

**SUPPLEMENTARY MATERIAL FOR
 "SPATIAL STRUCTURAL CHANGE"
 BY FABIAN ECKERT AND MICHAEL PETERS**

This Technical Appendix contains detailed information on how we calibrate and solve our quantitative model. The information in this Appendix can be used to replicate the results in the paper.

A. SOLVING THE MODEL

We solve the model in two steps. First, we take the number of workers in each location, L_r , as fixed and solve for the labor market equilibrium in each region ("inner loop"). Second, given the set of equilibrium wages $(\bar{\chi}_{rt}, w_{rMt})$, we update local populations given the wages implied by the labor market equilibrium ("outer loop"). This procedure converges to a set of wages and workers in each location that are consistent with the equilibrium conditions of our model. Note that relative to the main part of the paper, we model the agricultural sectors as an Armington sector with regional varieties that have a very high elasticity of substitution, σ_A , and introduce trade costs in *both* sectors. We first describe the "inner loop" solving for wages and agricultural profits, taking workers in each region (L_r) as given. When describing the inner loop, we drop time subscripts to save on notation since the inner loop is all contemporaneous. We choose the agricultural price in region 1 as the numeraire throughout so that $P_{1A} = 1$.

A.1 Inner Loop

Step 0: Guess manufacturing wages w_{rM} and agricultural profitability $\bar{\chi}_r$ in each region.

Step 1: Compute average wages in each location:

$$\bar{w}_r = \left(w_{rM}^\zeta + \bar{\chi}_r^\zeta \right)^{1/\zeta}.$$

Step 2: Solve for sectoral employment shares:

$$s_{rM} = \left(\frac{w_{rM}}{\bar{w}_r} \right)^\zeta \quad \text{and} \quad s_{rA} = \left(\frac{\bar{\chi}_r}{\bar{w}_r} \right)^\zeta.$$

Step 3: Compute the human capital stock in each location-sector:

$$H_{rs} = \Gamma_\zeta L_r s_{rs}^{\frac{\zeta-1}{\zeta}}.$$

Step 5: Compute factory gate prices in each sector²⁸:

$$p_{rA} = \frac{\bar{\chi}_{rt}}{A_{rt}} \left(\frac{\gamma + 1 - \alpha}{1 - \alpha} \right)^{1-\alpha} \left(\frac{H_{rA}}{T_r} \right)^\alpha \quad \text{and} \quad p_{rM} = \frac{\sigma_M}{\sigma_M - 1} \frac{w_{rM}}{Z_r}.$$

Step 6: Compute the number of firms in manufacturing:

$$N_r = \frac{1}{\sigma f_E} H_{rM}.$$

Step 7: Compute price indexes in both sectors:

$$P_{rM}^{1-\sigma_s} = \sum_j N_j (p_{jM} \tau_{jr})^{1-\sigma_s} \quad \text{and} \quad P_{rA}^{1-\sigma_s} = \sum_j (p_{jA} \tau_{jr})^{1-\sigma_s}.$$

Step 8: Compute real income

$$\bar{v}_r = \frac{\bar{w}_r}{P_{rA}^\phi P_{rM}^{1-\phi}}.$$

Step 9: Compute Expenditure Shares Use real income \bar{v}_r to compute sectoral expenditure shares for each region and sector

$$\vartheta_{rM} = (1 - \phi) - \nu^{\text{RC}} \bar{v}_r^{-\eta} \quad \text{and} \quad \vartheta_{rA} = 1 - \vartheta_{rM}.$$

Note: Here we ensure in the code that these shares are bounded between zero and one. If ϑ_{rM} below 0 set to small positive numbers if ϑ_{rM} above 1 we set to number just below 1.

Step 10: Construct sectoral trade shares:

$$\pi_{rj}^M = N_r \left(\frac{p_{rM} \tau_{rj}}{P_{j,M}} \right)^{1-\sigma_M} \quad \text{and} \quad \pi_{rj}^A = \left(\frac{p_{rM} \tau_{rj}}{P_{j,A}} \right)^{1-\sigma_A}.$$

Step 11: Compute total local sales in each location and sectors as:

$$R_{rM} = \sum_j \pi_{rj}^M \vartheta_{jM} \Gamma_\zeta \bar{w}_r L_r \quad \text{and} \quad R_{rA} = \sum_j \pi_{rj}^A \vartheta_{jA} \Gamma_\zeta \bar{w}_r L_r$$

²⁸If p_{rA} is the price at which agricultural goods in r are sold, total profits of type z_A^i are

$$\chi(z_A^i) = (1 - \alpha) \left(\frac{1 - \alpha}{\gamma + 1 - \alpha} \right)^{1-\alpha} p_{rA} A_{rt} \left(\frac{T_r}{H_{rA}} \right)^\alpha z_A^i.$$

Hence, the revenue productivity per efficiency unit in region r is

$$\bar{\chi}_{rt} = p_{rA} A_{rt} \left(\frac{1 - \alpha}{\gamma + 1 - \alpha} \right)^{1-\alpha} \left(\frac{T_r}{H_{rA}} \right)^\alpha.$$

Step 12: Note that all sales are paid to labor in both sectors since agricultural land is owned by agricultural workers and free entry implies that profits are zero. Also note that average earnings are the same in both sectors. But then we can update the sectoral employment shares:

$$s_{rs}^{NEW} = \frac{R_{rs}}{R_{rA} + R_{rM}}$$

must hold. We use this to update wages in each sector:

$$\bar{\chi}_r^{NEW} = \frac{R_{rA}}{\Gamma_\zeta L_r} \left(s_{rs}^{NEW} \right)^{\frac{1-\zeta}{\zeta}} \quad \text{and} \quad w_{rM}^{NEW} = \bar{\chi}_r^{NEW} \left(\frac{R_{rM}}{R_{rA}} \right)^{1/\zeta}$$

Note: As a last step we update \bar{w}_r and use it as a natural convergence criterion.

Iterate on steps 1-11 until sectoral wages have converged.

A.2 Outer Loop

Once the inner loop has converged, we use the number of period $t - 1$ workers in each region, L_{rt-1} , and the spatial law of motion,

$$L_{jt} = \sum_{r=1}^R m_{rjt} n_{rt-1} L_{rt-1},$$

to update the population in each location. The fraction of workers who moves from r to j , m_{rj} , is given by:

$$m_{rjt} = \frac{\left(\mu_{rj}^{-1} \mathcal{V}_{jt} \mathcal{B}_{jt} \right)^\varepsilon}{\sum_k \left(\mu_{rk}^{-1} \mathcal{V}_{kt} \mathcal{B}_{kt} \right)^\varepsilon},$$

where \mathcal{V}_{rt} is given by:

$$\mathcal{V}_{rt} = \int V(y, P_{rAt}, P_{rMt}) dF_{rt}(y) = \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \bar{v}_{rt}^\eta - v \ln(P_{rAt}/P_{rMt}),$$

and where $\mu_{rj}^{-1} = d_{rj}^{-\theta}$. Note that amenities have an endogenous component and are given by: $\mathcal{B}_r := B_r L_r^{-\rho}$.

After updating the population in this way, we go back into the inner loop. We have found an equilibrium when the population distribution is consistent with the equilibrium wages given L_{jt} .

Note that we update last periods population for exogenous population growth before moving on to the next period.

B. CALIBRATING THE MODEL

In this section, we describe how we calibrate the parameters of our model.

B.1 Calibrating Regional Fundamentals $[Z_{r1880}, A_{r1880}, B_{r1880}, T_{r1880}]_r$

We have data on agricultural employment shares s_{rA} , local population L_r , local land prices r_r , and local average wages \bar{w}_r .

Step 1: Compute Human Capital in each Region Given the agricultural employment share for each region, $\{s_{rs}\}$, and the total population, L_r , we can compute total human capital in each region by sector s as

$$H_{rs} = L_r \Gamma_\zeta s_{rs}^{\frac{\zeta-1}{\zeta}}.$$

Step 2: Compute Sectoral Prices Similarly, we can compute wages for each sector by solving these two equations:

$$\frac{w_{rM}}{\bar{\chi}_r} = \left[\frac{1 - s_{rA}}{s_{rA}} \right]^{\frac{1}{\zeta}} \quad \bar{w}_r = \left(w_{rM}^\zeta + \bar{\chi}_r^\zeta \right)^{1/\zeta},$$

together with the data on average wages, \bar{w}_r and sectoral employment shares s_{rs} .

Step 3: Loop Solving for Factory Gate Prices in each Sector Now we solve for the regional equilibrium prices:

Step 3.1: Guess price index in each region and sector, P_{rA} and P_{rM}

Step 3.2: Compute “modified” trade shares

We compute the following modified trade shares, that do not have prices in the numerator:

$$\tilde{\pi}_{rj}^s = \left(\frac{\tau_{rj}}{P_{rj}} \right)^{1-\sigma_s} = \frac{\pi_{rjs}}{p_{rs}^{1-\sigma_s}}.$$

Also note that for manufacturing, we define $p_{rM} = \tilde{p}_{rM} N_r^{\frac{1}{1-\sigma_M}}$ where \tilde{p}_{rM} is the true factory gate price of a variety from a firm in region r . Note that σ_s now depends on s .

Step 3.3: Compute Expenditure Shares

The expenditure share on manufacturing goods in region r is given by

$$\vartheta_{rM} = 1 - \phi - \nu^{RC} \left(\frac{\bar{w}_r}{P_{rA}^\phi P_{rM}^{(1-\phi)}} \right)^{-\eta}.$$

Step 3.4: Compute Implied Factory Gate Price

We can now infer the factory gate prices using sectoral market clearing:

$$(OA.1) \quad p_{rM}^{NEW} = \left[\frac{1}{w_{rM} H_{rM}} \sum_j \tilde{\pi}_{rjM} \vartheta_{jM} \Gamma_\zeta \bar{w}_r L_r \right]^{\frac{1}{\sigma_M-1}}$$

and

$$(OA.2) \quad p_{rA}^{NEW} = \left[\frac{1}{\bar{\chi}_r H_{rA}} \sum_j \tilde{\pi}_{rjA} \vartheta_{jA} \Gamma_\zeta \bar{w}_r L_r \right]^{\frac{1}{\sigma_A - 1}}.$$

Step 3.5: Update Price Index

Given $\{p_{rs}^{NEW}\}_r$, we calculate the implied price index in region r and sectors s

$$P_{rs}^{1-\sigma_s} = \sum_j (p_{r's}^{NEW} \tau_{jr})^{1-\sigma_s}.$$

In updating we ensure that $P_{1A} = 1$ is enforced, i.e., the agricultural price index is unity in region one. We then iterate on this procedure until convergence.

Step 4: Compute Number of Firms We can infer the mass of manufacturing firms active in each region from manufacturing labor supply in efficiency units

$$N_r = \frac{1}{\sigma_M f_E} H_{rM}.$$

Step 5: Compute Implied Fundamentals We can now infer regional manufacturing productivity:

$$(OA.3) \quad Z_r = \frac{\sigma_M}{\sigma_M - 1} \frac{w_{rM}}{p_{rM}} N_r^{\frac{1}{1-\sigma_M}}$$

For the agricultural sector, note first that we can infer agricultural land given information on land values:

$$(OA.4) \quad R_t T_r = \alpha \bar{\chi}_{rt} H_{rAt},$$

Given T_r , $\bar{\chi}_{rt}$, and agricultural factory-gate prices p_{rA} , we can infer agricultural productivity A_{rt} from

$$A_{rt} = \frac{\bar{\chi}_{rt}}{p_{rA}} \left(\frac{\gamma + 1 - \alpha}{1 - \alpha} \right)^{1-\alpha} \left(\frac{H_{rA}}{T_r} \right)^\alpha$$

Step 8: Infer Amenities Region Now we infer regional amenities. We choose the set of regional amenities $\{\mathcal{B}_r\}_r$ to ensure that the 1880 distribution of workers across locations is implied by the 1860 distribution of workers and the implied migration shares of the model. In other words, we choose $\{\mathcal{B}_r\}_r$ to ensure the following law of

motion,

$$L_{r1880} = \sum_j m_{jr1880} L_{j1860} n_{j1860},$$

holds for the observed levels of 1860 and 1880 employment given the model-implied moving probabilities:

$$m_{jrt} = \frac{\left(\mu_{jr}^{-1} \left(\frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} (v_{rt})^\eta - v \ln \left(\frac{P_{rA}}{P_{rM}} \right) \right) \mathcal{B}_r \right)^\varepsilon}{\sum_q \left(\mu_{jq}^{-1} \left(\frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} (v_{qt})^\eta - v \ln \left(\frac{P_{qA}}{P_{qM}} \right) \right) \mathcal{B}_q \right)^\varepsilon} \equiv \frac{\Omega_{jrt} \tilde{\mathcal{B}}_r}{\sum_q \Omega_{jq} \tilde{\mathcal{B}}_q}.$$

The previous equation provides a simple iterative updating equation:

$$\tilde{\mathcal{B}}_r^{k+1} = L_{r1880} \left[\sum_j \frac{\Omega_{jrt}}{\sum_q \Omega_{jq} \tilde{\mathcal{B}}_q^k} L_{j1860} n_{j1860} \right]^{-1},$$

where k indexes the iteration number. Note that everyone on the right hand side is either data or fully determined by the 1880 equilibrium equations given 1880 employment in each location. So we guess parameters, compute the static equilibrium in 1880 given 1880 observed employment, and then update amenities until $\tilde{\mathcal{B}}_r^{k+1} \approx \tilde{\mathcal{B}}_r^k$. Note that amenities are only identified up to scale.

Once we have identified $\tilde{\mathcal{B}}_r$, we use ε and ρ to isolate the fundamental part of amenities in each location \bar{B}_{rt} .

B.1.1 Computing Macroeconomic Aggregates

In this section, we describe how we measure GDP and average consumer prices.

Growth in GDP per Capita We measure the growth rate of real aggregate GDP per capita using the Fisher chained index. The Fisher Index is defined by

$$g_t^F = \sqrt{\frac{P_{ct-1} C_t / L_t}{P_{ct-1} C_{t-1} / L_{t-1}} \times \frac{P_{ct} C_t / L_t}{P_{ct} C_{t-1} / L_{t-1}}},$$

where P_{ct} is the price of the consumption good at time t , C_t is the total quantity of consumption, and L_t is the total population.

In the context of our model with regions r and sectors s , we construct the following auxiliary indices of spending per capita, evaluated at the respective prices:

$$S_{t-1}(P_{t-1}) = \frac{\sum_{r=1}^R \sum_{s=1}^S P_{rst-1} C_{rst-1}}{\sum_{r=1}^R L_{rt-1}} \quad \text{and} \quad S_{t-1}(P_t) = \frac{\sum_{r=1}^R \sum_{s=1}^S P_{rst} C_{rst-1}}{\sum_{r=1}^R L_{rt-1}}.$$

and

$$S_t(P_{t-1}) = \frac{\sum_{r=1}^R \sum_{s=1}^S P_{rst-1} C_{rst}}{\sum_{r=1}^R L_{rt}} \quad \text{and} \quad S_t(P_t) = \frac{\sum_{r=1}^R \sum_{s=1}^S P_{rst} C_{rst}}{\sum_{r=1}^R L_{rt}}.$$

where $P_{rst}(C_{rst})$ denotes the prices (consumption quantities) of sector s goods in region r at time t .

We then combine these expressions into the corresponding Fisher Index as follows:

$$g_t^F = \sqrt{\frac{S_t(P_{t-1})}{S_{t-1}(P_{t-1})} \times \frac{S_t(P_t)}{S_{t-1}(P_t)}}.$$

In our model, C_{rst} can be computed as

$$\begin{aligned} C_{rAt} &= \frac{\vartheta_{rAt} \Gamma_{\zeta} \bar{w}_{rt} L_{rt}}{P_{rAt}} \\ C_{rMt} &= \frac{(1 - \vartheta_{rAt}) \Gamma_{\zeta} \bar{w}_{rt} L_{rt}}{P_{rMt}}. \end{aligned}$$

Relative Prices To compute the time-series of the relative price of manufacturing to agricultural goods, we compute chained sectoral price indices and then take their ratio. More specifically, consider sector s and time-period between $t - 1$ and t . Let P_{st}^L and P_{st}^P denote the Laspeyres and Paasche indices, respectively. These are given by

$$P_{st}^L = \frac{\sum_r P_{rst} C_{rst-1}}{\sum_r P_{rst-1} C_{rst-1}} \quad P_{st}^P = \frac{\sum_r P_{rst} C_{rst}}{\sum_r P_{rst-1} C_{rst}}.$$

The Fisher Index for sector s is then given by $P_{st}^F = \sqrt{P_{st}^L \times P_{st}^P}$. The time-series of the relative price is then given by $\mathcal{P}_t^{M-A} = P_{Mt}^F / P_{At}^F$.

B.2 Validating the First-Order Approximation

In Figure 7, we decomposed local wage growth into the four components highlighted in Proposition 4. Using Proposition 3, the first-order approximations for local wage growth and local structural change are given by:

$$(OA.5) \quad d \ln \bar{w}_{rt} = \phi_M(s_{rA}) d \ln \mathcal{Z}_{rt} + (1 - \phi_M(s_{rA})) d \ln \mathcal{A}_{rt}$$

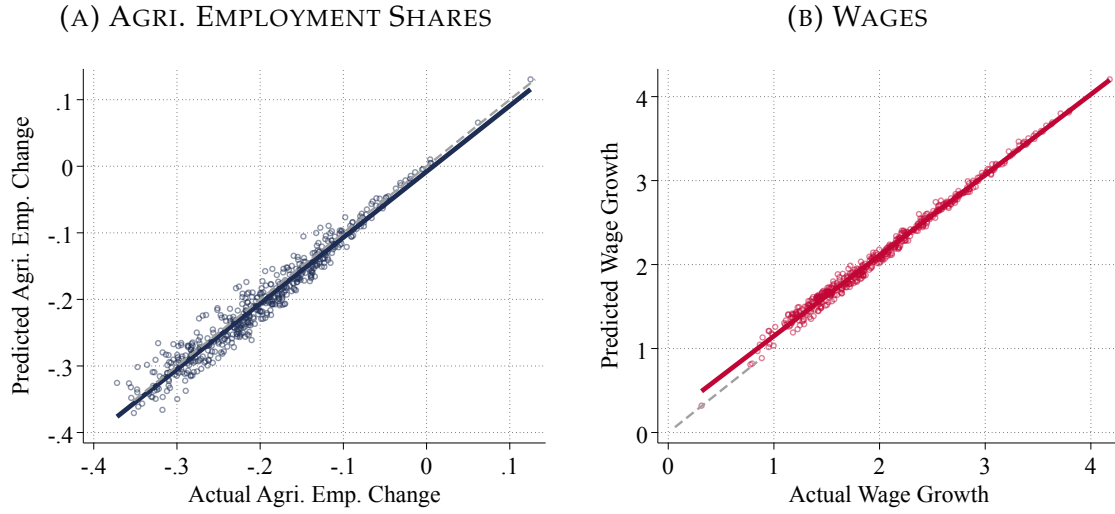
$$(OA.6) \quad ds_{rAt} = \psi(s_{rA}) (d \ln \mathcal{Z}_{rt} - d \ln \mathcal{A}_{rt}),$$

where

$$\begin{aligned} d \ln \mathcal{Z}_{rt} &= \frac{\sigma_M - 1}{\sigma_M} d \ln Z_{rt} + \frac{1}{\sigma_M} d \ln \mathcal{D}_{rMt}, \\ d \ln \mathcal{A}_{rt} &= \frac{\sigma_A - 1}{\sigma_A} d \ln A_{rt} + \frac{1}{\sigma_A} d \ln \mathcal{D}_{rAt} - \alpha \frac{\sigma_A - 1}{\sigma_A} d \ln \left(\frac{L_{rt}}{T_r} \right) - \frac{1}{\sigma_A} d \ln L_{rt} \end{aligned}$$

Figure OA.1 shows that these first-order approximations closely match the full nonlinear solution of the model. The right panel plots local wage growth predicted by equation (OA.5) against the model-simulated local wage growth; the left panel plots predicted versus simulated changes in the agricultural employment share using equation (OA.6). Each dot represents a commuting zone. The dashed gray line marks the 45-degree line, while the solid line depicts the best linear fit. If the model followed the first-order approximations exactly, all observations would lie on the 45-degree line. The figure demonstrates that the approximations perform remarkably well, validating the

FIGURE OA.1: ASSESSING THE ACCURACY OF THE FIRST-ORDER-APPROXIMATIONS



Notes: The left (right) panel shows the correlation between actual agricultural employment share changes (actual wage growth) in the model and predicted agricultural employment share changes (predicted wage growth) based on Proposition 4. Each dot is a commuting zone. The dashed grey line is a 45 degree line. The solid line in each panel is a weighted fit line using 1880 total employment as weights.

decomposition analysis reported in Section 3.

B.3 Normalizing the Level of Utility

Our model of spatial mobility requires expected utility \mathcal{V}_{rt} to be positive. Given the expression for \mathcal{V}_{rt} :

$$\mathcal{V}_{rt} = \frac{1}{\eta} \Gamma_{\xi} \bar{v}_{rt}^{\eta} - v \ln(P_{rAt}/P_{rMt}),$$

this requires real income \bar{v}_{rt} and manufacturing prices P_{rMt} to be sufficiently large.

In terms of fundamental parameters, we can always achieve this by the appropriate choice of productivity parameters Z and T . In particular, given our empirical strategy we can choose the level of the observed wages $\{\bar{w}_{rt}\}$ to determine the level of utility. To see this, suppose we found a set of prices p_{rM} and p_{rA} that rationalized the observed data for a given set of wages $\{\bar{w}_r\}_r$. Now suppose we multiply all wages by a factor m , i.e. $\tilde{w}_{rt} = m\bar{w}_{rt}$. Given the observed employment shares, this also implies that $\tilde{w}_{rs} = mw_{rs}$.

It is then easy to see that $\tilde{p}_{rM} = m^{\frac{1}{1-\phi}} p_{rM}$ and $\tilde{p}_{rA} = p_{rA}$ still rationalize the same data. Note that the agricultural prices do not change given the fact that we take P_{rA} as a numeraire for one particular region. This implies that $\tilde{P}_{rA} = P_{rM}$. Moreover, if $\tilde{p}_{rM} = m^{\frac{1}{1-\phi}} p_{rM}$, then $\tilde{P}_{rM} = m^{\frac{1}{1-\phi}} P_{rM}$. This implies that, by construction,

$$\left(\frac{\bar{w}_r}{P_{rM}^{1-\phi} P_{rA}^{\phi}} \right)^{-\eta} = \left(\frac{\tilde{w}_{rt}}{\tilde{P}_{rM}^{1-\phi} \tilde{P}_{rA}^{\phi}} \right)^{-\eta}.$$

Hence, conditional on parameters, the expenditure share $\vartheta_{r,M}$ stays the same. But for given $\vartheta_{r,M}$ the previous mapping is linearly homogeneous in \bar{w} and p_{rM} and hence

does not change. The implied productivity \tilde{Z}_r is therefore given by (see equation OA.3)

$$\tilde{Z}_r = \frac{\sigma_M}{\sigma_M - 1} \frac{\tilde{w}_{rM}}{\tilde{p}_r^M} = \frac{\sigma_M}{\sigma_M - 1} \frac{m w_{rM}}{m^{\frac{1}{1-\phi}} p_{rM}} = m^{1-\frac{1}{1-\phi}} \frac{\sigma_M}{\sigma_M - 1} \frac{w_{rM}}{p_{rM}} = m^{-\frac{\phi}{1-\phi}} Z_r < Z_r,$$

i.e. regional productivity is smaller by a factor $m^{-\frac{\phi}{1-\phi}}$. Intuitively, to rationalize the same allocations, "real income" $\frac{\tilde{w}_r}{p_{rM}^{1-\phi} p_{rA}^\phi}$ has to be the same. Given the choice of agriculture as a numeraire, income in terms of agricultural goods is higher. This implies that manufacturing prices have to rise, i.e. manufacturing productivity has to be lower.

Similarly, the new land supply \tilde{T}_r is be given by

$$\tilde{T}_r = \alpha \frac{H_{rA} \tilde{w}_{rA}}{r_r} = m T_r > T_r,$$

and the new level of agricultural productivity \tilde{A}_r is given by

$$\tilde{A}_r = \frac{\tilde{w}_{rA}^{1-\alpha} r^\alpha}{\tilde{p}_{rA} \alpha^\alpha} = \frac{(m w_{rA})^{1-\alpha} r^\alpha}{p_{rA} \alpha^\alpha} = m^{1-\alpha} A_r > A_r$$

Hence, both T_r and A_r are higher. Intuitively: to rationalize a given agricultural employment share with a higher wage, the model requires manufacturing productivity to be lower and land endowments to be higher, for the relative price in agriculture to be lower (recall that our model implies that the elasticity of substitution exceeds unity). This price effect is needed to counteract the non-homotheticity stemming from higher utility.

To see that inferred utility is indeed higher, note that expected utility is given by

$$\begin{aligned} \mathcal{V}_{rt} &= \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \bar{v}_{rt}^\eta - \nu \ln(p_{rA}/p_{rM}) \\ &= \frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \bar{v}_{rt}^\eta - \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \nu^{RC} \ln(p_{rA}/p_{rM}) \\ \text{(OA.7)} \quad &= \nu^{RC} \left(\frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \frac{1}{\nu^{RC} \bar{v}_{rt}^{-\eta}} - \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln(p_{rA}/p_{rM}) \right). \end{aligned}$$

Because $\bar{v}_{rt}^{-\eta}$ and p_{rA} stays the same by construction, expected utility as a function of the scaling parameter m , $\mathcal{V}_{rt}(m)$, is given by

$$\begin{aligned} \mathcal{V}_{rt}(m) &= \nu^{RC} \left(\frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \frac{1}{\nu^{RC} \bar{v}_{rt}^{-\eta}} - \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln \left(\frac{p_{rA}}{p_{rM} m^{\frac{1}{1-\phi}}} \right) \right) \\ &= \nu^{RC} \left(\frac{1}{\eta} \Gamma_{\frac{\zeta}{\eta}} \frac{1}{\nu^{RC} \bar{v}_{rt}^{-\eta}} + \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln(p_{rM}) \right) + \nu^{RC} \frac{1}{1-\phi} \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln(m) \\ &= \mathcal{V}_{rt}(1) + \nu^{RC} \frac{1}{1-\phi} \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln(m). \end{aligned}$$

Hence, multiplying all wages by m , shifts utility by a constant $\nu^{RC} \frac{1}{1-\phi} \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \ln(m)$. By choosing m , we can thus determine the level of utility.

We scale m in the initial allocation in 1880 to set the minimum of utility to unity. Hence, we are going to chose m as

$$m = \exp \left((1 - \mathcal{V}_{rt}(1)) \frac{1 - \phi}{\nu^{RC}} \frac{\Gamma_\zeta}{\Gamma_{\zeta/(1-\eta)}} \right).$$

And because productivity is growing over time, expected utility also increases. Hence, $\mathcal{V}_{rt}(m) > 0$ for all t .

B.4 Identification of Parameters

Market clearing for agricultural goods requires that

$$(OA.8) \quad w_{rA} H_{rA} = \sum_j \left(\frac{\tau_{rj} p_{rA}}{P_{jA}} \right)^{1-\sigma_A} \left(\phi + \nu^{RC} \left(\frac{\bar{w}_j}{P_{jM}^{1-\phi} P_{jA}^\phi} \right)^{-\eta} \right) \Gamma_\zeta \bar{w}_j L_j$$

Market clearing for manufacturing goods requires that

$$(OA.9) \quad w_{rM} = \frac{1}{\sigma_M f_E} \sum_j \left(\frac{\tau_{rj} p_{rM}}{P_{jM}} \right)^{1-\sigma_M} \left(1 - \phi - \nu^{RC} \left(\frac{\bar{w}_j}{P_{jM}^{1-\phi} P_{jA}^\phi} \right)^{-\eta} \right) \Gamma_\zeta \bar{w}_j L_j.$$

Note that w_{rA} , w_{rM} , \bar{w}_r and L_j are observed. Note that also that given $[\tau_{rj}]$, f_E and the fact that H_{rM} is observed, P_{jA} and P_{jM} are fully determined from p_{rA} and p_{rM} . Hence: equations (OA.8) and (OA.9) are $2R$ equations in $2R$ unknowns $[p_{rA}, p_{rM}]$. However, equations (OA.8) and (OA.9) are not independent because they imply that

$$\sum_r w_{rMt} H_{rMt} = \sum_j \vartheta_{jMt} \Gamma_\zeta L_{jt} \bar{w}_{jt} \quad \text{and} \quad \sum_r w_{rA} H_{rA} = \sum_j \vartheta_{jAt} \Gamma_\zeta \bar{w}_j L_j.$$

Hence, equations (OA.8) and (OA.9) are $2R - 1$ equations. This implies that given $\{w_{rM}, w_{rA}, L_r; \phi, \nu^{RC}, \eta\}$ we can identify $\{p_{rM}, p_{rA}\}$ subject to one choice of a numeraire. Alternatively, given $\{\bar{w}_r, s_{rA}, L_r; \phi, \nu^{RC}, \eta, \zeta\}$, we can identify $\{p_{rM}, p_{rA}\}$ subject to one choice of a numeraire.

B.4.1 Preference Parameters $\{\phi, \nu^{RC}, \eta\}$

All of the above was conditional on the 3 structural preference parameters $\{\phi, \nu^{RC}, \eta\}$. The cross-section in 1880 does not add any degree of freedom, because we treat $\{Z_{r1880}, A_{r1880}, B_{r1880}, T_{r1880}\}$ as structural residuals. Moreover, as we show below in Section B.4.3, we can without loss of generality normalize ν^{RC} by a choice of units. Hence, we need two moments to identify $\{\phi, \eta\}$. We use the following two:

1. The change in agricultural employment shares between 1880 and 1920: η
2. A long-run restriction: because ϕ is the asymptotic agricultural value added share,

we set $\phi \approx 0.01$.

B.4.2 Fixed cost f_E

Note that f_E drops out of the above equations: We can write everything as a function of $\tilde{p}_{rM} \equiv f_E^{-\frac{1}{1-\sigma_M}} p_{rM}$. Then we can solve the above for $\{\tilde{p}_{rM}, p_{rA}\}$ as f_E is no longer in the equations (note that f_E is also in the price index). Hence, f_E is akin to a shift in Z_r as it is essentially a common shift in the units of N_r . Hence, setting $f_E = 1$ is without loss of generality.

B.4.3 The preference shifter ν

In this section, we show that we can normalize ν without loss. The overall expenditure share on agricultural goods is given by

$$\vartheta_{rA} = \phi + \nu^{RC} \left(\frac{\bar{w}_r}{P_{rA}^\phi P_{rM}^{1-\phi}} \right)^{-\eta} \equiv \phi + \nu^{RC} \bar{y}_r^{-\eta},$$

where \bar{y}_r denote real income. Hence, ν^{RC} is not separately identified from the level of real income, which itself depends on the choice of units.

To see this directly, consider the following single-region economy with homogeneous workers, where agricultural and manufacturing production is given by $Y_A = AL$ and $Y_M = ZL$ respectively. Taking the agricultural good as the numeraire, it follows that $w = A$ and $P_M = \frac{w}{Z} = \frac{A}{Z}$. Real income is thus given by

$$\bar{y}_r = \frac{w}{1\phi(w/Z)^{1-\phi}} = \frac{A}{1\phi(A/Z)^{1-\phi}} = A^\phi Z^{1-\phi}.$$

The agricultural spending share is thus given by $\vartheta_{rA} = \phi + \nu(A^\phi Z^{1-\phi})^{-\eta}$, which shows that the level of ν is not separately identified from the level of A or Z . Intuitively, a given expenditure share can either be generated by a large demand shifter for agricultural goods with rich consumers (high productivity) or a small demand shifter with poor consumers.

In our application we normalize η to unity and pick the levels productivity (alternatively real income) to ensure that utility is positive (see Section B.3). The latter is important in our application because expected utility is allocative through agents' spatial mobility.

B.5 Human Capital or Productivity? Detailed Derivations

Suppose the human capital distribution in region r at time t is sector-specific and given by $F_{rst}(z) = \exp\left(-\left(\frac{z}{h_{rt}^s}\right)^{-\zeta}\right)$, where h_{rt}^s parametrizes the amount of sector- s human capital individuals in region r have access to. Our baseline model is nested as the case of $h_{rt}^s = 1$. Given w_{rMt} and $\bar{\chi}_{rt}$, it then follows that

$$(OA.10) \quad \bar{w}_{rt} = \left(\left(h_{rt}^A \bar{\chi}_{rt} \right)^\zeta + \left(h_{rt}^M w_{rMt} \right)^\zeta \right)^{1/\zeta}$$

and

$$(OA.11) \quad s_{rA} = \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta \quad \text{and} \quad H_{rAt} = \Gamma_\zeta L_r h_{rt}^A s_{rA}^{\frac{\zeta-1}{\zeta}} = \Gamma_\zeta L_r h_{rt}^A \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{\zeta-1}.$$

Manufacturing wages are still given by $w_{rMt} = \left(\frac{1}{\bar{J}_E} \right)^{1/\sigma_M} Z_{rt}^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \equiv \mathcal{Z}_{rt}$, so that

$$(OA.12) \quad h_{rt}^M w_{rMt} = h_{rt}^M \mathcal{Z}_{rt} \equiv \mathcal{Z}_{rt}^{HC}.$$

Similarly, agricultural profitability $\bar{\chi}_{rt}$ is given by

$$\bar{\chi}_{rt} = \left(\frac{1-\alpha}{\gamma+1-\alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt})^{\frac{\sigma_A-1}{\sigma_A}} \bar{T}_r^{\frac{\sigma_A-1}{\sigma_A}} H_{rAt}^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)\frac{\sigma_A-1}{\sigma_A}}.$$

Using the expression in equation (OA.11) yields

$$\begin{aligned} h_{rt}^A \bar{\chi}_{rt} &= h_{rt}^A \left(\frac{1-\alpha}{\gamma+1-\alpha} \right)^{(1-\alpha)\frac{\sigma_A-1}{\sigma_A}} (\mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt})^{\frac{\sigma_A-1}{\sigma_A}} \bar{T}_r^{\frac{\sigma_A-1}{\sigma_A}} (\Gamma_\zeta L_r h_{rt}^A)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)\frac{\sigma_A-1}{\sigma_A}} \\ &\quad \times \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)\frac{\sigma_A-1}{\sigma_A}(\zeta-1)}. \end{aligned}$$

Defining

$$(OA.13) \quad \mathcal{A}_{rt}^{HC} \equiv \left[\left(h_{rt}^A \right)^{1-\alpha} \left(\frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)} \right]^{\frac{\sigma_A-1}{\sigma_A}}$$

yields

$$(OA.14) \quad \frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} = \left(\frac{\mathcal{A}_{rt}^{HC}}{\bar{w}_{rt}} \right)^{\frac{1}{\kappa}}.$$

where

$$\kappa = \zeta - (\zeta - 1) (1 - \alpha) \frac{\sigma_A - 1}{\sigma_A}.$$

Equations (OA.10), (OA.12), and (OA.14) imply that average earnings \bar{w}_{rt} are given by

$$(OA.15) \quad 1 = \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^\zeta + \left(\frac{h_{rt}^M w_{rMt}}{\bar{w}_{rt}} \right)^\zeta = \left(\left(\frac{\mathcal{A}_{rt}^{HC}}{\bar{w}_{rt}} \right)^{\frac{1}{\kappa}} \right)^\zeta + \left(\frac{\mathcal{Z}_{rt}^{HC}}{\bar{w}_{rt}} \right)^\zeta.$$

Similarly, note that

$$h_{rt}^A \bar{\chi}_{rt} = \mathcal{A}_{rt}^{HC} \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)\frac{\sigma_A-1}{\sigma_A}(\zeta-1)} = \mathcal{A}_{rt}^{HC} \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{\bar{w}_{rt}} \right)^{-(\kappa-1)} = \mathcal{A}_{rt}^{HC} s_{rAt}^{-\frac{\kappa-1}{\zeta}},$$

where the last equality uses equation (OA.11). Hence,

$$\frac{s_{rA}}{1-s_{rA}} = \left(\frac{h_{rt}^A \bar{\chi}_{rt}}{h_{rt}^M w_{rMt}} \right)^\zeta = \left(\frac{\mathcal{A}_{rt}^{HC}}{\mathcal{Z}_{rt}^{HC}} \right)^\zeta s_{rA}^{-(\kappa-1)}$$

so that

$$(OA.16) \quad \frac{s_{rA}^\kappa}{1-s_{rA}} = \left(\frac{\mathcal{A}_{rt}^{HC}}{\mathcal{Z}_{rt}^{HC}} \right)^\zeta.$$

Given the human capital adjusted revenue productivity terms \mathcal{A}_{rt}^{HC} and \mathcal{Z}_{rt}^{HC} , equations (OA.15) and (OA.16) are the same equations as in Proposition 3. This then yields the following result:

Proposition 6. *Consider the model with human capital and let Z_{rt}^{HC} and A_{rt}^{HC} denote productivity in both sectors. The equilibrium allocations in this model are the same as the ones in our baseline model without human capital if we redefine, in our baseline model,*

$$Z_{rt}^{Base} \equiv (h_{rt}^M)^{\frac{\sigma_M}{\sigma_M-1}} Z_{rt}^{HC} \quad \text{and} \quad A_{rt}^{Base} = (h_{rt}^A)^{1-\alpha} A_{rt}^{HC}.$$

Proof. Given data on s_{rA} and \bar{w}_r , we can uniquely solve for \mathcal{A}_{rt}^{HC} and \mathcal{Z}_{rt}^{HC} . We now show that given these data and the other structural parameters there is a one-to-one mapping between \mathcal{A}_{rt}^{HC} and \mathcal{Z}_{rt}^{HC} and Z_{rt}^{Base} and A_{rt}^{Base} . The fact that $Z_{rt}^{Base} \equiv h_{rt}^M Z_{rt}^{HC}$ follows immediately from equation (OA.12), which implies that

$$\begin{aligned} \mathcal{Z}_{rt}^{HC} = h_{rt}^M \mathcal{Z}_{rt} &= h_{rt}^M \left(\frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} \left(Z_{rt}^{HC} \right)^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \\ &= \left(\frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} \left((h_{rt}^M)^{\frac{\sigma_M}{\sigma_M-1}} Z_{rt}^{HC} \right)^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M} \\ &= \left(\frac{1}{\tilde{f}_E} \right)^{1/\sigma_M} \left(Z_{rt}^{Base} \right)^{(\sigma_M-1)/\sigma_M} \mathcal{D}_{rMt}^{1/\sigma_M}. \end{aligned}$$

To derive the expression for A_{rt}^{Base} note that (see equation OA.13)

$$\begin{aligned} \mathcal{A}_{rt}^{HC} &= \left[\left(h_{rt}^A \right)^{1-\alpha} \left(\frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt}^{HC} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)} \right]^{\frac{\sigma_A-1}{\sigma_A}} \\ &= \left[\left(\frac{1-\alpha}{\gamma+1-\alpha} \right)^{1-\alpha} \mathcal{D}_{rAt}^{\frac{1}{\sigma_A-1}} A_{rt}^{Base} \bar{T}_r^\alpha (\Gamma_\zeta L_r)^{-\left(\alpha+\frac{1}{\sigma_A-1}\right)} \right]^{\frac{\sigma_A-1}{\sigma_A}}. \end{aligned}$$

Furthermore, it can be shown that \mathcal{D}_{rAt} and \mathcal{D}_{rMt} also only depend on A_{rt}^{Base} and Z_{rt}^{Base} conditional on the data s_{rA} and \bar{w}_r and the other structural parameters. \square

B.6 Cross-sectional Estimates of the Engel Elasticity η

Targeting the time series of the agricultural employment share implied an estimate of $\eta = 0.8$. However, our model also implies a log-linear relationship between individuals' expenditure share on agricultural products and their total expenditure that can be used to estimate η from cross-sectional microdata:

$$(OA.17) \quad \ln \vartheta_A(y, P_{rM}) = \ln(-\nu P_{rM}^{-1}) - \eta \ln y,$$

where we used our estimate $\phi \approx 0$.

We use the 1936 Consumer Expenditure Survey (CEX) by the US Bureau of Labor Statistics obtained from the Inter-university Consortium for Political and Social Research (ICPSR) to provide direct evidence on the log-linear relationship between expenditure shares and total expenditure.²⁹ The CEX contains micro data on expenditure of individuals on a large variety of categories and a swath of individual characteristics. We use the household files. We obtain information on households' total expenditure, expenditure on food, urban/rural status, size, interview data, occupation and industry of household head, race of household head, and county of residence.

In Figure OA.2, we show that this log-linear relationship is a good description of the data. In the left panel, we display the cross-sectional distribution of food shares. Empirically this variation is substantial, ranging from 5% to 80%. In the right panel, we show the empirical relationship between log expenditure and log food shares as a binned scatter plot. As implied by our theory, the elasticity between food shares and expenditure is indeed essentially constant across the entire range of the distribution of expenditure.

The slope coefficient falls in between $\eta \in (0.315, 0.362)$ depending on which additional controls are chosen, implying that our "macroestimate" of $\eta = 0.79$ is higher than the microestimate that exploits cross-sectional variation.

B.7 Elasticity of Substitution

The Allen-Uzawa elasticity of substitution is given by

$$EOS = \frac{\frac{\partial^2 e(P_{rA}, P_{rM}, V)}{\partial P_{rA} \partial P_{rM}} e(P_{rA}, P_{rM}, V)}{\frac{\partial e(P_{rA}, P_{rM}, V)}{\partial P_{rA}} \frac{\partial e(P_{rA}, P_{rM}, V)}{\partial P_{rM}}},$$

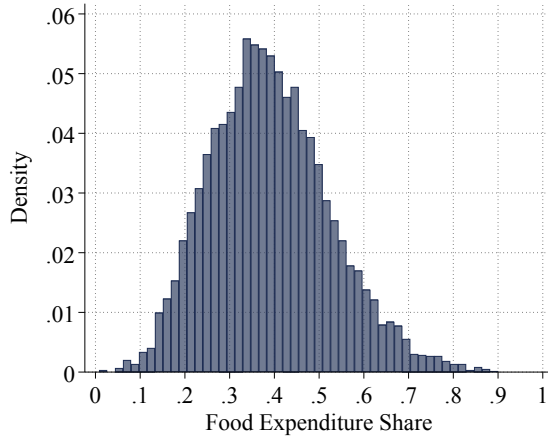
where $e(P_{rA}, P_{rM}, V)$ is the expenditure function given by

$$(OA.18) \quad e(P_{rA}, P_{rM}, V) = (V + \nu \ln(P_{rA}/P_{rM}))^{1/\eta} \eta^{1/\eta} P_{rA}^\phi P_{rM}^{1-\phi}.$$

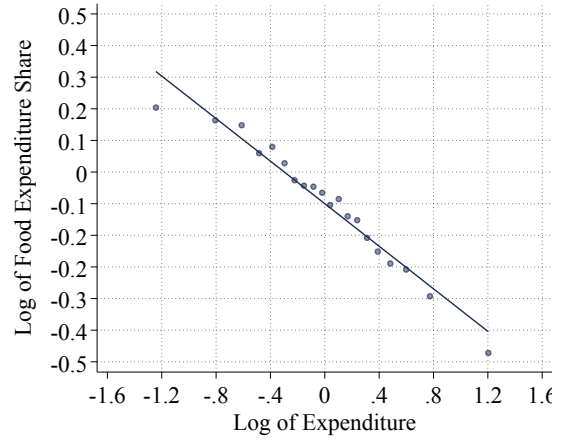
²⁹We note that our theory is written in terms of value added. The expenditure data is in terms of final expenditure data. Herrendorf, Rogerson, and Valentinyi (2013) show that in general there is no direct mapping between the preference parameters of the value added and the final good demand system. However, Fan et al. (2023) show that for the class of PIGL preferences used here, the Engel elasticity η is portable between the final good and value added demand system.

FIGURE OA.2: HETEROGENEITY IN FOOD EXPENDITURE SHARES

(A) FOOD EXP. SHARE DENSITY



(B) FOOD EXP. SHARES AND EXP. LEVEL



Notes: The left panel shows the cross-sectional distribution of individual expenditure shares on food. The right panel plots the binned scatterplot of the (log) food expenditure share against (log) total expenditure, controlling for location and family size fixed effects.

Using equation (OA.18), one can show that

$$EOS = 1 - \eta \frac{(\vartheta_A - \phi)(\vartheta_M - (1 - \phi))}{\vartheta_A \vartheta_M} = 1 + \eta \frac{(\vartheta_A - \phi)^2}{\vartheta_A (1 - \vartheta_A)}.$$