that

$$\int_{\bar{\xi}} \max_{j\in R}\left\{v_{jt+1}^{\nu} + \xi_j - \eta_{rj}\right\} dF\left(\bar{\xi}\right) = \frac{1}{\varepsilon}\ln\left(\sum_{j=1}^{R}\exp\left(v_{jt+1}^{\nu} - \eta_{rj}\right)^{\varepsilon}\right). \tag{OA-32}$$

Hence, v_{rt}^{ν} solves the recursive equation

$$v_{rt}^{\nu} = u_{rt}^{\nu}\mathcal{A}_r + (1-\psi)\,\beta E_t\left[v_{rt+1}^{\nu}\right] + \psi\beta\frac{1}{\varepsilon}E_t\left[\ln\left(\sum_{j=1}^{R}\exp\left(v_{jt+1}^{\nu} - \eta_{rj}\right)^{\varepsilon}\right)\right]. \tag{OA-33}$$

For given parameters and a given stochastic process for utilities $\{u_{rt}^{\nu}\}_{rt\nu}$ (which vary across types ν and locations r) one can solve (OA-33) for $\{v_{rt}^{\nu}\}_{rt\nu}$.[49] Given $\{v_{rt}^{\nu}\}_{rt\nu}$ the moving probabilities between periods t and $t+1$ from r to j are given by

$$m_{jrt+1}^{\nu} = \frac{\exp\left(v_{jt+1}^{\nu} - \eta_{rj}\right)^{\varepsilon}}{\sum_l \exp\left(v_{lt+1}^{\nu} - \eta_{rl}\right)^{\varepsilon}}. \tag{OA-34}$$

Given these moving probabilities m_{jrt+1}^{ν}, the resulting law of motion for the local population is given by the same equation as in the baseline model

$$L_{rt}^{\nu} = (1-\psi)\,L_{rt-1}^{\nu} + \psi\sum_{j=1}^{R} L_{jt-1}^{\nu} m_{jrt}^{\nu}. \tag{OA-35}$$

OA-4.1.2 Firms' problem

Let the profits of a manufacturing firm in region r at time t be given by π_{rt}. Note that because of symmetry these profits are equalized across firms within a location. The present discounted value of profits is then given by

$$\Pi_{rt} = \pi_{rt} + E_t\left[\sum_{n=1}^{\infty}\left(\frac{1-\delta}{1+r}\right)^n \pi_{rt+n}\right] = \pi_{rt} + \frac{1-\delta}{1+r}E_t\left[\Pi_{rt+1}\right],$$

[49] Equation (OA-33) highlights why it is convenient to specify both the idiosyncratic preference shocks ξ and the moving costs additively. Under these assumptions, $\int_{\bar{\xi}} \max_{j\in R}\left\{v_{jt+1}^{\nu} + \xi_j - \eta_{rj}\right\} dF\left(\bar{\xi}\right)$ has a the analytic expression given in (OA-32). With the multiplicative formulation as in the baseline model, the recursive formulation would be the following. Let k_{rt}^{ν} denote the value of type ν in region r after the productivity shock $\mathcal{Q}_{rt}$ is realized but before agents had the option to move and before the realization of ξ_j. Then

$$k_{rt}^{\nu} = (1-\psi)\left(u_{rt}^{\nu}\mathcal{A}_{rt} + \beta E_t\left[k_{rt+1}^{\nu}\right]\right) + \psi\int_{\bar{\xi}} \max_j\left(u_{jt}^{\nu}\mathcal{A}_{jt}\xi_j + \beta E_t\left[k_{jt+1}^{\nu}\right]\right) dF\left(\bar{\xi}\right).$$

In contrast to (OA-32), $\int_{\bar{\xi}} \max_j\left(u_{jt}^{\nu}\mathcal{A}_{jt}\xi_j + \beta E_t\left[k_{jt+1}^{\nu}\right]\right) dF\left(\bar{\xi}\right)$ does not have an analytical expression.

where r is the exogenous discount rate and δ is the exogenous probability of exit. Free entry requires that

$$\Pi_{rt} = w_{rMt} h^E_{rt} = w_{rMt} f_E N^{-\lambda}_{rt-1}.$$

Hence, if the free entry condition holds between t and $t+1$ in region r,

$$w_{rMt} f_E N^{-\lambda}_{rt-1} = \pi_{rt} + \frac{1-\delta}{1+r} E\left[w_{rMt+1} f_E N^{-\lambda}_{rt}\right].$$

Rearranging terms yields

$$1 = \frac{\pi_{rt}}{f_E w_{rMt}} N^{\lambda}_{rt-1} + \frac{1-\delta}{1+r}\frac{1}{f_E}\left(\frac{N_{rt-1}}{N_{rt}}\right)^{\lambda} E\left[\frac{w_{rMt+1}}{w_{rMt}}\right].$$

As in the baseline model, monopolistic competition with constant markups still implies that profits are a constant fraction of aggregate revenue and labor market earnings, i.e. $\pi_{rt} = \frac{1}{\rho-1}\frac{w_{rM} H_{rP}}{N_r}$. Hence, the optimal number of varieties N_{rt} is given by

$$N_{rt} = \frac{1}{f_E}\frac{1}{\rho-1} H_{rPt} N^{\lambda}_{rt-1} + \frac{1-\delta}{1+r}\frac{1}{f_E} N^{\lambda}_{rt-1} N^{1-\lambda}_{rt} E\left[\frac{w_{rMt+1}}{w_{rMt}}\right]. \tag{OA-36}$$

The equilibrium mass of varieties in region r at time t is therefore given by a function

$$N_{rt} = N\left(\underbrace{H_{rPt}}_{+}, \underbrace{N_{rt-1}}_{+}, \underbrace{\frac{1-\delta}{1+r} E\left[\frac{w_{rMt+1}}{w_{rMt}}\right]}_{+}\right).$$

As in the baseline model, H_{rPt} is still a sufficient static for the level of *contemporaneous* market size. With forward-looking firms, however, the expected growth rate of regional skill prices $E\left[\frac{w_{rMt+1}}{w_{rMt}}\right]$ also matters because these capture the long-run value of the firm, which - because of free entry - is tied to the future entry costs and hence grows at the same rate as the local wage. Finally, again as in the baseline model, N_{rt} is also dependent on N_{rt-1} so that the process of regional productivity features persistence. The solution to (OA-36) coincides with the one of the baseline model if $\delta = 1$ or $r \to \infty$.

OA-4.1.3 Trade and labor market equilibrium

The trade and labor market equilibrium is similar to the baseline model. Because workers still receive a constant share γ and β of sectoral revenue, aggregate income in region r is still given by $PY_r = \frac{w_{Ar} H_{Ar}}{\gamma} + \frac{\rho}{\rho-1} w_{Mr} H_{rP}$. Hence, the trade market equilibrium is still given by

$$\frac{\rho}{\rho-1} w_{rMt} H_{rPt} = \sum_{j=1}^{R}\left(\frac{\tau_{rj} P_{rMt}}{\overline{P}_{jMt}}\right)^{1-\sigma}(1-\alpha)\left(\frac{w_{jAt} H_{jAt}}{\gamma} + \frac{\rho}{\rho-1} w_{jMt} H_{jPt}\right) \tag{OA-37}$$

$$\frac{w_{rAt} H_{rAt}}{\gamma} = \sum_{j=1}^{R}\left(\frac{\tau_{rj} P_{rAt}}{\overline{P}_{jAt}}\right)^{1-\sigma}\alpha\left(\frac{w_{jAt} H_{jAt}}{\gamma} + \frac{\rho}{\rho-1} w_{jMt} H_{jPt}\right), \tag{OA-38}$$

where $P_{rMt} = \frac{1}{Q_{rt}}\frac{1}{\varsigma} w_{rMt} N_{rt}^{-\frac{1}{\rho-1}}$, $P_{rAt} = \frac{1}{Q_{rt}} w_{rAt}\left(\frac{H_{rAt}}{T_r}\right)^{1-\gamma}$ and $\overline{P}_{jst} = \left(\sum_r \left(\tau_{rj} P_{rst}\right)^{1-\sigma}\right)^{1/(1-\sigma)}$, where N_{rt} is implicitly determined from (OA-36).

Similarly, the labor market clearing conditions are given by

$$
\begin{aligned}
H_{rAt} &= \Gamma_\theta \sum_{\nu=I,F} L^\nu_{rt} \left(\phi^\nu_A\right) \left(\frac{w_{rAt}}{\overline{w}^\nu_{rt}}\right)^{\theta-1} && \text{(OA-39)}\\
H_{rPt} + f_E N^{-\lambda}_{rt-1} \left(N_{rt} - (1-\delta) N_{rt-1}\right) &= \Gamma_\theta \sum_{\nu=I,F} L^\nu_{rt} \left(\phi^\nu_M\right) \left(\frac{w_{rMt}}{\overline{w}^\nu_{rt}}\right)^{\theta-1}, && \text{(OA-40)}
\end{aligned}
$$

where, again, N_{rt} is the solution determined by (OA-36). Taking $\{L^\nu_{rt}\}$ as given, equations (OA-37), (OA-38), (OA-39), (OA-40) and (OA-36) are $5\times R\times T$ equations in the $5\times R\times T$ unknowns $\{w_{rAt}, w_{rMt}, H_{rAt}, H_{rMt}, N_{rt}\}$. Note that these equations are not independent across periods because N_{rt} depends on w_{rt+1}.

OA-4.1.4 Steady state

Consider a stationary equilibrium of this economy. In such an equilibrium both the number of varieties and the local wages are constant. Hence, the number of varieties N_{rt} defined in (OA-36) is given by

$$
N_{rt} = \left(\frac{\frac{1}{\rho-1}\frac{1}{f_E}}{1-\frac{1-\delta}{1+r}\frac{1}{f_E}}\right)^{\frac{1}{1-\lambda}} H^{\frac{1}{1-\lambda}}_{rPt}.
$$

Hence, compared to (15), the full model with forward-looking agents contains an additional scaling variable $\left(1-\frac{1-\delta}{1+r}\frac{1}{f_E}\right)^{-1}$ in the denominator. However, the cross-sectional implications are exactly the same as in the baseline model because $N_{rt} \propto H^{\frac{1}{1-\lambda}}_{rPt}$.

OA-4.2 The Consequences of the Myopia Assumption

The main conceptual difference between this full dynamic model and the simplified baseline analysis can be seen in the law of motions for the dynamic state variables $\{L^\nu_{rt}\}$ and $\{N_{rt}\}$, which I replicate here for convenience:

1. ***Spatial labor supply:*** The dynamic labor supply equation is

$$
L^\nu_{rt} = (1-\psi) L^\nu_{rt-1} + \psi \sum_{j=1}^{R} L^\nu_{jt-1} m^\nu_{jrt+1},
$$

   where

$$
m^\nu_{jrt+1} = \begin{cases} \dfrac{\left(\eta_{rj}\mathcal{A}_j \frac{\overline{w}^\nu_{rt}}{P^{1-\alpha}_{rMt}P^\alpha_{rAt}}\right)^\varepsilon}{\sum_k \left(\eta_{rj}\mathcal{A}_j \frac{\overline{w}^\nu_{rt}}{P^{1-\alpha}_{rMt}P^\alpha_{rAt}}\right)^\varepsilon} & \text{with myopic agents}\\[2ex] \underbrace{\dfrac{\exp\left(v^\nu_{jt+1}-\eta_{rj}\right)^\varepsilon}{\sum_l \exp\left(v^\nu_{lt+1}-\eta_{rl}\right)^\varepsilon}}_{\text{Future wages}} & \text{with forward looking agents} \end{cases},
$$

   and v^ν_{jt+1} solves the dynamic value function in (OA-33).

2. ***Local variety creation:*** The law of motion for local varieties is given by

$$
N_{rt} = \begin{cases} \frac{1}{\rho-1}\frac{1}{f_E} H_{rPt} \left(N_{rt-1}\right)^\lambda & \text{with myopic agents}\\ N_{rt} \text{ solves } N_{rt} = \left(\frac{1}{f_E}\frac{1}{\rho-1}H_{rPt} + \frac{1-\delta}{1+r}\frac{1}{f_E}N^{1-\lambda}_{rt}\underbrace{E\left[\frac{w_{rMt+1}}{w_{rMt}}\right]}_{\text{Future wages}}\right) N^\lambda_{rt-1} & \text{with forward looking agents} \end{cases},
$$

Hence, in the case with forward looking agents, solving for the regional population at time t, L^{ν}_{rt}, and the mass of local varieties at time t, N_{rt} , requires the entire future path of local utilities $\{u^{\nu}_{rt}\}_{rt}$ and local wages $\{w^{\nu}_{rMt+1}\}_{rt}$. The equilibrium path is a dynamic fixed point. In the baseline model with myopic agents, neither L_{rt} nor N_{rt} depends on any future variables but the system is purely backward looking.

These forward-looking decisions complicate the quantitative analysis considerably. In particular, calculating the equilibrium path involves a dynamic fixed point problem. Moreover, in the presence of local productivity shocks $\mathcal{Q}_{rt}$, one has to integrate over the entire future distribution of potential shocks. To see why this is challenging, suppose first that there are no productivity shocks, that is $\mathcal{Q}_{rt} = \mathcal{Q}_r$. To calculate the equilibrium for a given set of structural parameters and for a given set of initial conditions $\left[L^I_{r0}, L^F_{r0}, N_{r0}\right]$, one has to iterate on the entire sequence of the cross-sectional population distribution. More specifically, for a given guess of $\{L^{\nu}_{rt}\}^{\infty}_{t=1}$, one can use equations (OA-37), (OA-38), (OA-39), (OA-40) and (OA-36) to solve for equilibrium wages, prices and varieties. To do so, one has to find another dynamic fixed point to make N_{rt} consistent with equilibrium wage growth. Finally, the sequence $\{L^{\nu}_{rt}\}^{\infty}_{t=1}$ has to be consistent with the implied sequence of wages and hence per-period utility.

A pseuo-code looks as follows:

1. Guess a path of $\{L^{\nu}_{rt}\}^{\infty}_{t=1}$
   (a) Solve for $\{w_{rst}, N_{rt}\}^{\infty}_{t=1}$. To do so, guess $\{N_{rt}\}^{\infty}_{t=1}$, solve for $\{w_{rst}\}^{\infty}_{t=1}$ taking $\{N_{rt}\}^{\infty}_{t=1}$ as given and update $\{N_{rt}\}^{\infty}_{t=1}$ given the resulting $\{w_{rst}\}^{\infty}_{t=1}$
   (b) Given $\{w_{rst}\}^{\infty}_{t=1}$, calculate the path of local utilities $\{u^{\nu}_{rt}\}_{rt}$.
   (c) Given $\{u^{\nu}_{rt}\}_{rt}$, calculate the optimal migration choices $\left\{\hat{L}^{\nu}_{rt}\right\}^{\infty}_{t=1}$.
2. Find a fixed point between $\left\{\hat{L}^{\nu}_{rt}\right\}^{\infty}_{t=1}$ and $\{L^{\nu}_{rt}\}^{\infty}_{t=1}$.

Hence, the outer loop consists of a fixed point in $2 \times R \times T$ variables. The inner loop consists of a fixed point in $R \times T$ variables. In the presence of stochastic shocks $\{\mathcal{Q}_{rt}\}_{r,t}$, this procedure has to be done for each sequence of shocks and then integrated over the distribution of $\{\mathcal{Q}_{rt}\}_{r,t}$. Doing so as part of the estimation algorithm makes the quantitative analysis extremely challenging.

OA-4.3 Dynamic models of migration in the literature

The recent literature on quantitative models of economic geography contains two contributions that are particularly relevant, because they also employ dynamic model of regional migration and labor reallocation but allow explicitly for forward-looking agents. In this section I discuss why the these methods are not directly applicable to my context.

"The Geography of Development" (Desmet et al., 2018) The paper by Desmet et al. (2018) is the closest in spirit to my paper. They also study a model, where individuals face frictions to spatial mobility and firms can invest in future productivity growth. Hence, in principle, both individuals and entrepreneurs face forward looking problems like in the dynamic model above.

One of the theoretical contributions of Desmet et al. (2018) is that they manage to impose sufficient theoretical restrictions to ensure that all policy functions end up being static. On the innovation side, *"they assume that innovation spills over locally after one period, and new firms can freely enter at any point in time. Hence, entry drives down firms' profits to zero in all future periods, and they only innovate to the extent that maximizes their current profits"* (Nagy, 2021, Footnote 17). This specification was developed in Desmet and Rossi-Hansberg (2014). For individuals, they assume a form of migration costs that also make individuals "current utility maximizers".

In fact, in their model, the endogenous population distribution is given by (see Desmet et al. (2018, Equations (3) and (7))

$$\frac{H(r)\,\overline{L}_t(r)}{\overline{L}} = \frac{u_t(r)^{1/\Omega}\, m_2(r)^{-1/\Omega}}{\int u_t(v)^{1/\Omega}\, m_2(v)^{-1/\Omega}\, dv} = \frac{\left(\frac{a_t(r)y_t(r)}{m_2(r)}\right)^{1/\Omega}}{\int \left(\frac{a_t(v)y_t(v)}{m_2(v)}\right)^{1/\Omega} dv}, \tag{OA-41}$$

where $H(r)\,\overline{L}_t(r)$ is the total population in region r at time t[50], $\overline{L}$ is the total world population, $y_t(r)$ is real consumption in r, $a(r)$ denotes regional amenities and $m_2(r)$ is a regional disutility term that captures migration costs. Hence, as in my baseline model, individuals' spatial mobility decisions only depend on current utility - future wages do not enter in their decision making.

Note that equation (OA-41) has another important implication that makes it not suitable for my analysis: the population distribution does not does not depend on $\overline{L}_{t-1}(r)$ and hence does not feature any persistence. Recall that the law of motion for the population in my analysis was given by

$$L^{\nu}_{rt} \quad = \quad (1-\psi)\, L^{\nu}_{rt-1} + \psi \sum_{j=1}^{R} L^{\nu}_{jt-1} \frac{\left(V_r \eta_{jr} \overline{u}^{\nu}_{rt}\right)^{\varepsilon}}{\sum_k \left(V_k \eta_{jk} \overline{u}^{\nu}_{rt}\right)^{\varepsilon}}. \tag{OA-42}$$

If $\psi = 1$ (i.e. individuals are free to move each period) and $\eta_{jr} = \eta_r$ (i.e. migration frictions are not origin-destination specific but only destination specific) this equation reduces to

$$L^{\nu}_{rt} \quad = \quad \frac{\left(V_r \eta_r \overline{u}^{\nu}_{rt}\right)^{\varepsilon}}{\sum_k \left(V_k \eta_k \overline{u}^{\nu}_{rt}\right)^{\varepsilon}} \sum_{j=1}^{R} L^{\nu}_{jt-1} = \frac{\left(V_r \eta_r \overline{u}^{\nu}_{rt}\right)^{\varepsilon}}{\sum_k \left(V_k \eta_k \overline{u}^{\nu}_{rt}\right)^{\varepsilon}} L^{\nu},$$

where $L^{\nu} = \sum_{j=1}^{R} L^{\nu}_{jt-1}$ denotes the total population. This expression is isomorphic to equation (OA-41). In terms of my context this implies that the initial "drop" in refugees would not show any persistence, which is counterfactual. In fact, the Poisson mobility shock ψ is one of the parameters I estimated and the data pointed to number substantially smaller than 1. The assumption of individuals being myopic thus allows me to estimate the more general migration equation in (OA-42) rather than having to rely on (OA-41).

"Trade and labor market dynamics: General equilibrium analysis of the china trade shock" (Caliendo et al., 2019) Another recent contribution that makes substantial progress regarding the problem of forward-looking migration decisions is Caliendo et al. (2019). They study the labor market implications of changes in trade policy in settings where individuals face sectoral and regional mobility frictions and are forward looking. The methodological contribution of Caliendo et al. (2019) is that they demonstrate that many questions of interest - like counterfactual changes in trade policy - can be answered without having to estimate many parameters of the model (a procedure that they refer to as "dynamic hat algebra").

There are two reasons why the insights from Caliendo et al. (2019) are not applicable to my setting. The first is conceptual: the Caliendo et al. (2019) methodology relies on the fact that regional productivity is an exogenous fundamental. While it can be varying, their method is not directly applicable to a setting where local productivity responds endogenously to market conditions. This endogeneity of local productivity, however, is a key aspect of my analysis.

The second is empirical: a key methodological insight of Caliendo et al. (2019) is to use a "revealed preference" approach, where observable data can be directly substituted in some equilibrium conditions. In their words: *"Aside from data that directly map into the model's equilibrium conditions, the only parameters we need in order to solve the full transition of the dynamic model are the trade elasticities, the migration elasticity, and the intertemporal discount factor."* Caliendo et al. (2019, p. 742) This is particularly relevant to solve agents' dynamic migration problem in (OA-33): if individuals' migration choices can be observed

[50] In their model $H(r)$ is the amount of land in location r and $\overline{L}_t(r)$ is population per unit of land at r in period t.

in the initial period (say $t = 0$), these "choice probabilities" m^{ν}_{rj0} can be used to infer the initial value of life-time utility V^{ν}_{r0}, that rationalizes such choices. Given these initial value functions, the recursive equation in (OA-33) can then be iterated forward and hence jointly solved with the distribution of wages. In order to apply this procedure to my setting, I would need to observe the matrix of migration choices for both types ν. However, not only do I treat the type ν as an unobserved latent variable, but even if I were to abstract from such unobserved types, I do not have data on migration flows for all origin-destination pairs in the initial period, i.e. in 1950 after the inflow of refugees.

OA-4.4 Implications

The key question is of course to what extent my substantive results hinge on the assumption of myopic agents. Without actually computing the solution to the full dynamic model, it is difficult to answer this question. However, it is worthwhile to point out that my empirical strategy suggests that for many outcomes of interest of this study, namely the aggregate and regional effects of refugee inflows on income per capita and industrialization, the discrepancy between the myopic and the full dynamic model might be small.

The reason is the following: as highlighted in the characterization of the dynamic equilibrium in Section OA-4, both models have the exact same solution given a time path for the population distribution $\{L^{\nu}_{rt}\}$ and the regional varieties $\{N_{rt}\}$. Hence, the two models' implications only differ to the extent that they imply different paths for these dynamic state variables.

However, in my empirical strategy, I discipline my structural parameters directly by targeting moments, that are tightly linked to the evolution of exactly such state variables: the impact on short- and long-run population growth, the auto-correlation of the refugee share and the impact on income per capita, again both in the short- and long-run. In addition I also directly target moments related to agents' observed choices, namely the ratio of new to gross flows and the share of migration to out-of-state counties.

If a fully dynamic model were to be estimated to the same moments, the implied structural parameters would of course be different, but the response of the model to the population inflows would be - by construction - the same. It is therefore not immediate that the two models would have very different implications if such flows had been absent. It is in that sense that my empirical strategy of relying on indirect inference might deliver credible results even if the model was misspecified.

This is, of course, not to say that these different models would have similar results to *all* shocks. An announcement of a sizable refugee inflows in five years time would naturally have immediate consequences in the fully dynamic model while nothing would happen in the myopic model. Similarly, a quantification of the welfare consequences might yield very different results.