Online Appendices for “Migration, Specialization, and Trade: Evidence from Brazil’s March to the West”

Not for Publication

Heitor S. Pellegrina and Sebastian Sotelo*

Table of Contents

*Email: hpelleg3@nd.edu and ssotelo@umich.edu.

List of Tables

List of Figures

OA Data

This section describes how we construct our data set. Appendix Tables O.1 and O.2 present summary statistics of the key aggregate variables.

Employment and Migration. Our data on migration and employment come from decadal demographic and economic censuses administered by the Brazilian statistical institute IBGE (*Instituto Brasileiro de Geografia e Estatistica*). For 1970, 1980, 1991, 2000 and 2010, we have micro-level data on migration and employment. For 1950 and 1960, we digitized state-level aggregates from historical census publications. The information in the micro-level data is divided in two questionnaires: one applied to the universe of the population, which asks basic questions about education and family structure, and one applied to a sample of households, which asks detailed information on migration, employment and income. In 1970 and 1980, 25 percent of the population was sampled for the detailed questionnaire. For 1990, 2000 and 2010, about 25 percent of the population was sampled in smaller municipalities and 10 percent in larger ones. In this detailed questionnaire, we observe both total income, which includes transfers from the government, and income from a worker's main activity, which includes any source of income, in addition to wages. In our analysis, we use the latter.

For the census of 1980 onward, we observe the current and the previous municipality of residence of each individual, if they have migrated within the previous 10 years.[34] We use this variable to define migration in the reduced form elasticities that we presented in Section 4, but as we show later in this appendix, our results are robust to alternative measures of migration. Since less than 0.1 percent of

[34] The exception is the census of 2000, which asks individuals their previous state of residence, their municipality of residence in 1995, but not their previous municipality of residence.

Brazil's population was born abroad, we remove international migrants from our sample. For 1970, we have micro data with information on the state of birth and the state of residence. For 1950 and 1960, we use information on the total population in each state who were born in each state of Brazil, and also information on total employment per economic activity. In our structural calibration, we use the state of residence and the state of birth to measure the flow of workers between states.

For 1980, 1990, 2000 and 2010, we use our micro data to construct the migration flows of workers who were born in state s and live in state s' and work in activity k, which we denote by $L_{ss'k,t}$. For 1950, we observe only the migration flows from a state s to a state s', which is given by $L_{ss',t}$. Given that we do not directly observe $L_{ss'k,t}$ in 1950, we therefore use entropy methods widely applied in the construction of input-output matrices to obtain $L_{ss'k,t}$. Specifically, we apply the algorithm developed in Ireland and Kullback (1968). In our case, the algorithm consists in searching for values of $L_{ss'k,t}$ based on a guessed value $\tilde{L}_{ss'k,t}$ until $L_{ss'k,t}$ are consistent with aggregate data on $L_{s'k,t}$ and $L_{s,t}$. We use data on $L_{ss'k,1980}$ (adjusting the overall population size to the one in 1950) as guesses of $\tilde{L}_{ss'k,1950}$ to construct the values of $L_{ss'k,1950}$.[35]

To account for population growth, we normalize the population in Brazil to 1 in 1950 and 1920, and to 44 in the rest of the world, according to data from the World Bank. For 1950 onward, we keep the population in the rest of the world fixed at 44, but we adjust Brazil's population so that the ratio of the population in Brazil, relative to the rest of the world, matches the data. That gives us a population of 1.18 in 1980 and 1.22 in 2010. The difference in fertility rates between migrants and non-migrant families is small: according to the census of 2010, migrants families have on average 0.02 additional children, so we impose the same fertility decisions for all households. We impute the rest of the world's land endowment each year (1950, 1980, 2010) so that the relative land-labor ratio between Brazil and the rest of the world, i.e., $(H_{BR,t}/L_{BR,t})\,/\,(H_{F,t}/L_{F,t})$ is reproduced by the model in each year. Within the rest of the world, we allocate labor proportionally to value added across sectors.

Gross Output. To construct gross output and value added per meso-region and activity, we apply a procedure that ensures that our aggregates are consistent with the ones measured in dollars by FAO and United Nations. We compute shares of value added by meso-region and activity and multiply such values by aggregate values from FAO and UN. When necessary, we apply the algorithm developed in Ireland and Kullback (1968) to construct more disaggregated values.

Specifically, our data on agricultural revenues come from PAM (*Produção Agrícola Municipal*), which is maintained by the Brazilian census bureau. PAM provides municipality level data since 1974 for more than 20 crops and state level data since 1930s for a subset of crops. For cattle, we use data from the agricultural census.[36] We converted the revenues measured in these data sets into value added and computed the share of value added coming from each agricultural activity for each meso-region. We

[35] In 1950 we do not observe labor employment in soybeans. We therefore complemented our numbers with special reports from EMBRAPA on historical production of soybeans.

[36] All data on agricultural revenues come from *Produção Agrícola Municipal* (PAM) and the agricultural census, which are based on surveys. They are therefore not generated by imputation.

then multiplied these shares of value added by agricultural activity—given at the meso-region level—by the share of value added in agriculture coming from each meso-region relative to the total value added in agriculture in Brazil, which is measured by IPEA. Lastly, we multiply the share of value added by meso-region and agricultural activity by the aggregate value added in agriculture measured in dollars from the UN. For the manufacturing and service sectors, we use the share of value added for each meso-region measured by IPEA, we then multiply such shares by the aggregate value added measured by UN. Lastly, with data on value added, we computed gross output using the share of value added in the World Input-Output Database.

For 1950, we do not observe value added by economic activity at the meso-region level, but only aggregates at the state-level. As with migration, we use the entropy method developed in Ireland and Kullback (1968) to adjust our values. We search for values of value added $VA_{jk,t}$ based on guesses of value added $\tilde{VA}_{jk,t}$ until they are consistent with observed value added by state and activity $VA_{sk,t}$ and value added by meso-region $VA_{j,t}$. We use the values from 1980 as guesses of $\tilde{VA}_{jk,t}$ for 1950, disciplining the adjustment based on the aggregate values at the state-level.

Trade Flows. The data on trade flows by agricultural activity come from FAO. The data is disaggregated by good, according to the Harmonized System at the 6 digit level. We classified the trade flows according to the agricultural activities included in our analysis. We focused on the unprocessed versions of each good. For example, for tobacco, we excluded manufactured cigars and, for wheat, we excluded pastry related goods. Since the data from FAO starts in 1960, we extrapolate exports and imports back using data on aggregate exports and imports from IPEA.

For 2010, we use export data by state from Comexstat, a website organized by the Ministry of Development, Industry and Foreign Trade (MDIC). For each state, we observe how much was exported and imported from abroad. According to MDIC, the trade data at the state level is registered according to the location of production. For domestic trade flows, we digitized data on trade flows between states from the annual statistical yearbook reports from the Brazilian government of 1947, 1948, 1949, 1972, 1973 and 1974. For 1999, we use estimates of trade flows between states from Vasconcelos (2001) based on state merchandise and services taxes.

Travel Distance. To construct travel distance between regions in each year, we rely on data from the government regarding the expansion of highways in Brazil.[37] To accomplish this, we implement the Fast Marching Method, which is explained in detail in Allen and Arkolakis (2014). Initially, we divide Brazil into a grid and categorize each square based on the availability of roads. We assume that travelling through a grid with a paved road is possible at a speed of 60 km/h, on a non-paved road at an average speed of 30 km/h, and through a grid with no road at an average speed of 15 km/h. Our results are sensible, as evidenced by a comparison with Google Maps. For example, the driving distance between the meso-regions of São Paulo and Rio de Janeiro is 5.8 hours in 2010 in our data, compared to 6

[37] After the 1950s, highways were the predominant (and often the only) mode of transportation in most of Brazil.

hours according to Google Maps. Similarly, the driving distance between the meso-region of São Paulo and Cuiabá in the western state of Mato Grosso is 24 hours according to our data, while Google Maps estimates it to be 21 hours.

Land Settlements - INCRA. The National Institute for Colonization and Agrarian Reform (INCRA) is responsible for monitoring all government-funded land settlement projects. It maintains records of various details, including the settlement area, location, creation year, and assignment year (even for settlements established prior to INCRA's establishment in 1970). Our analysis involves computing the proportion of total land within each meso-region that has been assigned to families through federal land settlement projects. In certain regions, particularly in the West, this proportion can be as high as 0.40. Panel (a) in Online Appendix Figure O.7 presents the distribution of land settlements in the East and in the West.

Suitability. Our measure of suitability exploits the data from FAO-GAEZ, described in greater detail in Farrokhi and Pellegrina (2023). This data provides, for each crop, information on the suitability of land in a given region for two distinct technology classes: one that is intensive in the use of inputs such as fertilizers, seeds, and machinery, and another that is less reliant on these inputs. For instance, Bustos et al. (2016) use the differences between the land suitability scores for these two types of technologies to predict the produticitivty growth of soybean and corn cultivation in Brazil during the 2010s. Given that our analysis covers a longer time span (dating back to at least the 1970s), during which certain regions may have adopted more traditional methods while others have embraced modern techniques, we take the following approach: We run regressions of $\bar{Q}_{sk} = f_k\left(Q_{sk}^{FAO-H}, Q_{sk}^{FAO-L}\right) + \epsilon_{sk}$, where f_k is a non-parametric function of the FAO-GAEZ data, Q_{sk}^{FAO-H} is the suitability measure for crop k in high suitability (measure in total quantity achievable of a given crop in region i), Q_{sk}^{FAO-L} is the respective suitability for crop k using low input methods, and $\bar{Q}_{sk}$ is the average total output between 1950 and 2010 in region s in crop k—where a region is a state. We then project $\bar{Q}_{sk}$ and use these projections as our measures of suitability based on estimates of f_k—i.e., we recover $\widehat{\bar{Q}_{sk}} = \widehat{f}_k\left(Q_{sk}^{FAO-H}, Q_{sk}^{FAO-L}\right)$ and define $S_{sk} = \widehat{\bar{Q}_{sk}}$.

OB Robustness Checks on Fact 2

Alternative Specifications. In Online Appendix Table O.6, we experiment with the initial date of our sample. Earlier initial dates leave fewer years for the construction of the instrument. Our main specification starts with 1980, which strikes a balance between more data for the regression and enough years since 1950 to construct the instruments, but the table shows our results largely remain with different intial dates. In addition, the table shows results change little when we control for bilateral factors in the "zero stage" regression (equation A.1).

Alternative Explanatory Variables. Online Appendix Table O.8 shows that our conclusions remain largely the same if we employ alternative right hand side variables in the estimation of our main equation (2). To be more specific, we can substitute S_{sk} by the total farming population, revenue, or quantity produced in a given origin, time and activity.

Alternative Outcomes. Online Appendix Table O.9 presents the impact of migration composition on alternative outcomes. It shows that the impact on revenues per worker is similar to that on quantity per workers. Importantly for our argument, we find that the impact of migrants' composition on prices is negative, which is consistent with the idea that a "better" composition of migrants reduces the cost of producing a given good. We also interpret this as evidence that migrants' influence the comparative advantage not because of their access to better market opportunities, but due to better knowledge about production methods.

Specific Crops do not drive the results. Online Appendix Table O.7 examines whether any single crop is driving our results. To do so, we estimate equation (A.1) dropping every individual crop, separately, from our regression. In addition, the last column shows our results when we drop capital-intensive crops.

Meso-region level regressions. We now present estimates of equation (2), but using meso-region level data. Specifically, we estimate:

$$\ln\left(\frac{Q_{ik,t}}{L_{ik,t}}\right) = \delta_{k,t} + \delta_{i,t} + \kappa \ln\left(\sum_{i' \neq i} \frac{M_{i'i,t}}{M_{i,t}} S_{i'k}\right) + X'_{ik,t}\beta + u_{ik,t}.$$

There are two challenges with running this regression at the meso-region level. First, we are unable to construct the historical panel with migration flows between any two destination and origin regions, since we do not observe bilateral migration flows for 1950, 1960, and 1970 at the meso-regional level. Second, we are unable to compute the stock of workers ($M_{i'i,t}$ and $M_{i,t}$) who migrated from a region i to a destination i', all the way back to the 1950s, since we only observe a workers' previous meso-region, and not their meso-region of birth. (At the state-level, we observe the state of birth of an individual, which allows us to compute the accumulated number of people who migrated to a given destination.) Note that we do construct $S_{i'k}$ at the meso-region level for the regressions. We therefore run:

$$M_{i'ik,t} = \delta^0_{i',t} + \delta^0_{j,t} + \sum_{\tau=1980}^{t} \alpha^0_{\tau,t} \times \left(I^{-r(s)}_{i',\tau} \frac{I^{-r(i')}_{i,\tau}}{I^{-r(i')}_{\tau}}\right) + X'_{i'i}\beta^0_t + u^0_{i'ik,t}. \tag{O.1}$$

Online Appendix Table O.12 presents results. Due to lack of statistical power, we are unable to add the same rich set of controls as in our baseline specification. Yet, despite the substantially different approach for the construction of the specifications, results are qualitatively similar.

OC Robustness Checks on Fact 3

Geographic units of analysis and definition of migration. Online Appendix Table O.21 replicates Table 3 using information at the state level, with which we can modify the definition of migration, since we observe the state of birth of an individual—this is a departure from our analysis in the main body of the paper in which we use a worker's previous region, which is the definition of migration that we can establish at the meso-region level since we do not observe the meso-region of birth.

Role of specific crops. Online Appendix Table O.22 shows that no single crop drives the results that we obtain. It reports PPML and IV estimates using RAIS and Census data dropping every individual crop. In doing so, we also demonstrate our effects remain when we drop capital-intensive crops.

Socio-economic status and previous networks. Focusing on the earnings equation (3), Online Appendix Table O.23 columns (2) and (5) show that our results hold when we control for previous networks, as measure by the total population who migrate from origin i to the same destination j to produce k in $t-1$. Additionally, it controls for socio-economic status in columns (3) and (6).

Individual level regressions. Online Appendix Table C.4 report results on the earnings equation (3) when we use individual level data, instead of aggregating the data at the cell level as we do in our main analysis in the body of the paper. In particular, with the individual level data we can control more flexibly for socio-economic status (column 3) and for the time spent in a given meso-region (column 4).

Manufacturing sector. We check if the relationships that we uncovered for the case of agriculture are also found more generally in other sectors of the economy. Online Appendix Table O.25 shows that if we run the same regressions with workers employed in manufacturing activities, we find similar patterns in the data.

RAIS data and Experience. We ran our specifications using data from RAIS to check whether our results are also present in alternative datasets. Online Appendix Tables O.13, O.17, O.18, O.19, O.20, O.24, and O.23 report results using the RAIS dataset. We notice that the right hand side variable is still based on the census, since they are lagged 30 years and it is not possible to obtain such measures from RAIS, which is only available with information on the crop of production after the 1990s. In particular, Online Appendix Table O.24 shows that the mechanisms that we study in the paper are still present for workers without any experience in a given crop.

Composition of workers in the origin. Online Appendix Table O.26 shows our results are not driven by a correlation between the number and the productivity (quality) of workers in the origin. In all specifications, we add as controls the average productivity of workers as measured by the output per worker and the revenue per worker. In addition, in the IV specifications, we control for $\sum_{i\neq j}\mu_{ji,t-1}^{-1}\left(Q_{jk,t-1}/N_{jk,t-1}\right)$

and $\sum_{i \neq j} \mu_{ji,t}^{-1} \left(R_{jk,t-1}/N_{jk,t-1}\right)$, where $Q_{jk,t}$ and $R_{jk,t}$ are the output in quantity and value, respectively, and $N_{jk,t}$ is the total worker employment. The goal is to control, as allowed by observables, for the quality of workers. Reassuringly, our results are, if anything, stronger when we add these variables as controls.

Non-parametric relationship. Online Appendix Figure O.1 presents local polynomial regressions of farmers and income that correspond to regressions (3) and (4). We first absorb origin-destination-year and destination-crop-year fixed effects; we then run a polynomial regressions on the residuals of the dependent and independent variables of interest in equations (3) and (4).

OD Constructing the Policy Counterfactuals

Section (7.2) discusses the effects of major observable government policies on comparative advantage. This section provides details on our approach to measure the impact of these policies.

OD.1 Counterfactual Expansion of Highways

We first estimate, using the migration costs $\tilde{\mu}_{ss,t}$ backed out from the model:

$$\ln \tilde{\mu}_{ss',t} = \alpha_0 + \alpha_{1,t} dist_{ss',t} + \epsilon_{ss',t} \tag{O.2}$$

using the instrumental variable approach developed in Morten and Oliveira (2016), which exploits the radial highways constructed to connect the new capital of Brazil, Brasília, constructed during the 1950s. We then construct counterfactual migration costs based on

$$\ln \tilde{\mu}_{ss',t}^{CF} = \widehat{\alpha}_{1,t} \left(dist_{ss',t}^{CF} - dist_{ss',t} \right) + \ln \tilde{\mu}_{ss',t}$$

where $dist_{ss',t}^{CF}$ is the travel time estimates when there is no expansion of highways in the west. Online Appendix Figure (O.5) present the baseline presence of highways and the expansion of highways without the expansion towards the West.

OD.2 Counterfactual Land Settlement

To construct our counterfactual measures of productivity growth in land supply, we run the following regression, using change in the land supply productivity we back out from the model, $\Delta \ln g_{i,t}$:

$$\Delta \ln g_{i,t} = \alpha_{0,t} + \alpha_{1,t} \ln LS_i + \epsilon_{ss',t} \tag{O.3}$$

where LS_i is the share of land in meso-region i that has been acquired through land redistribution programs from the federal government. We then recover counterfactual growth in land productivity between 2010 and 1950 and between 1980 and 1950 based on

$$(\Delta \ln g_{i,t})^{CF} = \widehat{\alpha}_{1,t} \left(\ln LS_i^{CF} - \ln LS_i \right) + \Delta \ln g_{i,t}$$

where $\widehat{\alpha}_{1,t}$ is the OLS estimates of equation (O.3) and $\ln LS_{i,t}^{CF}$ is the counterfactual expansion of land settlements. To construct LS_i^{CF}, we proceed as follows. For every region in the West, we compute $LS_i^{CF} = LS_i \times \left(\bar{LS}^E / \bar{LS}^W \right)$ where $\bar{LS}^E$ is the average share of land settlements distributed in the East and $\bar{LS}^W$ is the average share of land settlements in the West. For regions in the East, $LS_i^{CF} = LS_i$. Panel A of Online Appendix Figure O.7 shows the distribution of land settlements in the East and in the West—it indicates a substantially larger distribution in the West. The panel also shows the counterfactual expansion of land. Panel (b) shows that the expansion of land settlements is well correlated with the productivity growth in the land supply in the West, as backed out by the model.

OE The Non-Quantitative Model from Section 5

This section presents the steady-state equations of our non-quantitative model, defined by equations (5) to (22).

OE.1 The steady state equilibrium

A steady-state equilibrium is a competitive equilibrium in which the sequence of geographies is constant $\Gamma_t = \Gamma$, and all endogenous variables are constant. It is defined by the following set of equations:

$$c_{jk} = \gamma_k^{\gamma_k} (1-\gamma_k)^{1-\gamma_k} \left(w_{jk}^{1-\gamma_k} r_j^{\gamma_k} \right)^{\alpha_k} P_j^{1-\alpha_k} \tag{O.4}$$

$$p_{jk} = \left(\sum_i (c_{ik}\tau_{ijk}/A_{ik})^{1-\eta_k} \right)^{\frac{1}{1-\eta_k}} \tag{O.5}$$

$$\pi_{ijk} = \left(\frac{c_{ik}\tau_{ijk}/A_{ik}}{p_{jk}} \right)^{1-\eta_k} \tag{O.6}$$

$$A_{ik} = \bar{A}_{ik} \ell_{ik}^{\xi} \tag{O.7}$$

$$P_j = \prod_{k=1}^{K} p_{jk}^{a_k} \tag{O.8}$$

$$W_{ijk} = \frac{w_{jk} s_{ik}}{P_j} \tag{O.9}$$

$$\lambda_{ijk} = \left(\frac{W_{ijk}/\mu_{ijk} \cdot \upsilon_j^{\delta}}{\upsilon_i} \right)^{\theta} \tag{O.10}$$

$$\upsilon_i = \exp(\Upsilon_i) \tag{O.11}$$

$$s_{ik} = \bar{s}_k \ell_{ik}^{\beta} \tag{O.12}$$

$$E_{ik} = \sum_{i'} s_{i'k} \lambda_{i'ik} L_{i'} \tag{O.13}$$

$$\ell_{ik} = \sum_{i'} \lambda_{i'ik} L_{i'} \tag{O.14}$$

$$Y_{jk} = \sum_i \pi_{jik} a_k X_i \tag{O.15}$$

$$w_{ik} E_{ik} = \alpha_k (1-\gamma_k) Y_{ik} \tag{O.16}$$

$$r_i H_i = \sum_k \alpha_k \gamma_k Y_{ik} \tag{O.17}$$

$$\bar{L} = \sum_{i,k} \ell_{ik} \tag{O.18}$$

$$\Upsilon_i = \frac{1}{\theta} \ln \left(\sum_{j,k} \left(W_{ijk}/\mu_{ijk} \cdot \upsilon_j^{\delta} \right)^{\theta} \right) \tag{O.19}$$

$$X_j = \sum_k w_{j,k} E_{j,k} + r_j H_j + \sum_{k=1}^{K} (1-\alpha_k) Y_{j,k} \tag{O.20}$$

OF Characterization of the Equilibrium

This section characterizes the equilibrium of our model. First it discusses uniqueness of equilibria in special cases of the steady state. Then it draws implications of the knowledge externality for growth and welfare.

OF.1 The Steady State in a Single Sector Model with Many Regions

This section maps the steady state of our model to the set of equations in Allen and Donaldson (2022) and applies the sufficient conditions given there. We consider a version of our full model described in Online Appendix OL, in which $K = 1$, production only requires value added ($\alpha = 1$), and labor is the only factor of production ($\gamma = 0$). In that case, similarly to Allen and Donaldson (2022), we can write the system of equations for the steady state as follows, where—differently from our model—we define $W_i = \bar{u}_i L_i^{-\chi} \dfrac{w_i}{P_i}$ to facilitate derivations:

1. Total payments to labor satisfy

$$w_i^{\sigma} E_i L_i^{\xi(1-\sigma)} = \sum_j \left(\tau_{ij}/\bar{A}_i \bar{u}_j\right)^{1-\sigma} w_j^{\sigma} \left(W_j\right)^{1-\sigma} L_j^{\chi(1-\sigma)} E_j \tag{O.21}$$

2. Trade is balanced

$$w_i^{1-\sigma} L_i^{\chi(\sigma-1)} \left(W_i\right)^{\sigma-1} = \sum_j \left(\tau_{ji}/\bar{u}_i \bar{A}_j\right)^{1-\sigma} L_j^{\xi(1-\sigma)} w_j^{1-\sigma} \tag{O.22}$$

3. A location's population is equal to the population arriving in that location

$$L_i = \sum_j \mu_{ji}^{-\theta} \upsilon_j^{-\theta} V_i^{\theta} L_j \tag{O.23}$$

4. A location's population equals the number of people leaving that region[38]

$$\upsilon_i^{\theta} = \sum_j \mu_{ij}^{-\theta} V_j^{\theta} \tag{O.24}$$

5. Agents are forward looking

$$V_j = W_j \upsilon_j^{\delta} \tag{O.25}$$

[38]Because we have a single sector, knowledge does not affect migration decisions, since knowledge increases proportionally the income of a worker regardless of her destination region. For that reason, the knowledge component, $s_i = \bar{s} L_i^{\beta}$, does not appear in equation (O.24).

6. Efficiency units of labor is given by

$$E_i = \sum_j \mu_{ji}^{-\theta} \upsilon_j^{-\theta} V_i^{\theta} L_j^{1+\beta} \tag{O.26}$$

As shown in Allen and Arkolakis (2014) and Allen and Donaldson (2022), assuming trade costs and migration are symmetric makes origin and destination fixed effects proportional. This gives the following two equations:

$$\underbrace{\left(w_i/\bar{A}_i L_i^{\xi}\right)^{1-\sigma}}_{\text{origin FE}} = \kappa^T \underbrace{\left(w_i E_i / P_i^{1-\sigma}\right)}_{\text{destination FE}} \tag{O.27}$$

and

$$\underbrace{\left(\upsilon_i\right)^{-\theta} L_i}_{\text{origin FE}} = \kappa^M \underbrace{\left(V_i\right)^{\theta}}_{\text{destination FE}}, \tag{O.28}$$

where κ^M and κ^T are constants that do not affect the equilibrium properties of the system.

Using equations (O.27) and (O.28), we can reduce the system (O.21)-(O.26) to the following 3 equations in E_i/L_i, L_i, and W_i:

$$\left(E_i/L_i\right)^{\tilde{\sigma}} \left(L_i\right)^{\tilde{\sigma}(1+\chi\sigma+\xi(1-\sigma))} \left(W_i\right)^{\tilde{\sigma}\sigma} = \sum_j K_{ij}^1 \left(W_j\right)^{\tilde{\sigma}(1-\sigma)} \left(L_j\right)^{\tilde{\sigma}(1+\chi(1-\sigma)+\xi\sigma)} \left(E_j/L_i\right)^{\tilde{\sigma}} \tag{O.29}$$

$$\left(E_i/L_i\right) \left(L_i\right)^{\frac{1}{1+\delta}} \left(W_i\right)^{-\frac{\theta}{1+\delta}} = \sum_j K_{ij}^2 \left(W_j\right)^{\frac{\theta}{1+\delta}} L_j^{\frac{\delta}{1+\delta}+\beta} \tag{O.30}$$

$$\left(L_i\right)^{\frac{1}{1+\delta}} \left(W_i\right)^{-\frac{\theta}{1+\delta}} = \sum_j K_{ij}^2 \left(W_j\right)^{\frac{\theta}{1+\delta}} \left(L_j\right)^{\frac{\delta}{1+\delta}} \tag{O.31}$$

where $\tilde{\sigma} \equiv (\sigma-1)/(2\sigma-1)$ and K_{ij}^1 and K_{ij}^2 capture the constants of the model, $\bar{A}_i$, $\bar{u}_i$, τ_{ij}, and μ_{ij}.

To apply their Propostion 2 from Allen and Donaldson (2022), we compute the matrix of the LHS coefficients in equations (O.29) to (O.31)

$$\Gamma = \begin{pmatrix} \tilde{\sigma} & \tilde{\sigma}(1+\chi\sigma+\xi(1-\sigma)) & \tilde{\sigma}\sigma \\ 1 & \frac{1}{1+\delta} & -\frac{\theta}{1+\delta} \\ 0 & \frac{1}{1+\delta} & -\frac{\theta}{1+\delta} \end{pmatrix}$$

and also the matrix of the RHS coefficients

$$A = \begin{pmatrix} \tilde{\sigma} & \tilde{\sigma}(1+\chi(1-\sigma)+\xi\sigma) & (1-\sigma)\tilde{\sigma} \\ 0 & \frac{\delta}{1+\delta}+\beta & \frac{\theta}{1+\delta} \\ 0 & \frac{\delta}{1+\delta} & \frac{\theta}{1+\delta} \end{pmatrix}.$$

These two matrices allow for a change of variables in the system of equations (O.29) to (O.31) required to apply Proposition 2. Specifically, the proposition states that, if the spectral radius $\rho\left(\left|A\Gamma^{-1}\right|\right)$ is strictly smaller than one, we have a unique equilibrium. We find that the spectral radius increases with β, indicating that it contributes to multiplicity of equilibria. Based on our values of χ, ξ, δ, and setting $\sigma = 5.5$, the spectral radius is 1.04, so that the sufficient condition is not satisfied.

Note, in addition, the following two properties of the system. If $K_{ij}^2 = 1$, $\forall i, j$, equations O.30 and O.31 imply that E_i/L_i is equalized across all regions i. The system is then reduced to two equations in W_i and L_i, which does not depend on β. That is, when there are no migration costs, accumulated knowledge spreads evenly across all locations (due to the preference shocks), and therefore cannot impact locations differentially. We conclude that migration costs are necessary for our knowledge mechanism to affect the uniqueness properties of the equilibrium.

OF.2 The Steady State in a Two-sector, One-region Economy

We next study the role of sectors by studying economies that feature two-sectors, but lack geography. Start with the model described in the main body of the paper (see details in Online Appendix OL). Suppose there are two sectors, denoted k and k', and a single region. Again, we assume no intermediate inputs ($\alpha = 0$) and only labor ($\gamma = 0$), but keeping the agglomeration externalities, $\xi > 0$. We will rearrange the demand and supply conditions of the model.

The labor market equilibrium for sector k in period t is as follows

$$w_{kt}E_{kt} = \left(\frac{p_{kt}}{P_t}\right)^{1-\sigma}\sum_{k'} w_{k't}E_{k't},$$

which we rewrite as:

$$w_{kt}\lambda_{kt}\bar{s}\lambda_{kt-1}^{\beta} = \frac{p_{kt}^{1-\sigma}}{p_{kt}^{1-\sigma}+p_{k't}^{1-\sigma}}\left\{w_{kt}\lambda_{kt}\bar{s}\lambda_{kt-1}^{\beta} + w_{k't}\lambda_{k't}\bar{s}\lambda_{k't-1}^{\beta}\right\},$$

where the efficient units of labor is $E_{kt} = \ell_{kt}\bar{s}\ell_{kt-1}^{\beta}$ and $\ell_{kt} = \lambda_{kt}\bar{L}$. Substituting the price $p_{kt} = w_{kt}\lambda_{kt}^{-\xi}A_{kt}^{-1}$, setting $\lambda_{kt} = \lambda_{kt-1} = \lambda_{k,SS}$, and rearranging shows how labor allocations relate to wages on the demand side on the steady state

$$\lambda_{k,SS}^{1+\beta+\xi(1-\sigma)} = \frac{w_{k,SS}^{-\sigma}A_{k,SS}^{\sigma-1}}{w_{k,SS}^{1-\sigma}A_{k,SS}^{\sigma-1}\lambda_{k,SS}^{-\xi(1-\sigma)} + w_{k',SS}^{1-\sigma}A_{k',SS}^{\sigma-1}\lambda_{k',SS}^{-\xi(1-\sigma)}}\left\{w_{k,SS}\lambda_{k,SS}\bar{s}\lambda_{k,SS}^{\beta} + w_{k',SS}\lambda_{k',SS}\bar{s}\lambda_{k',SS}^{\beta}\right\}.$$

Turning to the labor supply decisions, we obtain:

$$\lambda_{k,t} = \frac{\left(w_{k,t}\lambda_{k,t-1}^{\beta}\right)^{\theta}}{\left(w_{k,t}\lambda_{k,t-1}^{\beta}\right)^{\theta} + \left(w_{k',t}\lambda_{k',t-1}^{\beta}\right)^{\theta}},$$

which after rearranging and in the steady state, yields

$$\lambda_{k,SS}^{1-\theta\beta} = \frac{w_{k,SS}^{\theta}}{\left(w_{k,SS}\lambda_{k,SS}^{\beta}\right)^{\theta} + \left(w_{k',SS}\lambda_{k',SS}^{\beta}\right)^{\theta}}.$$

To study the equilibrium, we define $\tilde{\lambda} = \lambda_{k,SS}/\lambda_{k',SS}$, $\omega = w_{k,SS}/w_{k',SS}$ and $a = A_{k,SS}/A_{k',SS}$ and rewrite the system as

$$\tilde{\lambda} = \omega^{-\frac{\sigma}{1+\beta+\xi(1-\sigma)}} a^{\frac{\sigma-1}{1+\beta+\xi(1-\sigma)}}$$
$$\tilde{\lambda} = \omega^{\frac{\theta}{1-\theta\beta}}.$$

A weak set of sufficient conditions for uniqueness is that the demand elasticity be negative, while that of supply is positive, which guarantees they only cross once. Focusing on this set of conditions is also useful because it allows us to understand the distinct role of β in each. We state the result in this simple lemma:

Lemma 1. *If both* $\frac{\sigma}{1+\beta+\xi(1-\sigma)} > 0$ *and* $\frac{\theta}{1-\theta\beta} > 0$*, then the equilibrium is unique. Moreover,* $0 < \tilde{\lambda} < 1$ *and* $0 < \omega < 1$.

To understand the condition on demand, note that the efficient number of workers is a power $1+\beta$ of the simple count, $\lambda_{k,SS}$. Thus, a large value of β makes the marginal product of labor in sector k decrease more strongly when the raw count of workers there increases. Note that, for $\sigma > 1$, this effect is operates against that of the externality ξ, which instead makes the marginal product of labor less responsive to increases in employment. Thus, for $\xi = 0$, the condition is met for any value of β, while if $\xi > 0$, $\beta > \xi(\sigma - 1)$ guarantees that demand is downward sloping..

To understand the condition on supply, note that in the steady state, the supply of workers to sector k increases in wages but also to supply of workers to sector k itself (due to $\beta > 0$). With this cumulative effect, the elasticity of supply increases in β, so long as $\beta < 1/\theta$.

Because this simple version lacks geography, we verify that both conditions are met given our parameter values, using $\sigma = 0.5$. Note, however, that they would still be verified if we set chose $\sigma = 5.5$, our Armington elasticity.

An example of multiplicity. We close this section of the appendix by constructing an example in which multiplicity arises. To do so, we pick the following values $\sigma = 2$, $\beta = 1$, $\theta = 3$, $A_1 = 1.5$ and $A_2 = 1$. Importantly, we set $\xi = 0$, which means that multiplicity must come exclusively from knowledge accumulation. As shown in Figure O.4, Panel (a) besides the interior equilibrium, equilibria with zero labor allocation in a sector are possible. Note that, at these values, $\theta/(1-\theta\beta) < 0$, which violates the sufficient conditions given above.

Among these extreme equilibria, one equilibrium is characterized by $\lambda_1 = 0$, $W = A_2 = 1$. The other is characterized by $\lambda_1 = 1$,$W = A_1 = 1.5$, which is preferable because the economy specializes in a high-productivity activity.

Figure O.4, Panel (b) shows that the multiplicity disappears for $\beta = 0$, which respects the conditions met above.

OF.3 The Effect of Population Growth on Welfare levels

This section shows that that population changes affect the level of welfare (not its growth). Additionally, we do a simple calculation to provide a sense of how much welfare levels can improve in response to observed population growth. We then confirm quantitatively that these impacts are relatively small.

Proposition 2. *Consider the steady state equilibrium of the model described in Section 5 (Online Appendix OE contains all the equations). Consider an increase in total population of $\bar{L}' = \psi \bar{L}$. Then, real wages increase by a factor of $\psi^{\tilde{\beta}}$ and expected welfare increases by a factor of $\psi^{\frac{\tilde{\beta}}{1-\delta}}$, where $\tilde{\beta} = \beta - (1+\beta)\gamma + \xi/\alpha$.*

Proof.

We begin with a guess, which we verify later: $L_i^{'} = \psi L_i$, $\lambda_{ijk}' = \lambda_{ijk}$, $\pi_{ijk}' = \pi_{ijk}$, and $Y_{ik}' = \psi^{1+\beta} Y_i$. Using our guesses for λ_{ijk}' and L_i' and equation O.14, we obtain

$$\ell_{ik}' = \psi \ell_{ik}.$$

Using this finding, together with equation O.12 and our guess of L_i', in equation (O.13),

$$E_{ik}' = \psi^{1+\beta} \sum_j s_{jk} \lambda_{jik} L_j.$$

Using this result, and our guess for Y_{ik}', in equation (O.15), we verify our guess that $w_{ik}' = w_{ik}$. Furthermore, using equation (O.17), we conclude that $r_i' = \psi^{1+\beta} r_i$. Evaluating equation (O.20) at the counterfactual values of $w_{ik}' E_{ik}'$, r_i', and Y_{jk}', we conclude that $X_{jk}' = \psi^{1+\beta} X_{jk}$.

We next determine c_{ik}', p_{ik}', and P_i'.[39] Equations (O.4), (O.5), and (O.8), together with our results for r_i' and A_{ik}', imply that

$$\begin{aligned} c_{ik}' &= \psi^{(1+\beta)\gamma - \xi \frac{1-\alpha}{\alpha}} c_{ik} \\ p_{ik}' &= \psi^{(1+\beta)\gamma - \frac{\xi}{\alpha}} p_{ik} \\ P_i' &= \psi^{(1+\beta)\gamma - \frac{\xi}{\alpha}} P_i. \end{aligned}$$

With these results at hand, equation (O.6) implies $\pi_{ijk}' = \pi_{ijk}$.

Finally equations (O.9) and (O.12) imply[40]

$$W_{ijk}' = \psi^{\tilde{\beta}} W_{ijk},$$

[39] To obtain this result, guess that $c_{ik}' = \psi^a c_{ik}$, which holds if $a = (1+\beta)\gamma - \xi \frac{1-\alpha}{\alpha}$.

[40] To obtain this result, guess that $\upsilon_i' = \psi^b \upsilon_i$, which holds if $b = \tilde{\beta}/(1-\delta)$.

while equation (O.19) and (O.11) imply

$$\upsilon_i' = \psi^{\frac{\tilde{\beta}}{1-\delta}} \upsilon_i,$$

where

$$\tilde{\beta} = \beta - (1+\beta)\,\gamma + \xi/\alpha.$$

Next, we verify our guess for λ_{ijk}'. Using equation (O.10), and previous results

$$\lambda_{ijk}' = \left(\frac{\psi^{\tilde{\beta}} \psi^{\delta\tilde{\beta}/(1-\delta)} W_{ijk}/\mu_{ijk} \upsilon_j^{\delta}}{\psi^{\frac{\tilde{\beta}}{1-\delta}} \upsilon_i} \right)^{\theta} = \lambda_{ijk}.$$

Finally, we verify the proportional growth of population in each region:

$$\begin{aligned} L_i' &= \sum_j \lambda_{jik}' L_j \\ &= \sum_j \lambda_{jik} \psi L_j = \psi L_i, \end{aligned}$$

and we note that this satisfies the adding up constraint (O.18).

Discussion. Notice that it is possible for $\tilde{\beta} < 0$. The reason is that in this model land is in fixed supply. As population grows, workers become more productive, both through knowledge accumulation, β, and external economies, ξ/α. However, land becomes more expensive, $(1+\beta)\,\gamma$.

Examples. Having characterized this result and verified our initial guesses, we move on to provide a back-of-the-envelop calculation of the impact of population growth on welfare and real wages, before moving on to our full quantification. We consider two cases: i) land is not a factor of production, $\gamma = 0$ and ii) a case where land is used as in the agricultural sector $\gamma = 0.2$. In both cases, we consider a population growth factor of $\psi = 1.16$, which equals the accumulated population growth over 30 years at current growth rates.

Case 1: $\gamma = 0$

In this case,

$$\tilde{\beta} = 0.05 + 0.06/0.5 = 0.17$$

For real wages, our results imply

$$W_{ijk}'/W_{ijk} = 1.16^{0.17} \approx 1.025,$$

whereas for welfare

$$\upsilon_i'/\upsilon_i = 1.6^{\frac{0.17}{1-0.172}} \approx 1.031.$$

Thus, with 16 percent population growth, the next generation experiences 1.025 higher real wages, and 1.031 higher expected welfare.

Case 2: $\gamma > 0$

In this case, we adopt $\beta = 0.05$,$\gamma = 0.2$, $\xi = 0.06$, $\alpha = 0.5$. Then

$$\begin{aligned}\tilde{\beta} &= 0.06 - (1.05) \cdot 0.2 + 0.06/0.5 \\ &= -0.04.\end{aligned}$$

For real wages, our results imply

$$W_{ijk}'/W_{ijk} = 1.16^{-0.04} \approx 0.994,$$

whereas for welfare

$$\upsilon_i'/\upsilon_i = 1.6^{\frac{-0.04}{1-0.17}} \approx 0.993.$$

Thus, with 16 percent population growth, the next generation experiences about 0.6 percent lower real wages and 0.7 percent lower expected welfare.

Using the full quantitative model (described in Online Appendix OL, we simulate a 16 percent increase in population in steady state, and we find an average drop of 0.22 percent in wages and 0.34 percent in welfare. The key differences between our quantitative model and the two examples above are i) there are endogenous congestion forces, ii) land is only used in agriculture, iii) input-output linkages and value added share vary across sectors, iv) knowledge externalities only operate in agriculture.

Because, in all these cases, the growth responses are quantitatively small, we do not place emphasis on them in the main body of the paper.

OF.4 Geography, access to knowledge, and knowledge accumulation

This section characterizes how geography shapes a region's human capital accumulation in equilibrium, starting again from the model described in Section 5 (Online Appendix OE contains all the equations). It contains two main results. The first result is a direct consequence of Section OF.3: Skill accumulation and Endogenous growth generated by population growth is independent of geography. In particular, let

$$e_{ik,t} \equiv \frac{E_{ik,t}}{\ell_{ik,t}} \tag{O.32}$$

be human capital per worker in region i, industry k. The results in Section OF.3 show that in a steady state with larger population (by a factor of $\Lambda > 1$), effective labor force per worker (i.e. knowledge) is larger by a factor of Λ^{β}:

$$
\begin{aligned}
e'_{ik}/e_{ik} &= \frac{E'_{ik}/\ell'_{ik}}{E_{ik}/\ell_{ik}} \\
&= \Lambda^{\beta},
\end{aligned}
$$

which does not depend on i and is therefore independent of geography. Our assumed functional form for migration plays an important role, since there are no complementarities at the ijk level, other than the migration costs. This result shows that forces that might generate growth in the long run, such as population growth, do not interact with geography.

In contrast, our second result describes access to human capital for a region i, industry k as a function of geography, at any time t. Using equations (16) and (17) in equation (O.32), we obtain

$$
e_{ikt} = \sum_{j} \frac{\mu_{jikt}^{-\theta} \upsilon_{jt}^{-\theta} \left[\bar{s}_k \ell_{jk,t-1}^{\beta}\right]^{1+\theta} L_{jt-1}}{\sum_{j'} s_{j'kt}^{\theta} \mu_{j'ikt}^{-\theta} \upsilon_{j't}^{-\theta} L_{j't-1}},
$$

where the variables on the right-hand side do not pertain to i, with the exception of moving costs.

This expression formalizes two key ideas. First, region i characteristics do not shape its access to human capital from the rest of the country (including i itself). The reason is that local wages, costs of living, and expected welfare of children are not relatively more important to high $s_{j,k}$ workers from other regions j; in other words, they do not generate additional sorting into region i, activity k. Since this is true for every region j, more attractive regions do not attract disproportionately high skill workers. In particular, agglomeration economies, which operate by raising local wages, do not affect skill accumulation, since wages drop from this expression above. The second result is that geography does matter for access to human capital from other regions. If region i is relatively closer (than other regions i') to regions with i) a production structure that favors industry k, ii) lower expected opportunities for children, and iii) larger population, then its knowledge per worker will be larger. To see this clearly, note that assuming $\mu_{jik,t} = \mu_{ji',kt}, \forall i, i', j$,

$$
\frac{e_{ikt}}{e_{i'kt}} = 1,
$$

which shows that, absent geography, access to human capital is the same everywhere. This result shows that taste shocks are a dispersion force that leads to all regions benefiting equally from migrants coming from all other regions, unless geography limits migration.

OG The Small Open Economy Limit

We start by studying equilibrium prices when the Home economy is small, by adapting the procedure in Alvarez and Lucas (2007) to the case in which Home is a collection of regions. We begin with the model described in Section 5 in the main body of the paper, under the following parameter restrictions: $\eta_k = \eta$, $\alpha_k = \alpha$ and $\gamma_k = \gamma$, $\forall k$. We use the following assumptions, for each region in Home: (i) $L_{i,t-1} \to 0$, (ii) $\bar{A}_{ik,s}^{\eta-1}/L_{i,t-1}^{1-\xi(\eta-1)+\beta} \to \delta_{ik,s}$, where $\delta_{ik,s} \in (0,\infty)$, (iii) $H_{i,s}/L_{i,t-1}^{1+\beta} \to h_{i,s}$, where $h_{i,s} \in (0,\infty)$, (iv) $L_{i,t-1}/L_{i',t-1} \to \iota_{ii'}$, $\forall s = t, t+1, \ldots$ and (v) $\delta_{ik,s}$ and $h_{i,s}$ converge to constants as $s \to \infty$. Note that we place assumptions on exogenous objects or those that are predetermined at $t-1$. We then use the equilibrium conditions to characterize the evolution of endogenous objects for $s = t, t+1, \ldots$.

Assuming that wages and rents, as well as migration shares are well defined in the limit, i.e. $w_{ik,s} \in (0,\infty)$, $r_{ik,s} \in (0,\infty)$, and $\lambda_{ijk,s} \in (0,1)$—which we verify below—, the equilibrium price indexes for each region and sector is

$$p_{ik,t} = \left(\sum_j \left(c_{jk,t}\tau_{jik,t}\right)^{1-\eta} \bar{A}_{jk,t}^{\eta-1} \ell_{jk,t}^{\xi(\eta-1)} \right)^{1-\eta}$$

and assumptions (i) and (ii), imply that

$$p_{ik,t} \to c_{Fk,t}\tau_{Fik,t} / \left(\bar{A}_{Fk,t} \ell_{Fk,t}^{\xi} \right), \forall i, k.$$

where $c_{Fk,t} = \bar{c}_k w_{Fk,t}^{\alpha(1-\gamma)} r_{F,t}^{\gamma\alpha} P_{F,t}^{1-\alpha}$, and $w_{Fk,t}$, $r_{F,t}$, $P_{F,t}$, $\ell_{Fk,t}$ solve the labor and land market clearing conditions for Foreign and the definition of the price index in our limiting scenario. In what follows, we take $w_{Fk,t}$, $r_{F,t}$, $p_{Fk,t}$, and P_{Ft} as given.

We now characterize the equilibrium wages, rental rates, migration shares, labor supplies, and expected welfare for each region at Home. Doing so entails appropriately scaling endogenous outcomes at each time $s = t, t+1, \ldots$, such that the equilibrium conditions are well defined in the limit, and such that we can finally show prices are positive but finite (thus confirming in our guess).

Begin by noting that, given the guess that migration shares are well behaved, then labor supplies relative to regional size, $\ell_{ik,s-1}/L_{i,s-1}$, are well defined

$$\frac{\ell_{ik,s}}{L_{i,s-1}} = \sum_j \lambda_{jik,s} \cdot \frac{L_{j,s-1}}{L_{i,s-1}}, \tag{O.33}$$

when the relative population sizes, $L_{j,s-1}/L_{i,s-1}$ are well defined. Accounting for labor then states,

$$\frac{L_{i,s}}{L_{i',s}} = \frac{\sum_k \left\{ \sum_j \lambda_{jik,s} \cdot \frac{L_{j,s-1}}{L_{i,s-1}} \right\}}{\sum_k \left\{ \sum_j \lambda_{ji'k,s} \cdot \frac{L_{j,s-1}}{L_{i',s-1}} \right\}}, \tag{O.34}$$

which is also well defined so long as the $L_{j,s-1}/L_{j',s-1}$ terms are, $\forall j, j'$. Equations (O.33) and (O.34) can be applied sequentially, starting at $s = t$, together with assumption (iv) to show that $\ell_{ik,s}/L_{i,s-1}$ and $L_{i,s}/L_{i',s}$ are well defined for any $s = t, t+1, \ldots$. Similar logic shows that the ratios $L_{i,s}/L_{i,s-1}$, $\ell_{ik,s}/L_{i,s}$, and $\ell_{ik,s}/\ell_{ik',s}$ are also well defined.

Next we verify our guess of wages and rental rates. To do so, we will use labor and land market clearing in each period. As an intermediate step, equation (16) and the definition of $s_{ik,t}$ suggest that the appropriate scaling for effective labor supply is as follows

$$\frac{E_{ik,s}}{L_{i,s-1}^{1+\beta}} = \sum_{j} \lambda_{jik,s} \bar{s}_k \left(\frac{\ell_{jk,s-1}}{L_{j,s-1}} \right)^{\beta} \cdot \left(\frac{L_{j,s-1}}{L_{i,s-1}} \right)^{1+\beta}, \tag{O.35}$$

given that we have already established that each term on the right-hand side is well defined.

Turning to the labor market clearing condition (21), we obtain

$$w_{ik,s}^{1+\alpha(1-\gamma)(\eta-1)} \left(\frac{E_{ik,s}}{L_{i,s-1}^{1+\beta}} \right) \left(\frac{\ell_{ik,s}}{L_{i,s-1}} \right)^{\xi(1-\eta)} = \alpha (1-\gamma) \left(r_{i,s}^{\alpha\gamma} P_{i,s}^{1-\alpha} \right)^{1-\eta} \frac{\bar{c}_k^{1-\eta} \bar{A}_{ik,s}^{\eta-1}}{L_{i,s-1}^{1+\beta-\xi(\eta-1)}} \left(\frac{\tau_{iFk,s}}{P_{Fk,s}} \right)^{1-\eta} X_{Fk,s}. \tag{O.36}$$

which, together with assumption (ii) and equations (O.33)-(O.35), gives an equilibrium condition for $w_{ik,s}$ and verifies that it is strictly positive and finite.

In turn, from land market clearing we obtain

$$r_{i,s}^{1+(\eta-1)\alpha\gamma} \frac{H_{i,s}}{L_{i,t-1}^{1+\beta}} = \alpha\gamma P_{i,s}^{(1-\alpha)(1-\eta)} \ldots \tag{O.37}$$

$$\times \left[\sum_{k} w_{ik,s}^{\alpha(1-\gamma)(1-\eta)} \left(\frac{\ell_{ik,s}}{L_{i,s-1}} \right)^{\xi(\eta-1)} \left(\frac{\bar{c}_k^{1-\eta} \bar{A}_{ik,s}^{\eta-1}}{L_{i,s-1}^{1+\beta-\xi(\eta-1)}} \right) \left(\frac{L_{i,s-1}^{1+\beta}}{L_{i,t-1}^{1+\beta}} \right) \left[\left(\frac{\tau_{iFk,s}}{P_{Fk,s}} \right)^{1-\eta} X_{Fk,s} \right] \right],$$

which, together with assumption (iii), gives an equilibrium condition for $r_{i,s}$ and shows that it is strictly positive and finite.

Finally, from equations (O.19) and (O.11), steady state we can scale υ_i by $L_i^{\frac{\beta}{1-\delta}}$:

$$\left(\frac{\upsilon_i}{L_i^{\frac{\beta}{1-\delta}}} \right)^{\theta} = \sum_{j,k} \left\{ \left(\frac{w_{jk}}{P_j \mu_{ijk}} \right)^{\theta} \frac{\ell_{ik}^{\beta\theta}}{L_i^{\beta\theta}} \frac{L_j^{\frac{\beta}{1-\delta}\theta\delta}}{L_i^{\frac{\beta}{1-\delta}\theta\delta}} \left[\frac{\upsilon_j}{L_j^{\frac{\beta}{1-\delta}}} \right]^{\theta\delta} \right\}, \tag{O.38}$$

which defines a system in which the scaled values of υ_i are positive and finite, given that we have characterized the behavior of ℓ_{ik}/L_i and L_j/L_i above. We proceed likewise for $\upsilon_{i,s}$:

$$\left(\frac{\upsilon_{i,s}}{L_{i,s}^{\frac{\beta}{1-\delta}}} \right)^{\theta} = \sum_{j,k} \left(\frac{w_{jk,s} \bar{s}_k \left(\ell_{ik,s-1}/L_{i,s} \right)^{\beta}}{P_{j,s} \mu_{ijk,s}} \right)^{\theta} \frac{1}{L_{i,s}^{\frac{\beta\theta}{1-\delta}\delta}} L_{j,s+1}^{\frac{\beta}{1-\delta}\theta\delta} \left(\frac{\upsilon_{j,s+1}}{L_{i,s+1}^{\frac{\beta}{1-\delta}}} \right)^{\theta\delta} \tag{O.39}$$

which, applying recursively, an bearing in mind equation (O.38), shows that all appropriately scaled $\upsilon_{i,s}$ are positive and finite.

Using the last result, we confirm that migration shares are well defined. Using (15), we can show that

$$\lambda_{ijk,s} = \frac{\left(\frac{w_{jk,s}}{P_{j,s}\mu_{ijk,s}}\right)^{\theta} \ell_{ik,s-1}^{\beta\theta} \upsilon_{j,s+1}^{\delta\theta}}{\upsilon_{i,s}^{\theta}}$$

$$= \frac{\left(\frac{w_{jk,s}}{P_{j,s}\mu_{ijk,s}}\right)^{\theta} \left(\frac{\ell_{ik,s-1}}{L_{i,s-1}}\right)^{\theta\beta} \left(\frac{\upsilon_{j,s+1}}{L_{j,s}^{\beta/(1-\delta)}}\right)^{\delta\theta}}{\left(\frac{\upsilon_{i,s}}{L_{i,s-1}^{\beta/(1-\delta)}}\right)^{\theta}} \left[\left(\frac{L_{j,s}}{L_{j,s-1}}\right)^{\frac{\delta}{1-\delta}}\right]^{\beta\theta},$$

and since each term in the right-hand side has an appropriate limit, so do the shares $\lambda_{ijk,s}$. This completes the proof.

Note that the case of $\delta = 0$, is much simpler, since it is not necessary to characterize the limit of υ_i via equations (O.38) and (O.39).

OG.1 Proof of Proposition 1

We seek to understand the impact of changes in migration costs, $\mu_{ijk,s} \to \infty$, on the direction of trade

$$\frac{X_{iFk,t}/X_{iFl,t}}{X_{FFk,t}/X_{FFl,t}},$$

for region i within a small open economy. Using our model, changes in the direction of trade respond to changes in wages and scale (through agglomeration economies)

$$\frac{\widehat{X_{iFk,t}/X_{iFm,t}}}{X_{FFk,t}/X_{FFm,t}} = \frac{\hat{w}_{ik,t}^{(1-\eta)(1-\gamma)\alpha} \cdot \hat{\ell}_{ik,t}^{\xi(\eta-1)} / \left(\hat{w}_{im,t}^{(1-\eta)(1-\gamma)\alpha} \hat{\ell}_{im,t}^{\xi(\eta-1)}\right)}{\hat{w}_{Fk,t}^{(1-\eta)(1-\gamma)\alpha} \hat{\ell}_{Fk,t}^{\xi(\eta-1)} / \left(\hat{w}_{Fm,t}^{(1-\eta)(1-\gamma)\alpha} \hat{\ell}_{Fm,t}^{\xi(\eta-1)}\right)}.$$

We proceed in steps to construct both terms. Applying hat algebra to equation (15), we obtain the counterfactual changes in migration shares:

$$\hat{\lambda}_{ijk,s} = \frac{\left(\widehat{W_{ijk,s}/\mu_{ijk,s}} \hat{\upsilon}_{j,s+1}^{\delta}\right)^{\theta}}{\hat{\upsilon}_{i,s}^{\theta}}$$

$$= \frac{\left(\hat{w}_{jk,s}\hat{\ell}_{ik,s-1}^{\beta}/\hat{\mu}_{ijk,s}\hat{\upsilon}_{j,s+1}^{\delta}\right)^{\theta}}{\upsilon_{i,s}^{\theta}}$$

$$= \begin{cases} 0 & \text{if } i \neq j \\ \frac{\left(\hat{w}_{jk,s}\hat{\ell}_{ik,s-1}^{\beta}\hat{\upsilon}_{j,s+1}^{\delta}\right)^{\theta}}{\hat{\upsilon}_{i,s}^{\theta}} & \text{otherwise} \end{cases} \tag{O.40}$$

Next, equation (O.35) implies

$$\frac{\widehat{E_{ik,s}}}{L_{i,s-1}^{1+\beta}} = \frac{\lambda_{iik,s}\left(\frac{\ell_{ik,s-1}}{L_{i,s-1}}\right)^{\beta}}{\sum_{j'}\lambda_{j'ik,s}\left(\frac{\ell_{j'k,s-1}}{L_{j',s-1}}\right)^{\beta}\cdot\left(\frac{L_{j',s-1}}{L_{i,s-1}}\right)^{1+\beta}}\widehat{\lambda_{iik,s}}\left(\widehat{\frac{\ell_{ik,s-1}}{L_{i,s-1}}}\right)^{\beta}, \tag{O.41}$$

while equation (O.33) implies

$$\widehat{\frac{\ell_{ik,s}}{L_{i,s-1}}} = \frac{\lambda_{iik,s}}{\sum_{j'}\lambda_{j'ik,s}\cdot\frac{L_{j',s-1}}{L_{i,s-1}}}\widehat{\lambda_{iik,s}}. \tag{O.42}$$

Next, the labor market clearing condition (O.36) implies, taking the ratio relative to good m

$$\frac{\widehat{w_{ik,s}}^{1+\alpha(1-\gamma)(\eta-1)+\theta}\left(\mathcal{E}_{iik,s}\left(\hat{\ell}^{\beta}_{ik,s-1}\right)^{\theta}\left(\widehat{\frac{\ell_{ik,s-1}}{L_{i,s-1}}}\right)^{\beta}\right)\left(\widehat{\frac{\ell_{ik,s}}{L_{i,s-1}}}\right)^{\xi(1-\eta)}}{\widehat{w_{im,s}}^{1+\alpha(1-\gamma)(\eta-1)+\theta}\left(\mathcal{E}_{iim,s}\left(\hat{\ell}^{\beta}_{im,s-1}\right)^{\theta}\left(\widehat{\frac{\ell_{im,s-1}}{L_{i,s-1}}}\right)^{\beta}\right)\left(\widehat{\frac{\ell_{im,s}}{L_{i,s-1}}}\right)^{\xi(1-\eta)}} = 1, \tag{O.43}$$

where

$$\mathcal{E}_{iik,s} \equiv \frac{\lambda_{iim,s}\bar{s}_k\left(\ell_{im,s-1}\right)^{\beta}\cdot\left(L_{i,s-1}\right)}{\sum_{j'}\lambda_{j'im,s}\bar{s}_k\left(\ell_{j'm,s-1}\right)^{\beta}\cdot\left(L_{j',s-1}\right)}$$

is the share of the effective labor supply in i, k that comes from i itself.

From equation (O.42) and (O.40):

$$\frac{\widehat{\frac{\ell_{ik,s}}{L_{i,s-1}}}}{\frac{\ell_{im,s}}{L_{i,s-1}}} = \frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\frac{\left(\hat{w}_{ik,s}\left(\widehat{\frac{\ell_{ik,s-1}}{L_{i,s-1}}}\right)^{\beta}\right)^{\theta}}{\left(\hat{w}_{im,s}\left(\widehat{\frac{\ell_{im,s-1}}{L_{i,s-1}}}\right)\right)^{\theta}},$$

where

$$\mathcal{L}_{iik,s} \equiv \frac{\lambda_{iik,s}\cdot L_{i,s-1}}{\sum_{j'}\lambda_{j'ik,s}\cdot L_{j',s-1}}$$

is the share of raw labor in i, k that comes from i itself. Thus, substituting in (O.43)

$$\frac{w_{ik,s}^{1+\alpha(1-\gamma)\widehat{(\eta-1)}+\theta+\theta\xi(1-\eta)}\left(\mathcal{E}_{iik,s}\left(\hat{\ell}^{\beta}_{ik,s-1}\right)^{\theta}\left(\widehat{\frac{\ell_{ik,s-1}}{L_{i,s-1}}}\right)^{\beta}\cdot\right)}{w_{im,s}^{1+\alpha(1-\gamma)\widehat{(\eta-1)}+\theta+\theta\xi(1-\eta)}\left(\mathcal{E}_{iim,s}\left(\hat{\ell}^{\beta}_{im,s-1}\right)^{\theta}\left(\widehat{\frac{\ell_{im,s-1}}{L_{i,s-1}}}\right)^{\beta}\cdot\right)}\left(\frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iik,s}}\frac{\left(\hat{\ell}^{\beta}_{ik,s-1}\right)^{\theta}}{\left(\hat{\ell}^{\beta}_{im,s-1}\right)^{\theta}}\cdot\right)^{\xi(1-\eta)} = 1.$$

Given $\hat{\ell}_{ik,t-1} = 1\ \forall k$ (because it is predefined) and given the baseline exposure shares $\{\mathcal{E}_{iik,s}, \mathcal{L}_{ii,s}\}$, the two equations above provide a system of equations for the sequence of relative wages and relative labor allocations. In particular, note that when $s = t$, the relative wages are given by

$$\frac{\hat{w}_{ik,t}}{\hat{w}_{im,t}} = \left[\frac{\mathcal{E}_{iik,s}}{\mathcal{E}_{iim,s}}\left(\frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\right)^{\xi(1-\eta)}\right]^{-\frac{1}{1+\alpha(1-\gamma)(\eta-1)+\theta+\theta\xi(1-\eta)}},$$

and changes in labor specialization are

$$\frac{\widehat{\ell_{ik,t}}}{\ell_{im,t}} = \frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\frac{(\hat{w}_{ik,s})^{\theta}}{(\hat{w}_{im,s})^{\theta}}.$$

Therefore, changes in specialization are given by

$$\frac{X_{iFk,t}\widehat{/X}_{iFm,t}}{X_{FFk,t}/X_{FFm,t}} = \left\{\frac{\widehat{w_{ik,t}}}{w_{im,t}}\right\}^{(1-\eta)(1-\gamma)\alpha}\left\{\frac{\widehat{\ell_{ik,t}}}{\ell_{im,t}}\right\}^{\xi(\eta-1)},$$

that is,

$$\frac{X_{iFk,t}\widehat{/X}_{iFm,t}}{X_{FFk,t}/X_{FFm,t}} = \underbrace{\left\{\left[\frac{\mathcal{E}_{iik,s}}{\mathcal{E}_{iim,s}}\left(\frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\right)^{\xi(1-\eta)}\right]^{-\frac{1}{1+\alpha(1-\gamma)(\eta-1)+\theta+\theta\xi(1-\eta)}}\right\}^{(1-\eta)(1-\gamma)\alpha}}_{=\left\{\frac{\widehat{w_{ik,t}}}{w_{im,t}}\right\}^{(1-\eta)(1-\gamma)\alpha}}$$
$$\times\underbrace{\left\{\frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\left[\frac{\mathcal{E}_{iik,s}}{\mathcal{E}_{iim,s}}\left(\frac{\mathcal{L}_{iik,s}}{\mathcal{L}_{iim,s}}\right)^{\xi(1-\eta)}\right]^{-\frac{\theta}{1+\alpha(1-\gamma)(\eta-1)+\theta+\theta\xi(1-\eta)}}\right\}^{\xi(\eta-1)}}_{=\left\{\frac{\widehat{\ell_{ik,t}}}{\ell_{im,t}}\right\}^{\xi(\eta-1)}}$$

To interpret this expression, note that the first term captures how wages influence exports to the rest of the world. The migration shock we study limits the inflow of workers to region i and, therefore, acts like a negative labor supply shock. The strength of this effect depends on how dependent the region was on labor from other regions. The second term captures the changes in productivity that come from the labor supply shock. Notice that this second term appears only when there are external agglomeration effects.

In the special case in which $\xi = 0$, we obtain the result in Proposition 1:

$$\frac{X_{iFk,t}\widehat{/X}_{iFm,t}}{X_{FFk,t}/X_{FFm,t}} = \left[\frac{\mathcal{E}_{iik,s}}{\mathcal{E}_{iim,s}}\right]^{-\frac{\alpha(1-\gamma)(1-\eta)}{1+\theta+\alpha(1-\gamma)(\eta-1)}}.$$

OH Econometric specifications implied by the model

OH.1 Derivation of the estimating equations

In this section we provide the steps to map our model to estimating equations (3) and (4), starting from our quantitative model described in Online Appendix OL.

Obtaining equation (24). Beginning with equation (3), our model states $\text{income}_{ijk,t} = w_{jk,t}s_{ik,t}$, and allowing for a general measurement error $\tilde{u}^I_{ijk,t}$, we obtain:

$$\ln \text{income}_{ijk,t} = \ln s_{ik,t} + \ln w_{jk,t} + \tilde{u}^I_{ijk,t}.$$

We begin with a general functional form for $s_{ik,t}$, and we discuss the restrictions we place on it to achieve identification and to simulate the model. Thus, starting with

$$s_{ik}\left(\ell_{ik,t-1}\right) = \bar{s}_k \bar{s}_i \bar{s}_{ik} \ell^{\beta}_{ik,t-1},$$

our regression becomes

$$\begin{aligned}\ln \text{income}_{ijk,t} &= \ln\left(\bar{s}_{k,t}\bar{s}_{i,t}\bar{s}_{ik,t}\ell^{\beta}_{ik,t-1}\right) + \ln w_{jk,t} + \tilde{u}^I_{ijk,t} \\ &= \underbrace{\bar{u}^I_{ij,t} + \ln \bar{s}_{i,t}}_{\iota^I_{ij,t}} + \underbrace{\ln w_{jk,t} + \ln \bar{s}_{k,t} + \bar{u}^I_{k,t}}_{\iota^I_{jk,t}} + \beta \ln \ell^{\beta}_{ik,t-1} + \underbrace{\ln \bar{s}_{ik,t}\ell^{\beta}_{ik,t-1} + \tilde{u}^I_{ijk,t}}_{\equiv u^I_{ijk,t}}.\end{aligned}$$

Note that our empirical specification contains an indicator at the (ij,t) level, $\iota^I_{ij,t}$, and the (jk,t) level, $\iota^I_{jk,t}$, which allows us to soak up origin-destination and destination-crop components of the measurement error, potentially correlated with $w_{jk,t}$, which we write as $\tilde{u}_{ijk,t} = \bar{u}^I_{ij,t} + \bar{u}^I_{jk,t} + u^I_{ijk,t}$. To the extent that such variation exists in the data, we remove it from the estimation and in our quantitative exercises focus exclusively in what pertains to our theory. This is how one arrives at equation (24).

Identification. OLS identification would require that the intercept of the function $s_{ik}(\cdot)$ be log-separable in to an i- and a k-specific term. Note that this already allows for the intercept to differ across regions and crops, only with the restriction that such variation be log-separable.

Our IV strategies generate variation in $\ell_{ik,t-1}$ that is orthogonal to permanent productivity components, such as $\bar{s}_{ik,t}$. Our instrument based on Burchardi et al. (2019) generates shifts in $\ell_{ik,t-1}$ entirely driven by coincidences in push and pull factors, controlling for i, k fixed effects in the zero stage. Our instrument based on Harris (1954), likewise, generates shifts in $\ell_{ik,t-1}$ based only on proximity to other regions that have high employments in that crop, excluding nearby regions to limit the influence of spatially correlated productivity.

Obtaining equation (25). To obtain (4), we proceed analogously. Begin with the fact that $\ell_{ijk,t} = \lambda_{ijk,t} L_{i,t-1}$. Then

$$
\begin{aligned}
\ln \ell_{ijk,t} &= \ln \lambda_{ijk,t} + \ln L_{i,t-1} + \tilde{u}^L_{ijk,t} \\
&= \theta \ln \left[\frac{w_{jk,t} s_{ik,t} \upsilon^{\delta}_{j,t+1}}{\tilde{P}_{j,t} \mu_{ijk,t} \upsilon_{i,t}} L^{-\chi}_{j,t} \right] + \ln L_{i,t-1} + \tilde{u}^L_{ijk,t} \\
&= \theta \ln w_{jk,t} + \theta \ln s_{ik,t} - \theta \left(\ln \tilde{P}_{j,t} - \chi \ln L_{j,t} \right) - \theta \ln \mu_{ijk,t} \\
&+ \theta \delta \Upsilon_{j,t+1} - \theta \ln \Upsilon_{i,t} + \ln L_{i,t-1} + \tilde{u}^L_{ijk,t} \\
&= \theta \ln w_{jk,t} + \theta \ln \left(\bar{s}_{i,t} \bar{s}_{k,t} \bar{s}_{ik,t} \ell^{\beta}_{ikt-1} \right) - \theta \left(\ln \tilde{P}_{j,t} - \chi \ln L_{j,t} \right) \\
&- \theta \ln \left(\bar{\mu}_{ij,t} + \bar{\mu}_{jk,t} \right) + \theta \delta \Upsilon_{j,t+1} - \theta \ln \Upsilon_{i,t} + \ln L_{i,t-1} + \tilde{u}^L_{ijk,t} \\
&= \underbrace{\theta \ln w_{jk,t} - \theta \left(\ln \tilde{P}_{j,t} - \chi \ln L_{j,t} \right) - \theta \ln \bar{\mu}_{jk,t} + \theta \ln \bar{s}_{k,t} + \theta \delta \Upsilon_{j,t+1}}_{\iota^L_{jk,t}} \\
&+ \underbrace{\bar{u}^L_{ij,t} - \theta \ln \bar{\mu}_{ij,t} + \ln L_{i,t-1} - \theta \ln \Upsilon_{i,t} + \ln \bar{s}_{i,t}}_{\iota^L_{ij,t}} + \theta \beta \ln \ell_{ik,t-1} + \underbrace{\ln \bar{s}_{ik,t} + \tilde{u}^L_{ijk,t}}_{\equiv u^L_{ijk,t}},
\end{aligned}
$$

which corresponds to equation (25). As in deriving equation (24), note that the i-specific and k-specific components of the intercept of the $s(\cdot)$ function are captured by our fixed effects. A more general term $\bar{s}_{ik,t}$ is captured in the error. Again, as before, log-separability ensures proper identification via OLS, while our IV strategies circumvent the problem entirely by generating variation in $\ell_{ik,t-1}$ that is orthogonal to $\bar{s}_{ik,t}$. In deriving the equation, we also assume $\mu_{ijk,t} = \bar{\mu}_{ij,t} \bar{\mu}_{jk,t}$, but we could also allow for measurement error in migration costs to become part of the error in the regression, as we discuss in the paper.

OH.2 A "Fixed Types" Model

This Section explores the implications of a model in which individuals have assigned "types" (i.e. they can only work on one activity) at the time of their location and occupation decision. We show that, if workers sort according to income, or if they sort at random, such a model yields econometric implications that are inconsistent with the results found in Section 4. For simplicity, we focus on the case where $\delta = 0$.

In both sections below, we consider the following common setup. Workers are born with a type k— and associated productivity $s_{i,kt}$— in a number proportional to their parents, $L_{i,kt-1}$. Idiosyncratic shocks are drawn from the same distribution as before, with dispersion parameter θ

OH.3 A "fixed types" model with sorting according to income

OH.3.1 Migration Probability.

Migration probabilities $\tilde{\lambda}_{ij,kt}$ reflect the following problem

$$\max_{j}\left\{\ln\frac{w_{jk,t}s_{ik,t}}{\tilde{P}_{j,t}\mu_{ijk,t}}+\varepsilon_j/\theta\right\},$$

where ε is drawn from a T1EV distribution with shape parameter θ. The solution is:

$$\tilde{\lambda}_{ij,kt}=\frac{W^{\theta}_{ij,kt}\mu^{-\theta}_{ijk,t}}{\tilde{\Xi}^{\theta}_{i,kt}}$$

where, recall, $W_{ijk,t}\equiv\frac{w_{jk,t}s_{ik,t}}{\tilde{P}_{j,t}}$ and $\tilde{\Xi}^{\theta}_{ik,t}\equiv\sum_{j'}W^{\theta}_{ij'k,t}\mu^{-\theta}_{ij'k,t}$.

OH.3.2 Econometric specification.

The income of a worker of type k from i going to region j is given by

$$\text{Income}_{ijk,t}=w_{jk,t}s_{ik,t},$$

Hence, the corresponding regression equation is:

$$\ln\text{Income}_{ijk,t}=\ln w_{jk,t}+\ln s_{ik,t}+u^{I}_{ijk,t},\tag{O.44}$$

where we allow again for measurement error in income.

Likewise, the migration equation states that

$$\ell_{ijk,t}=\tilde{\lambda}_{ijk,t}\ell_{ik,t-1},$$

and the corresponding regression equation is:

$$\ln\ell_{ijk,t}=\kappa\ln w_{jk,t}+\kappa\ln s_{ik,t}-\kappa\ln P_{j,t}-\kappa\ln\mu_{ijk,t}+\ln\ell_{ik,t-1}-\kappa\tilde{\Xi}_{ik,t}.\tag{O.45}$$

Direct comparison of equation (O.44) with the results from (24) shows that worker heterogeneity according to origin location is necessary to reproduce them. Moreover, one needs to impose $s_{ik,t}=\bar{s}_i L^{\beta}_{ik,t-1}$, to replicate our regression. In turn, using this assumption in (O.45) shows that the coefficient of $L_{ik,t-1}$ should be $1+\kappa\beta$, which is at odds with the results from equation (25). In our specification, in which we essentially use one cohort, one cannot separately identify $L_{ik,t-1}$ from the fixed effect that would capture $\tilde{\Xi}_{ik,t}$.

OH.3.3 A "fixed types" model with random sorting

Migration Probability. Migration probabilities $\tilde{\lambda}_{ijk,t}$ reflect the following problem

$$\max_j \{\varepsilon_j\}.$$

The standard solution is:

$$\tilde{\lambda}_{ijk,t} = \frac{1}{I}$$

where I is the number of regions.

Econometric specification. The income of a worker of type k from i going to j is

$$\text{Income}_{ijk,t} = w_{j,kt} s_{ik,t},$$

Hence, the corresponding regression equation is:

$$\ln \text{Income}_{ijk,t} = \ln w_{jk,t} + \ln s_{ik,t} + u^I_{ijk,t}, \tag{O.46}$$

where we allow again for measurement error in income.

Likewise, the migration equation states that

$$\ell_{ijk,t} = \frac{1}{I} \ell_{ik,t-1},$$

and the corresponding regression equation is:

$$\ln L_{ijk,t} = \ln L_{ik,t-1}. \tag{O.47}$$

Equation (O.46) shows that one requires the same conditions as in the previous Section to rationalize the income equation. Moreover, equation (O.46) shows that the coefficient of $L_{ik,t-1}$ should be 1, which is at odds with the results from equation (25).

OI A Simple Microeconomic Foundation

Suppose productivity in the production of crop k reflects the number of crop-specific tasks that workers know how to perform. There are a continuum of such potential tasks, indexed by $t \in [0, \infty]$. Workers either know how to perform these tasks or not, as indicated by $\iota(t) \in \{0,1\}$:

$$s_{k,t} = \int_0^\infty \iota(t)\, dt.$$

When young, a worker has T units of time to meet old workers. As a consequence of the meeting, the young worker learns one task at random, which is specific to the crop the old worker is working on. The total number of meetings between young workers and old workers in sector k is α_k, given by

$$\alpha_{k,t} = \bar{\alpha}_k Y_{i,t}^{1-\gamma} \ell_{ik,t}^{\gamma},$$

where $Y_{i,t}$ is the total size of the young population. This corresponds to the meeting rate of a random search model.

The young workers meets a random number of old, crop-k workers, given by $N_{k,t}$, which is distributed as a Poisson random variable, with parameter $\Phi_{k,t} = \alpha_{k,t} \frac{T}{Y_{i,t}}$. At the end of the day, workers pool all the information they obtained (see, e.g. Porcher, 2022). Because tasks are learned at random and there is a continuum of them, we ignore the possibility that two young workers bring the same task to the pool. Thus, the total amount of tasks learned by the workers in crop k is

$$\Phi_{k,t} Y_{i,t} = \bar{\alpha}_k Y_{i,t}^{1-\gamma} \ell_{ik,t}^{\gamma} T$$

Letting $\gamma = \beta$, $\bar{s}_k = \bar{\alpha}_k$, and $\bar{s}_{i,t} = Y_{i,t}^{1-\gamma} T$, delivers a version of our knowledge function

$$s_{k,t} = \bar{s}_k \bar{s}_{i,t} \ell_{ik,t}^{\beta}.$$

OJ A Model with Lineages

This section introduces a model in which children learn both from the learning externality and directly from their parents. Suppose that a worker's productivity depends on their origin region (as we posit) and on his lineage. In particular, a worker from i who works in activity k, whose parents worked in activity m has productivity

$$s_{ik}^{m} = \bar{s}_k \varsigma_k^m \ell_{ik,t-1}^{\beta}, \tag{O.48}$$

where

$$\varsigma_k^m = \begin{cases} \varsigma > 1 & \text{if } m = k \\ 1 & \text{otherwise} \end{cases} \tag{O.49}$$

Upon migration to region j, this worker's income are

$$\text{income}_{ijk,t}^{m} = w_{jk,t} \bar{s}_k \ell_{ik,t-1}^{\beta} \varsigma_k^m.$$

Since in our data we lack information on the parent's trade, m, we can only observe $\text{income}_{ijk,t}$, which is the average income of workers coming from origin i producing in jk, who potentially come from different lineages. As a result, the following regression

$$\ln \text{income}_{ijk,t} = \iota_{jk,t} + \beta \ln \ell_{ik,t-1} + u_{ik,t}, \tag{O.50}$$

includes the term ς_k^m in the error term $u_{ik,t}$. The worry here is that there is a correlation between employment in the origin $\ell_{ik,t-1}$ and the error, induced by ς_k^m.

Based on this model, we perform the following calibration exercise to see how the presence of lineages in the model affects the need for our mechanism to rationalize the data. Online Table O.27 shows the result. Moving from Columns (1) to (6), we consider values of ς from 1 to 2 (i.e. up to a 100 percent increase in productivity from choosing the same sector as the parent). For each value of ς, we calculate then what is the value of β (the strength of our mechanism) that is required to match a coefficient of 0.05 (in Columns (1) to (3)) and of 0.08 (in Columns (4) to (6)) when we run our earnings regression, described in equation (O.50). When matching a reduced-form coefficient of 0.05, a productivity increase of at least 100 percent is required to eliminate the need for our mechanism. When matching a slightly larger 0.08 coefficient, even a 100 percent increase retains close to half of our value of β.

To examine the plausibility of this value, we refer to previous work that has documented the intergenerational stickiness of sectoral employment, and we find that the simulated shares are substantially larger. For example, Long and Ferrie (2013) calculate using US and UK data (Table 1: 1949-55 to 1972-73) that the fraction of children whose parents are in the same broad sector is 54 percent in the US and 43 percent in the UK. For the 1850s, the corresponding values are 41 percent and 62 percent. The fact that these numbers are relatively stable over time give some confidence that they can be informative for a country at a different level of development, such as Brazil.

To evaluate the plausibility of the lineage model, the last row of the Table O.27 computes the same statistic, i.e., the fraction of workers who remain in the same activity as their parents in our lineage model. For a value of $\varsigma = 0$, our model generates shares that are comparable to those in Long and Ferrie (2013). However, even a small increase in the value of ς (e.g. 0.5), with which more than half of our mechanism remains, produces a substantially larger fraction of children choosing the same sector as their parents. This is despite the fact that our our definition of activity is much narrower than that of Long and Ferrie (2013), which would suggest the fraction we compute should be smaller. Recall, moreover, that matching a reduced-form coefficient of 0.05 is a relatively conservative choice, since that value is in the middle range of our empirical estimates. Matching a slightly larger coefficient of 0.08, the lineage model leaves ample room for our mechanism, while still producing employment fractions that are not aligned with the data. Given the evidence from other countries, we conclude that the lineage model tends to generate what seems like implausible intergenerational stickiness, and we therefore prefer our baseline formulation.

OK Migration and the Gains from Trade

To complete our evaluation of the impact of migration on international trade, we assess how it affects the gains from trade (hereafter, GFT).

For simplicity, we derive these expressions using the model in Section 5 (see Online Appendix OE for all the equations), under the following parameter restrictions: $\delta = 0$, $\gamma_k = 0$, $\forall k$. Given that $\delta = 0$, we focus on a single period, drop time indexes and note that s_{ik} is predetermined.

Under these restrictions, we define the corresponding utility of being born in i as

$$\begin{aligned}\Xi_i &= v_i \\ &= \left[\sum_j \sum_k \left(W_{ijk,t}/\mu_{ijk}\right)^\theta\right]^{1/\theta}.\end{aligned}$$

In our model, a measure of welfare is the expected utility attained by a person born in region i, Ξ_i. Denoting by $\hat{\Xi}_i$ the change in welfare from going to autarky, the GFT are $1 - \hat{\Xi}_i$.

We first show analytically that, in a one-activity model, regional terms of trade and migration opportunities generate new sources of gains from trade. We then show that migration and comparative advantage are key drivers of the GFT in our multi-sector environment. Our results complement the approach in Galle et al. (2023) by bringing in geographic mobility frictions, which do not play a prominent role in their analysis.

OK.1 Gains from Trade with One Activity

We start with a result that, following directly from the definitions of Ξ_i and of trade shares π_{ii}, highlights that changes in expected utility depend on changes in real wages in all regions to which workers can migrate; these changes, in turn, can be computed using observed regional trade shares.

Proposition 3. *Using observed trade shares, one can compute the losses from full trade autarky, i.e.* $\tau_{ij} \to \infty, \forall i \neq j$, *as*

$$\hat{\Xi}_i^{B \to B,A} = \left(\sum_j \lambda_{ij} \pi_{ii}^{\frac{\theta}{\alpha(\eta-1)}}\right)^{1/\theta}, \tag{O.51}$$

where λ_{ij} *are observed migration shares and* π_{ii} *are observed domestic expenditure shares, and "B" denotes baseline and "B, A" denotes trade autarky starting from B.*

Absent migration, i.e. when $\lambda_{ii} = 1$ and $\lambda_{ij} = 0$ for any $j \neq i$, equation (O.51) collapses to the canonical formula for the GFT (see Arkolakis et al., 2012). The next proposition relates the GFT in the baseline to those in a situation where migration across regions is not allowed.

Proposition 4. *For region i, the autarky losses in the baseline economy* $\hat{\Xi}_i^{B \to B,A}$ *and the no-migration* *economy* $\hat{\Xi}_i^{N\to N,A}$ *are related by the following equation*

$$\left(\hat{\Xi}_i^{B\to B,A}\right)^\theta = \underbrace{\lambda_{ii} T_i^{-\theta}\left(\hat{\Xi}_i^{N\to N,A}\right)^\theta}_{\textit{domestic contribution}} + \underbrace{\sum_{j\neq i} \lambda_{ij}\left(\pi_{jj}^B\right)^{\frac{\theta}{\alpha(\eta-1)}}}_{\textit{migration opportunities contribution}}, \tag{O.52}$$

where $T_i = \left(\pi_{ii}^B/\pi_{ii}^N\right)^{1/\alpha(1-\eta)}$.

Note first that two components contribute to the baseline losses from full autarky.[41] The first component is the losses from autarky that would occur without migration, $\hat{\Xi}_i^{N\to N,A}$, whose weight is given by the fraction of workers who stay in i, λ_{ii}. The coefficient T_i corrects for the fact that migration, by itself, worsens the terms of trade for regions that receive workers. The second component measures the contribution of migration opportunities: Additional welfare losses also arise from migration destinations in which real wages drop when there is no trade. Note that if region i is a large receiver of migrants — and hence T_i is small — the losses from autarky tend to be smaller in the baseline economy than in the economy without migration, and so migration attenuates the losses from autarky (i.e., migration reduces the gains from trade).

OK.2 Gains from Trade with Multiple Activities

The forces we have uncovered in Section OK.1—which state how migration opportunities shape the full GFT with one activity—carry over to a model with multiple activities. Namely, (i) changes in real wages in other regions contribute to the GFT, and (ii) migration by itself induces changes in local real wages via changes in the terms of trade.[42]

But with multiple activities, these migration-related forces also interact with comparative advantage. For one thing, larger proportions of workers sort into region-activity combinations with high efficiency relative to the rest of the world, governing the initial $\lambda_{ij,k}$ shares. For another, comparative advantage activities tend to experience larger reductions in real wages from going to trade autarky. We now examine the quantitative importance of all these forces in our baseline model.

Panel (a) in Appendix Figure O.12 shows that the full GFT are large (26 percent on average) but vary substantially across regions. Using equation (O.52), we can compute the share of the full GFT accounted for exclusively by migration opportunities (i.e., the welfare loss from real wage losses in migration destinations).[43] The contribution of migration opportunities ranges from minimal to almost the

[41] We continue to write changes in real wages as a function of domestic trade shares π_{ii} (although the first term is not directly observable) to both keep the symmetry with equation (O.51) and emphasize that some of these components are directly observable.

[42] Dix-Carneiro and Kovak (2017) show that exposure to a trade liberalization shock has protracted, negative effects across Brazilian labor markets. Our results show that negative outcomes in a local labor market also impact welfare for workers who could potentially migrate to that labor market.

[43] That is, we set $T_i^{-\kappa}\left(\hat{\Xi}_i^{N\to N,A}\right)^\theta$ equal to 1 in Equation (O.52) and divide by the full GFT. Section OK.4 shows the equivalent of expression (O.52) in a model with many activities, which we use to make these calculations.

full losses from trade, with an average of 61 percent.

As shown in panel (b), the migration-opportunity contribution is correlated with the fraction of people (in the model) who leave when migration is available, $1 - \sum_k \lambda_{ii,k}$, as one expects from Equation (O.52). But the correlation is far from perfect, which highlights that comparative advantage and regional heterogeneity also shape the importance the migration-opportunity component. Since the West tends to receive workers from the East, the migration-opportunity contribution is smaller on average for the West than the East (49 percent and 62 percent; see Panel (c)). This share is particularly low in large urban centers in the East and expanding agricultural regions in the West.

To assess the impact of comparative advantage on the full GFT, we compute the ratio of the GFT in a model with one activity relative to a model with multiple ones.[44] As panels (d) and (e) in Appendix Figure O.12 show, the contribution of multiplicity of activities to the GFT is strongly associated with comparative advantage in agriculture relative to the rest of the world. For many regions in the West, including multiple activities almost doubles the GFT. Since comparative advantage is inherently a local characteristic of region i, it operates chiefly by strengthening the response of $i's$ wages to trade and less so by shifting migration opportunities (i.e., it tends to shift the "domestic" component in equation (O.52)).

Lastly, we study in Panel (f) how migration and comparative advantage interact to determine the GFT. On the y-axis, we measure how the contribution of migration opportunities changes, when going from a single-activity model to one with multiple activities. On the x-axis we measure again the GFT with many activities relative to the GFT with one activity. The figure shows a clear negative relation, with a slope of -0.51, which means that the larger the role of comparative advantage in the GFT, the smaller the share of migration opportunities in the GFT (relative to a single-activity model). Hence, although migration opportunities create further gains from trade in a many-activity model (relative to a single-activity one), these gains do not rise as fast as the additional gains coming from domestic markets.

OK.3 International Gains from Trade and the March

The previous section establishes the contribution of different mechanisms in the case of full GFT and full migration autarky. This section returns to our main counterfactual and studies how the March to the West interacted with international GFT.

Appendix Figure O.13 (a) maps the international GFT across regions. For Brazil as a whole, the international GFT are 5.0 percent, reflecting that it is a relatively closed economy. Within the country, nevertheless, there are regions for which international trade is crucial, and limiting it can cut down welfare by as much as 11 percent. Panel (b) presents the impact of limiting East-West migration on

Although a decomposition such as (O.52)—which exactly links welfare losses in the two scenarios—is not available in a model with multiple activities, one can still easily separate the contribution of migration opportunities to the total gains from trade, as we do in this exercise. The correction term T adds little quantitatively to our results. The average $T_i^{-\kappa}$ across meso-regions is approximately 1.01 with a standard deviation of 0.06.

[44] We calculate the one-sector losses using equation (O.51). Levchenko and Zhang (2016) perform a similar comparison underscoring the quantitative importance of comparative advantage.

the gains from international trade. The average international GFT drops by a modest 3.1 percent (0.15 percentage points out of the 5.0 percent baseline), but the differentials range from -19 to 7 percent across regions. The impact of migration on the international GFT is particularly large for the Central-West region, which hosts a large production share of Brazil's new export activities, and which also received the majority of Eastern migrants. We now proceed to disentangle the forces behind these international GFT differentials.

Echoing the results in Section OK.2, we begin by computing the contribution of East-West migration to international GFT in our baseline economy. Across regions, on average 22 percent of the international GFT are associated with East-West migration opportunities. In the counterfactual economy, these opportunities are not available to workers, which tends to lower the international GFT across the board.

To understand the regional variation in Panel (b), the interaction between migration and comparative advantage is key. Consider first what happens to real wages in each region. From Figure O.3, we know that Eastern workers sort disproportionately into agriculture when they migrate to the West and, especially, the Central-West. This means that Eastern migrants sort according to the West's international comparative advantage, which makes Western sales more reliant on international markets, rather than domestic ones. In the West, therefore, the drop in real wages from going to international trade autarky is larger in the baseline, when migration is allowed. The exact opposite happens (i) in a few regions in the northeast—which also have a comparative advantage in agriculture relative to ROW, but instead receive the Eastern agricultural workers in the no-migration counterfactual—and (ii) in the manufacturing regions in the Amazon, such as Manaus. In the rest of the East, because changes in labor supply are small, these effects are quite muted. Finally, note that an additional consequence of limiting migration is to make local real wages, relative to the ones associated with migration opportunities, have a larger weight in the expected welfare of workers born in each region.

Putting these forces together, we conclude that the international GFT in high-population regions in the East were not greatly affected by migration, which explains why aggregate GFT are insensitive to it. But the large heterogeneity we observe across other regions is driven by how migration interacts with the forces of comparative advantage.

OK.4 Proofs

This Section contains the proofs to the Propositions in Section OK.1. It also contains analytical expressions corresponding to the discussion in Sections OK.2 and OK.3. In what follows, we set $\gamma_k = 0$, $\forall k$, $\delta = 0$, $\xi = 0$ and drop time indexes.

Let W_i denote the real wage in region i, $W_i = w_i/P_i$. Inverting the domestic trade share, we obtain:

$$W_i = \pi_{ii}^{\frac{1}{\alpha(1-\eta)}} A_i^{\frac{1}{\alpha}},$$

which implies the following changes in real wages in response in changes to fundamentals:

$$\hat{W}_i = (\hat{\pi}_{ii})^{\frac{1}{\alpha(1-\eta)}} \hat{A}_i^{\frac{1}{\alpha}}. \tag{O.53}$$

Likewise, the implied changes to expected welfare are

$$\hat{\Xi}_i = \left[\sum_j \lambda_{ij} \left(\hat{W}_j \hat{s}_i \hat{\mu}_{ij}^{-1} \right)^{\theta} \right]^{1/\theta}, \tag{O.54}$$

where λ_{ij} are observed migration shares.

We introduce the following notation to indicate four scenarios: (i) B is our baseline with observed trade costs and migration costs, (ii) B, A is the scenario in which, starting from B, we take region i to full trade autarky, (iii) N is the scenario in which, starting from B, we take region i to full migration autarky, and (iv) N, A corresponds to the scenario in which, starting from N, we take region i to full trade autarky. Note, in what follows, that $\hat{A}_i = \hat{s}_i = 1$.

OK.4.1 Proof of Proposition 2

Observing that $\hat{\pi}_{ii}^{B \to B,A} = \pi_{ii}^{-1}$, i.e., the inverse of the observed trade shares, direct substitution of (O.53) in (O.54) yields

$$\hat{\Xi}_i^{B \to B,A} = \left[\sum_j \lambda_{ij} \pi_{jj}^{\frac{\theta}{\alpha(\eta-1)}} \right]^{1/\theta}, \tag{O.55}$$

which completes the proof.

OK.4.2 Proof of Proposition 3

Start by noting that we can write the welfare change from going to autarky, starting from no migration, $N \to N, A$ as

$$\begin{aligned} \hat{\Xi}_i^{N \to N,A} &= \frac{\Xi_i^{N,A}}{\Xi_i^{N}} \\ &= \frac{\Xi_i^{N,A}}{\Xi_i^{B}} \left(\frac{\Xi_i^{N}}{\Xi_i^{B}} \right)^{-1}. \end{aligned} \tag{O.56}$$

We obtain expressions for each of the terms in the last equation.

Applying the same reasoning that led to equation (O.55), we obtain

$$\left(\frac{\Xi_i^{N}}{\Xi_i^{B}} \right)^{-1} = \left(\lambda_{ii}^{\frac{1}{\theta}} \left(\frac{\pi_{ii}^{N}}{\pi_{ii}^{B}} \right)^{\frac{1}{\alpha(1-\eta)}} \right)^{-1},$$

noting that π_{ii}^N is not observed and that π_{ii}^B is simply data. Likewise, we obtain:

$$\frac{\Xi_i^{N,A}}{\Xi_i^B} = \lambda_{ii}^{\frac{1}{\theta}} \left(\frac{\pi_{ii}^{N,A}}{\pi_{ii}^B}\right)^{\frac{1}{\alpha(1-\eta)}} = \lambda_{ii}^{\frac{1}{\theta}} \left(\pi_{ii}^B\right)^{\frac{1}{\alpha(\eta-1)}}.$$

Substituting the last two expressions in equation (O.56) we obtain

$$\begin{aligned}\hat{\Xi}_i^{N\to N,A} &= \left(\frac{1}{\pi_{ii}^B}\right)^{\frac{1}{\alpha(1-\eta)}} \left(\frac{\pi_{ii}^B}{\pi_{ii}^N}\right)^{\frac{1}{\alpha(1-\eta)}} \\ &= \left(\frac{1}{\pi_{ii}^B}\right)^{\frac{1}{\alpha(1-\eta)}} T_i, \end{aligned} \tag{O.57}$$

where the last line defines $T_i = \left(\pi_{ii}^B/\pi_{ii}^N\right)^{1/\alpha(1-\sigma)}$.

To obtain the result in Proposition 3, rewrite equation (O.55)

$$\hat{\Xi}_i^{B\to B,A} = \left[\lambda_{ii}\left(\pi_{ii}^B\right)^{\frac{\theta}{\alpha(\eta-1)}} + \sum_{j\neq i}\lambda_{ij}\left(\pi_{jj}^B\right)^{\frac{\theta}{\alpha(\eta-1)}}\right]^{1/\theta}$$

and use (O.57) to substitute for π_{ii}^B

$$\hat{\Xi}_i^{B\to B,A} = \left[\lambda_{ii}T^{-\kappa}\left(\hat{\Xi}_i^{N\to N,A}\right)^{\theta} + \sum_{j\neq i}\lambda_{ij}\left(\pi_{jj}^B\right)^{\frac{\theta}{\alpha(\eta-1)}}\right]^{1/\theta}. \tag{O.58}$$

OK.4.3 Gains from Trade in a Multisector Model

The key difficulty in the multi-sector case is that changes trade shares are no longer sufficient statistics for changes in real wages induced by changes in trade costs. Nevertheless, with CES preferences across activities and an elasticity of substitution different from one, one can use changes in observed expenditure shares to proceed.

Gains from Trade. Note first that we can rewrite trade shares as a function of real wages and expenditure shares

$$\begin{aligned}\pi_{ii,k} &= \left(\frac{w_{i,k}^{\alpha_k}P_i^{1-\alpha_k}/A_{i,k}}{P_{i,k}}\right)^{1-\eta} \\ \pi_{ij,k} &= \left(\frac{W_{i,k}^{\alpha_k}}{A_{i,k}}\frac{P_i}{P_{i,g}}\frac{P_{i,g}}{P_{i,k}}\right)^{1-\eta} \\ \pi_{ii,k} &= \left(\frac{W_{i,k}^{\alpha_k}}{A_{i,k}}\left(\frac{S_{i,k}}{\bar{a}_k}\right)^{\frac{1}{\sigma_g-1}}\left(\frac{S_{i,s}}{\bar{b}_s}\right)^{\frac{1}{\sigma-1}}\right)^{1-\eta},\end{aligned}$$

where we use

$$S_{i,k} = \bar{a}_k \left(\frac{P_{j,k}}{P_{j,s}} \right)^{1-\sigma_s}$$
$$S_{i,s} = \bar{b}_s \left(\frac{P_{i,s}}{P_i} \right)^{1-\sigma}$$

We begin by computing the GFT starting from the baseline and going to full trade autarky. Noting that

$$W_{i,k} = \pi_{ii,k}^{\frac{1}{\alpha_k(1-\eta)}} A_{i,k}^{\frac{1}{\alpha_k}} \left(S_{i,k}\right)^{\frac{1}{\alpha_k(1-\sigma_s)}} \left(S_{i,s}\right)^{\frac{1}{\alpha_k(1-\sigma)}},$$

we can compute changes in real wages, $\hat{W}_{i,k}$, and substitute them into the change in expected welfare $\hat{\Xi}_i^{B\to B,A}$:

$$\hat{\Xi}_i^{B\to B,A} = \left[\sum_j \sum_k \lambda_{ij,k} \left(\pi_{ii,k}^{\frac{1}{\alpha_k(\eta-1)}} \hat{S}_{i,k}^{\frac{1}{\alpha_k(1-\sigma_s)}} \hat{S}_{i,s}^{\frac{1}{\alpha_k(1-\sigma)}} \right)^{\theta} \right]^{1/\theta}$$

GFT Comparison to the Migration Autarky Scenario. Our goal now is to compare the gains from trade in our baseline scenario to one in which there is no migration. Unfortunately, an exact decomposition such as the one in Proposition 3 is not available. However, we will show that one can cleanly separate the gains arising from migration opportunities, as before.

First, we will see show how real wage changes determine the welfare change going to autarky in a no migration scenario. As before, note that we can decompose the welfare change as

$$\hat{\Xi}_i^{N\to N,A} = \frac{\Xi_i^{N,A}}{\Xi_i^B} \left(\frac{\Xi_i^N}{\Xi_i^B} \right)^{-1}.$$

The first and second terms are given by:

$$\frac{\Xi_i^{N,A}}{\Xi_i^B} = \left(\sum_k \lambda_{ii,k} \left(\frac{s_{i,k}^{N,A}}{s_{i,k}^B} \right)^{\theta} \left(\frac{1}{\pi_{ii,k}^B} \right)^{\frac{\theta}{\alpha_k(1-\eta)}} \left(\frac{S_{i,Sk}^{N,A}}{S_{i,Sk}^B} \right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}} \left(\frac{S_{i,S}^{N,A}}{S_{i,S}^B} \right)^{\frac{\theta}{\alpha_k(1-\sigma)}} \right)^{1/\theta}$$

and

$$\left(\frac{\Xi_i^N}{\Xi_i^B} \right)^{-1} = \left(\sum_k \lambda_{ij,k} \left(\frac{s_{i,k}^N}{s_{i,k}^B} \right)^{\theta} \left(\frac{\pi_{ii,k}^N}{\pi_{ii,k}^B} \right)^{\frac{\theta}{\alpha_k(1-\eta)}} \left(\frac{S_{i,Sk}^N}{S_{i,Sk}^B} \right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}} \left(\frac{S_{i,S}^N}{S_{i,S}^B} \right)^{\frac{\theta}{\alpha_k(1-\sigma)}} \right)^{-1/\theta}.$$

Putting them together, we obtain

$$\hat{\Xi}_i^{N\to N,A} = \left[\frac{\sum_k \lambda_{ii,k} \left(\frac{1}{\pi_{ii,k}^B} \right)^{\frac{\theta}{\alpha_k(1-\eta)}} \left(\frac{S_{i,Sk}^{N,A}}{S_{i,Sk}^B} \right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}} \left(\frac{S_{i,S}^{N,A}}{S_{i,S}^B} \right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}{\sum_k \lambda_{ii,k} \left(\frac{\pi_{ii,k}^N}{\pi_{ii,k}^B} \right)^{\frac{\theta}{\alpha_k(1-\eta)}} \left(\frac{S_{i,Sk}^N}{S_{i,Sk}^B} \right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}} \left(\frac{S_{i,S}^N}{S_{i,S}^B} \right)^{\frac{\theta}{\alpha_k(1-\sigma)}}} \right]^{1/\theta}.$$

Letting

$$\rho_{i,k} = \frac{\lambda_{ii,k}\left(\frac{\pi^N_{ii,k}}{\pi^B_{ii,k}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^N_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^N_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}{\sum_l \lambda_{ii,l}\left(\frac{\pi^N_{ii,k}}{\pi^B_{ii,k}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^N_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^N_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}$$

and

$$\xi_{i,k} = \frac{\left(\frac{1}{\pi^B_{ii,k}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^{N,A}_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^{N,A}_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}{\left(\frac{\pi^N_{ii,k}}{\pi^B_{ii,k}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^N_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^N_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}$$

we can rewrite the welfare change as

$$\hat{\Xi}_i^{N\to N,A} = \left[\sum_k \rho_{i,k}\xi_{i,k}\right]^{1/\theta}.$$

Thus $\{\rho_{i,k}, \xi_{i,k}\}$ and κ fully determine $\hat{\Xi}_i^{N\to N,A}$.

Note that we can write the baseline domestic trade share as

$$\left(\pi^B_{ii,k}\right)^{\frac{\theta}{\alpha_k(\eta-1)}} = \xi_{i,k}\rho_{i,k}T^{-1}_{i,k}$$

where $T_{i,k} \equiv \lambda_{ii,k}\left(\frac{S^{N,A}_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^{N,A}_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}} \Big/ \sum_l \lambda_{ii,l}\left(\frac{\pi^N_{ii,k}}{\pi^B_{ii,k}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^N_{i,Sk}}{S^B_{i,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^N_{i,S}}{S^B_{i,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}$. Finally, we can rewrite $\hat{\Xi}_i^{B\to B,A}$ to separate migration opportunities from the components that determine $\hat{\Xi}_i^{N\to N,A}$.

Now we compute $\hat{\Xi}_i^{B\to B,A}$ so

$$\begin{aligned}
\hat{\Xi}_i^{B\to B,A} &= \left[\sum_j\sum_k \lambda_{ij,k}\left(\frac{1}{\pi^B_{jj}}\right)^{\frac{\theta}{\alpha_k(1-\eta)}}\left(\frac{S^{B,A}_{j,Sk}}{S^B_{j,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^{B,A}_{j,S}}{S^B_{j,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}\right]^{1/\theta} \\
&= \left[\sum_k \lambda_{ii,k}\xi_{i,k}\rho_{i,k}T^{-1}_{i,k}\left(\frac{S^{B,A}_{j,Sk}}{S^B_{j,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^{B,A}_{j,S}}{S^B_{j,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}\right. \\
&\left.+\underbrace{\sum_{j\neq i}\sum_k \lambda_{ij,k}\left(\frac{1}{\pi^B_{jj}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}\left(\frac{S^{B,A}_{j,Sk}}{S^B_{j,Sk}}\right)^{\frac{\theta}{\alpha_k\left(1-\sigma_A\right)}}\left(\frac{S^{B,A}_{j,S}}{S^B_{j,S}}\right)^{\frac{\theta}{\alpha_k(1-\sigma)}}}_{\text{contribution of migration opportunities}}\right]^{1/\theta}.
\end{aligned} \tag{O.59}$$

Note that this decomposition is analogous to O.58, which forms the basis of Proposition 3.

In calculating the contribution of migration opportunities to the GFT in the multi-sector model, we rely on equation O.59.

OL The Quantitative Model

This section presents the full quantitative model. We highlight here the key differences relative to the simplified, non-quantitative model we introduced in Section 5 and, after that, state all the equations that define the equilibrium.

Technology is a Cobb-Douglas aggregator of intermediate inputs and value added, with value added share α_k for activity k. Value-added is a CES aggregator of land and labor, with elasticity ψ. Thus, equation (11) is now replaced by

$$c_{jk,t} = \left(\bar{v}_k w_{jk,t}^{1-\psi} + (1-\bar{v}_k)\, r_{j,t}^{1-\psi}\right)^{\frac{\alpha_k}{1-\psi}} \left(\tilde{P}_{j,t}\right)^{1-\alpha_k} \tag{O.60}$$

and for reference we define the value added share of labor as

$$v_{jk,t} = \frac{\bar{v}_k w_{jk,t}^{1-\psi}}{\bar{v}_k w_{jk,t}^{1-\psi} + (1-\bar{v}_k)\, r_{j,t}^{1-\psi}}.$$

We introduce a local disamenity that depends on total population in the location of residence

$$u_{i,t} = \bar{u}_{i,t} L_{i,t}^{-\chi}$$

Thus, equation (6) is replaced by

$$W_{ijk,t} = \frac{w_{jk,t} s_{ik,t} u_{i,t}}{P_{j,t}}. \tag{O.61}$$

The quantitative model also allows for more general preferences. In particular, there are three nests. The upper nest combines Agriculture, Manufacturing, and Services with constant elasticity σ. The middle nest allows for different activities k within each broad sector s, which are combined with constant elasticity σ_s. The lower tier combines regionally differentiated varieties of each activity k, with constant elasticity η_k. Thus, equations (13) and (14) are now replaced by

$$p_{jk,t} = \left(\sum_i \left(c_{ik,t}\tau_{ijk,t}/A_{ik,t}\right)^{1-\eta_k}\right)^{\frac{1}{1-\eta_k}}$$

$$P_{js,t} = \left(\sum_{k\in s} \bar{a}_{k,t} \left(p_{jk,t}\right)^{1-\sigma_s}\right)^{\frac{1}{1-\sigma_s}}$$

$$\tilde{P}_{j,t} = \left(\sum_s \bar{b}_{s,t} \left(P_{js,t}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$$

and the corresponding expenditure shares are now

$$\pi_{ijk,t} = \left(\frac{c_{ik,t}\tau_{ijk,t}/A_{ik,t}}{p_{jk,t}}\right)^{1-\eta_k},$$

which replaces 18, and

$$a_{jk,t} = \bar{a}_{k,t} \left(\frac{p_{jk,t}}{P_{js,t}} \right)^{1-\sigma_s},$$
$$b_{j,t} = \bar{b}_{s,t} \left(\frac{P_{js,t}}{\tilde{P}_{j,t}} \right)^{1-\sigma}.$$

To allow for expansions in land supply, we postulate local governments that produce land using intermediate inputs (in a decreasing returns technology) with productivity shifter $g_{i,t}$. The resulting land supply is

$$H_{i,t} = g_{i,t}^{\frac{\zeta}{\zeta-1}} \left(\frac{r_{i,t}}{\zeta \tilde{P}_{i,t}} \right)^{\frac{1}{\zeta-1}},$$

where ζ , the elasticity of land supply emerges from the technology of the firm. We assume the profits from land creation are allocated to the farmers who work in region i, proportionally to their earnings.

With this changes, the full system of equations that define the equilibrium now becomes

$$E_{ik,t} = \sum_{i'} s_{i'k,t} L_{i'k,t} \quad \text{(O.62)}$$

$$X_{i,t} = \sum_k w_{ik,t} E_{ik,t} + r_{i,t} H_{i,t} + \sum_k (1-\alpha_k) Y_{ik,t} \quad \text{(O.63)}$$

$$w_{ik,t} E_{ik,t} = \upsilon_{ik,t} \alpha_k Y_{ik,t} \quad \text{(O.64)}$$

$$r_{i,t} H_{i,t} = \sum_k (1-\upsilon_{ik,t}) \alpha_k Y_{jk,t} \quad \text{(O.65)}$$

$$H_{i,t} = g_{i,t}^{\frac{\zeta}{\zeta-1}} \left(\frac{r_{i,t}}{\zeta \tilde{P}_{i,t}} \right)^{\frac{1}{\zeta-1}} \quad \text{(O.66)}$$

$$\upsilon_{ik,t} = \frac{\bar{\upsilon}_k w_{ik,t}^{1-\psi}}{\bar{\upsilon}_k w_{ik,t}^{1-\psi} + (1-\bar{\upsilon}_k) r_{i,t}^{1-\psi}} \quad \text{(O.67)}$$

$$c_{ik,t} = \left(\bar{\upsilon}_k w_{ik,t}^{1-\psi} + (1-\bar{\upsilon}_k) r_{i,t}^{1-\psi} \right)^{\frac{\alpha_k}{1-\psi}} \left(\tilde{P}_{i,t} \right)^{1-\alpha_k} \quad \text{(O.68)}$$

$$Y_{ik,t} = \sum_j \pi_{ijk,t} a_{jk,t} b_{j,t} X_{j,t} \quad \text{(O.69)}$$

$$p_{jk,t} = \left(\sum_i (c_{ik,t} \tau_{ijk,t} / A_{ik,t})^{1-\eta_k} \right)^{\frac{1}{1-\eta_k}} \quad \text{(O.70)}$$

$$P_{js,t} = \left(\sum_{k \in s} \bar{a}_{k,t} (p_{jk,t})^{1-\sigma_s} \right)^{\frac{1}{1-\sigma_s}} \quad \text{(O.71)}$$

$$\tilde{P}_{j,t} = \left(\sum_s \bar{b}_{s,t} (P_{js,t})^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad \text{(O.72)}$$

$$\pi_{ijk,t} = \left(\frac{c_{ik,t} \tau_{ijk,t} / A_{ik,t}}{p_{jk,t}} \right)^{1-\eta_k} \quad \text{(O.73)}$$

$$a_{jk,t} = \bar{a}_{k,t} \left(\frac{p_{jk,t}}{P_{js,t}} \right)^{1-\sigma_s} \quad \text{(O.74)}$$

$$b_{j,t} = \bar{b}_{s,t} \left(\frac{P_{js,t}}{\tilde{P}_{j,t}} \right)^{1-\sigma} \quad \text{(O.75)}$$

$$L_{ik,t} = \sum_j \lambda_{jik,t} L_{j,t-1} \quad \text{(O.76)}$$

$$s_{ik,t} = \bar{s}_k L_{ik,t-1}^{\beta} \quad \text{(O.77)}$$

$$\lambda_{ijk,t} = \left(\frac{W_{ijk,t}/\mu_{ijk,t} \cdot v_{j,t+1}^{\delta}}{v_{i,t}} \right)^{\theta} \quad \text{(O.78)}$$

$$W_{ijk,t} = \frac{\tilde{w}_{jk,t} s_{ik,t} u_{j,t}}{\tilde{P}_{j,t}} \quad \text{(O.79)}$$

$$v_{i,t} = \exp(\Upsilon_{i,t}) \quad \text{(O.80)}$$

$$\Upsilon_{i,t} = \frac{1}{\theta} \ln \sum_{j,k} \left(W_{ijk,t}/\mu_{ijk,t} \cdot v_{j,t+1}^{\delta} \right)^{\theta} \quad \text{(O.81)}$$

$$A_{ik,t} = \bar{A}_{ik,t} L_{ik,t}^{\xi} \quad \text{(O.82)}$$

$$u_{i.t} = \bar{u}_{i,t} L_{i,t}^{-\chi} \quad \text{(O.83)}$$

$$\tilde{w}_{i,t} = w_{i,t} \left(\frac{\zeta \upsilon_{ik,t}}{\zeta - 1 + \upsilon_{ik,t}} \right) \quad \text{(O.84)}$$

OL.1 Steady state

Substitute equation (O.81) by

$$\exp \Upsilon_i = \left[\sum_{j,k} \left(\frac{w_{jk} s_{ik} u_j}{\tilde{P}_j \mu_{ijk}} \right)^{\theta} (\exp \Upsilon_j)^{\delta\theta} \right]^{\frac{1}{\theta}},$$

equation (O.76) by

$$\ell_{ik} = \sum_{i'} \lambda_{i'ik} L_{i'},$$

equation (O.77) by

$$s_{ik} = \bar{s}_k L_{i,k}^{\beta},$$

and equation (O.78) by

$$\lambda_{ijk} = \frac{\left(\left(\left(w_{jk} s_{ik} u_j / \left[\mu_{ijk} \tilde{P}_j \right] \right)^{\theta} (exp \Upsilon_j)^{\delta\theta} \right) \right)}{exp\left(\Upsilon_i\right)^{\theta}}.$$

The remaining equations are as in the dynamic equilibrium.

OM Parametrization and Inversion Algorithm

OM.1 Hub-and-Spoke

We parameterize migration costs using a hub-spoke structure (see Ramondo et al. (2016)). Specifically, for all regions within a state, workers have to travel through a state hub to reach any other region. With this structure, we can aggregate migration flows at the meso-regional level to the state level and still use equation 28 at the aggregate level of states in a theoretically consistent manner. To see this, define $\mu_{ijk} = \mu_{ss}\mu_i\mu_j\mu_{ss'k}$ and write migration flows as:

$$L_{ijk,t} = \left(\frac{w_{jk,t} s_{ik,t} u_{j,t} / \left(\tilde{P}_{j,t} \mu_{ss,t} \mu_{i,t} \mu_{j,t} \mu_{ss'k,t} \right) \cdot \upsilon_{j,t+1}^{\delta}}{\upsilon_{i,t+1}} \right)^{\theta} L_{i,t-1}.$$

Let $L_{ss'k,t} \equiv \sum_{i \in s} \sum_{j \in s'} L_{ijk,t}$ be the aggregate flow of workers from state s to state-activity $s'k$. Summing the expression above over origins in state s and destinations in state s', activity k gives

$$L_{ss'k,t} = \sum_{i \in s} \sum_{j \in s'} \left(\frac{w_{jk,t} s_{ik,t} u_{j,t} / \left(\tilde{P}_j \mu_{ss,t} \mu_{i,t} \mu_{j,t} \mu_{ss'k,t} \right) \cdot \upsilon_{j,t+1}^{\delta}}{\upsilon_{i,t+1}} \right)^{\theta} L_{i,t-1}.$$

Straightforward manipulations give

$$\ln L_{ss'k,t} = \alpha_{sk,t} + \alpha_{s'k,t} - \theta \ln \mu_{ss',t} - \theta \ln \mu_{ss'k,t},$$

where $\alpha_{sk,t} \equiv \sum_{i\in s} \left(s_{ik,t}/\mu_{i,t}v_{i,t+1}\right)^{\theta}$ and $\alpha_{s'k,t} \equiv \sum_{j\in s'} \left(v_{j,t+1}^{\delta} w_{jk,t} u_{j,t}/\mu_{j,t}\right)^{\theta}$.

OM.2 Parametrization

Based on the parametrization of trade and migration costs given by equations (26) and (27), to simulate the model we need to calibrate

- Preference and technology parameters that are constant over time $\{\eta_k, \sigma_s, \sigma, \alpha_k, \psi, \zeta\}$
- Migration choice and knowledge parameters $\{\theta, \beta, \bar{s}_k\}$
- Productivities and preference shifters $\{A_{i,kt}, g_{i,t}, \bar{v}_{k,t}, \bar{a}_{c,kt}, \bar{b}_{c,kt}\}$
- Trade costs $\{\delta_t^0, \delta_t^1, \delta_{kt}\}$
- Migration cost $\{\mu_t^0, \mu^1, \mu_{ss',t}, \mu_{ss',kt}\}$

We use data on

- Gross output and expenditure per region $Y_{ik,t}$ and $X_{ik,t}$[45]
- Trade flows between countries by sector $X_{ck,t}$, where c indexes a country $c \in \{B, ROW\}$
- Trade flows between states of Brazil $X_{ss',t}$
- Migration flows between states $L_{ss'k',t}$
- Share of workers staying at their meso-region $L_{i,t}/\sum_{j'} L_{ij',t}$
- National share of payments to labor v_k
- Total land use per region $H_{i,t}$

OM.3 Inversion Algorithm

Step 1 - Calibrate trade costs, preference shifters and prices

The goal in this step is to calibrate $(\delta_t^0, \delta_t^1, \delta_{kt}, \bar{a}_{ck,t}, \bar{b}_{ck,t})$ and recover model implied prices $(c_{j,kt}, \tilde{P}_{j,t})$ such that the model matches the gross output per region and activity and several statistics on trade flows for each period t. In what follows, define K_A as the set of agricultural activities. The algorithm proceeds, for each period $t = 1950, 1980$ and 2010, as follows:

1. Guess $\left(\delta_t^0\right)^g, \left(\delta_t^1\right)^g, \left(\delta_{kt}\right)^g, \left(\bar{a}_{ck,t}\right)^g \left(\bar{b}_{ck,t}\right)^g$ and $\left(c_{j,kt}\right)^g$
2. Construct trade costs $\widehat{\tau}_{ij,kt}$ using equation (26)

[45] We construct expenditure per region by creating a transfer from deficits to regions.

3. Using equations (O.69)-(O.75), generate predicted trade shares ($\widehat{\pi_{ij,k}}$) and predicted gross output ($\widehat{Y}_{jk}$)

4. Match trade flows and relative size of home and foreign countries within a given sector

   (a) If $|\widehat{\pi_{HF,k}} - \pi_{HF,k}^{data}| > \epsilon$, update $(\delta_k)^{g+1}$.

   (b) If $|\widehat{\pi_{FH,k}} - \pi_{FH,k}^{data}| > \epsilon$, update $(c_{F,k})^{g+1}$.

   (c) If any update is made, return to step 1. Otherwise, continue to next step.

5. Match service sector size for each region

   (a) If $|\widehat{Y}_{jS} - Y_{jS}^{data}| > \epsilon$, update $(c_{j,S})^{g+1}$

   (b) If any update is made, return to step 1. Otherwise, continue to next step.

6. Match relative size of sectors and gross output of regions within home

   (a) If $|\widehat{Y}_{ik} - Y_{ik}^{data}| > \epsilon$, update $(c_{i,k})^{g+1}$ for $i \neq F$. Ensure $\sum_{i \neq F} (c_{i,k})^{g+1} = 1$.

   (b) If $|\sum_{i \neq F} \widehat{Y}_{ik} - \sum_{i \neq F} Y_{ik}^{data}| > \epsilon$ for $k \in K_A$, update $(a_{i,k})^{g+1}$, for $i \neq F$. Ensure $\sum_k (a_{i,k})^{g+1} = 1$.

   (c) If $|\sum_{i \neq F} \widehat{Y}_{ig} - \sum_{i \neq F} Y_{ig}^{data}| > \epsilon$ for $g \in K_A$, update $(b_{i,k})^{g+1}$, for $i \neq F$. Ensure $\sum_k (b_{i,k})^{g+1} = 1$.

   (d) If $|\widehat{Y}_{Fk} - Y_{Fk}^{data}| > \epsilon$, update $(a_{i,k})^{g+1}$, for $i = F$. Ensure $\sum_k (a_{i,k})^{g+1} = 1$.

   (e) If $|\widehat{Y}_{Fg} - Y_{Fg}^{data}| > \epsilon$ for each $g \in K_A$, update $(b_{i,k})^{g+1}$, for $i = F$. Ensure $\sum_k (b_{i,k})^{g+1} = 1$.

   (f) If any update is made, return to step 1. Otherwise, continue to next step.

7. Using equations (O.69)-(O.75), generate predicted trade flows $\widehat{X}_{ss',t}$ between states in Brazil

   (a) Estimate model-implied trade elasticity $\widehat{\beta}_t$ based on

   $$\widehat{X}_{ss',t} = \beta_{s,t} + \beta_{s',t} + \widehat{\beta}_t \text{dist}_{ss',t},$$

   and data implied trade elasticity β_t using $\widehat{X}_{ss',t}^{data}$.

   (b) Compute share of trade within country

   $$\widehat{\pi}_{ss,t} = \frac{\sum_{s'} X_{ss,t}}{\sum_s \sum_{s'} X_{ss',t}}.$$

8. Match domestic trade elasticities and domestic trade shares. Using π_{ss}^{data}, which is the share of trade of a state with itself in the data, and $\widehat{\pi}_{ss,t}$, the corresponding value in the model:

(a) If $|\widehat{\beta}_t - \beta^{data}| > \epsilon$, update $\left(\delta_t^1\right)^{g+1}$

(b) If $|\widehat{\pi}_{ss,t} - \pi_{ss,t}^{data}| > \epsilon$, update $\left(\delta_t^0\right)^{g+1}$

(c) If any update is made, return to step 1. Otherwise, move to the next step.

9. Save calibrated $\widehat{\tau}_{ij,kt}$, $\widehat{\bar{a}_{ck,t}}$, and $\widehat{\bar{b}_{ck,t}}$ and model-implied prices $\widehat{\tilde{P}_j}$ and $\widehat{c_{j,kt}}$ to be used in the next step.

Step 2 - Calibrate migration costs and productivities

The goal in this step is to calibrate migration costs and amenities $(\mu_t^0, \mu^1, \mu_{ss',t}, \mu_{ss',kt}, \bar{u}_{i,t})$ and productivity parameters $\left(\bar{A}_{i,kt}, g_{i,t}, \bar{v}_{k,t}\right)$ such that the model solution to the dynamic problem is consistent with model-implied prices $\left(\widehat{\tilde{P}_j}, \widehat{c_{j,kt}}\right)$ and matches migration flows for each period $t = 1950, 1980, 2010$. The algorithm proceeds as follows:

1. Estimate $\widehat{\mu}_{ss',t}$ using

$$L_{ss,kt} = \beta_{s,kt} + \beta_{s',t} + \mu_{ss',t} + \epsilon_{ss',kt}.$$

2. Estimate $\widehat{\mu^1}$ based on

$$\widehat{\mu}_{ss',t} = \beta + \mu^1 \text{dist}_{ss',t} + \epsilon_{ss',t}$$

using the highways constructed to connect Brasilia as an IV (as in Morten and Oliveira, 2016).

3. Guess $\left(\mu_t^0\right)^g, \left(\mu_{ss',kt}\right)^g, \left(\bar{A}_{i,kt}\right)^g, \left(g_{i,t}\right)^g, \left(\bar{v}_{k,t}\right)^g$.

4. Construct migration costs based on equation (27).

5. Recover model-implied wages, land rents and share of labor consistent with data.

(a) Make guess of $(\bar{v}_{k,t})^g$.[46]

(b) Find model-implied $\widehat{w}_{ik,t}$ and $\widehat{r}_{i,t}$ using equations (O.62), (O.64), (O.65), (O.66), (O.67), (O.68), (O.76), (O.77), (O.78), and (O.81), $(\bar{v}_{k,t})^g$, $\left(\tilde{P}_j\right)^g$, $Y_{ik,t}^{data}$ and $H_{i,t}^{data}$.

(c) Generate model-implied $\widehat{v}_{k,t} = \left(\sum_{i \neq F} \widehat{v}_{ik,t} \alpha_{ik} \widehat{Y}_{ikt}\right) / \left(\sum_{i \neq F} \widehat{Y}_{ikt}\right)$.

(d) Compare $\widehat{v}_{k,t}$ with $v_{k,t}^{data}$ for 2010. If $|\widehat{v}_{k,t} - v_{k,t}^{data}| > \epsilon$, update $(\bar{v}_{k,t})^{g+1}$.

(e) If any update is made, return to step 5(a). Otherwise, move to next step.

6. Recover productivities.

(a) Using equation (O.68), guess $(\bar{v}_k)^g$ and model-implied prices $\widehat{c_{j,kt}}$, , $\widehat{w}_{ik}$, $\widehat{r}_i$, and $\widehat{\tilde{P}_j}$, recover $\widehat{\bar{A}}_{ik,t}$.

(b) Using equation (O.66) and model-implied prices $\widehat{r}_i$ and $\widehat{\tilde{P}_j}$, recover $\widehat{g}_{i,t}$.

[46] We only have data for $\widehat{v}_{k,t}$ in 2010, which we apply for 1950 and 1980.

7. Solve for the SS and the dynamic problem using $\widehat{A}_{ik,t}$, $\widehat{g}_{i,t}$, $\widehat{\delta}^0_t$, $\widehat{\delta}^1_t$, $\widehat{\delta_{kt}}$, $\widehat{\bar{a}_{ck,t}}$, $\widehat{\bar{b}_{ck,t}}$, $\widehat{\mu}_{ss',t}$, $\widehat{\mu^1}$, $\left(\mu^0_t\right)^g$ and $(\mu_{ss',kt})^g$ and equation (27).

8. Construct model-implied labor migration.

   (a) State-state-activity flows $\widehat{\lambda}_{ss',kt}$.

   (b) Own migration shares $\widehat{\lambda}_{ii,t} = \sum_k \lambda_{ii,kt} / \sum_j \sum_k \lambda_{ij,kt}$.

9. Match migration flows between states and out from origin region. In addition, match population living in every region.

   (a) If $|\widehat{\lambda}_{ss',kt} - \lambda^{data}_{ss',kt}| > \epsilon$, update $(\mu_{ss',kt})^{g+1}$. Ensure $\sum \mu_{ss',kt} = 1$ for $s = s'$ and $k = services$.

   (b) If $|\widehat{\lambda}_{ii,t} - \lambda^{data}_{ii,t}| > \epsilon$, update $\left(\mu^0_t\right)^{g+1}$.

   (c) If $|\widehat{\lambda}_{i,t} - \lambda^{data}_{i,t}| > \epsilon$, update $(\bar{u}_{i,t})^{g+1}$.

   (d) If any update is made, return to step 3. Otherwise, conclude the algorithm.

ON Online Appendix Figures and Tables

Table O.1: Aggregate Summary Statistics

	1950	1960	1970	1980	1990	2000	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a. Migration							
- East to East	0.851	0.842	0.825	0.759	0.735	0.719	0.687
- East to West	0.093	0.105	0.137	0.191	0.205	0.205	0.218
- East to West + West to East	0.115	0.128	0.152	0.212	0.229	0.235	0.255
- West to East	0.022	0.023	0.015	0.021	0.024	0.030	0.038
- West to West	0.033	0.030	0.024	0.029	0.036	0.046	0.058
b. Economic Aggregates							
- Brazil's GDP	1.000	1.000	1.000	1.000	1.000	1.000	1.000
- Exports	0.071	0.065	0.067	0.080	0.095	0.107	0.117
- Imports	0.071	0.068	0.075	0.093	0.069	0.122	0.125
- World's GDP	99.415	94.235	68.613	43.943	50.187	51.973	31.772

Notes: Panel (a) shows the share of different categories of migrants. We define migration based on a workers state of living and state of birth. It shows that East to West migration accounted for 9 percent of all interstate migrants in 1950. Panel (b) presents values normalized by Brazil's GDP in a given year.

Table O.2: Summary Statistics by Activity (in percentages)

				Percentage within Agriculture										
	services	mfg	agriculture	rest of agri	banana	cocoa	coffee	corn	cotton	livestock	rice	soy	sugarcane	tobacco
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
a. Value Added														
-1950	53.6	21.6	24.8	27.8	1.3	1.3	16.0	7.3	6.7	27.3	6.8	0.2	4.1	1.0
-1980	56.5	33.8	9.7	24.2	1.6	2.3	7.1	9.9	2.8	27.3	7.1	9.2	7.2	1.2
-2010	74.0	21.0	5.0	22.8	2.4	0.6	6.5	7.3	1.7	25.9	3.8	15.7	10.5	2.7
b. Workers														
-1950	65.3	7.5	27.2	35.4	0.7	0.8	11.2	10.1	3.7	27.0	5.6	0.2	4.2	0.9
-1980	68.7	14.0	17.3	20.6	0.7	1.9	8.0	14.0	6.1	27.0	11.3	3.9	4.7	1.8
-2010	84.5	7.5	8.0	27.7	1.3	1.2	12.4	7.5	0.1	27.7	2.9	3.1	13.0	3.1
c. Exports from Brazil to the ROW														
-1950	0.0	20.2	79.8	13.7	0.5	9.0	59.2	1.3	6.2	2.2	0.4	0.0	5.7	1.8
-1980	0.0	55.6	44.4	28.2	0.2	8.6	23.9	0.5	0.5	3.5	0.2	21.4	9.0	4.1
-2010	0.0	68.0	32.0	28.4	0.1	0.4	8.0	3.9	1.8	6.9	0.6	29.4	16.1	4.4
d. Imports of Brazil from the ROW														
-1950	0.0	82.9	17.1	99.4	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0
-1980	0.0	90.1	9.9	73.7	0.0	0.0	1.9	8.0	0.1	4.2	4.5	7.6	0.0	0.1
-2010	0.0	95.4	4.6	87.9	0.0	2.2	0.3	1.6	1.2	2.1	3.3	0.8	0.0	0.6

Notes: This table shows the distribution of economic activity that we match in our calibrated model. Data for workers per activity comes from Brazilian censuses. Data for value added per sector comes from United Nations since the 1970s and extrapolated back to 1950 based on data from IPEA-DATA. Data on trade comes from COMTRADE and IPEA-DATA. Labor employment by sector constructed based on the CENSUS.

Table O.3: Evolution of Revealed Comparative Advantage

	Brazil			East	West
	1950	1980	2010	2010	2010
	(1)	(2)	(3)	(4)	(5)
banana	1.92	1.29	0.36	0.43	0.00
cacao	17.08	14.98	1.26	1.51	0.01
coffee	35.42	17.58	15.34	18.33	0.29
corn	1.53	0.34	6.70	2.23	29.22
cotton	2.96	0.50	5.72	3.93	14.76
beef	2.43	3.17	9.53	5.95	27.59
rice	0.59	0.27	1.35	1.59	0.15
soy	0.00	15.33	22.56	15.14	59.96
sugarcane	3.35	6.08	29.03	33.21	7.97
tobacco	1.48	4.07	6.37	7.63	0.00
rest of agriculture	0.71	1.66	1.69	1.77	1.30
agriculture	3.37	3.71	4.47	4.06	6.55
manufacturing	0.27	0.63	0.73	0.76	0.57

Notes: This table shows the evolution of revealed comparative advantage as measured by the Balassa index:

$$RCA_{i,k} = \frac{X_{i,k}/\sum_k X_{i,k}}{\sum_i X_{i,k}/\sum_i \sum_k X_{i,k}},$$

where $X_{i,k}$ are exports from country i in activity k.

Table O.4: Migrants' Composition and Average Labor Productivity - First stage - w/ bilateral controls

	2010	2000	1990	1980	1970	1960	1950
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$I_{o-r(d),1950} \times \frac{I_{-r(o)d,1950}}{I_{-r(o),1950}}$	-0.048	-0.012	-0.047	-0.039	0.309***	0.368***	0.177***
	(0.100)	(0.123)	(0.105)	(0.092)	(0.075)	(0.047)	(0.023)
$I_{o-r(d),1960} \times \frac{I_{-r(o)d,1960}}{I_{-r(o),1960}}$	-0.162***	-0.196***	-0.276***	-0.307***	-0.288***	-0.086***	
	(0.057)	(0.070)	(0.060)	(0.053)	(0.044)	(0.027)	
$I_{o-r(d),1970} \times \frac{I_{-r(o)d,1970}}{I_{-r(o),1970}}$	0.084***	0.081**	0.146***	0.194***	0.188***		
	0.031	0.038	0.032	0.028	0.024		
$I_{o-r(d),1980} \times \frac{I_{-r(o)d,1980}}{I_{-r(o),1980}}$	0.055**	0.091***	0.120***	0.128***			
	(0.022)	(0.028)	(0.023)	(0.021)			
$I_{o-r(d),1990} \times \frac{I_{-r(o)d,1990}}{I_{-r(o),1990}}$	0.199***	0.349***	0.121***				
	(0.023)	(0.027)	(0.018)				
$I_{o-r(d),2000} \times \frac{I_{-r(o)d,2000}}{I_{-r(o),2000}}$	-0.252***	-0.400***					
	0.046	0.054					
$I_{o-r(d),2010} \times \frac{I_{-r(o)d,2010}}{I_{-r(o),2010}}$	0.019						
	(0.073)						
R^2	0.012	0.022	0.014	0.012	0.014	0.009	0.007
Obs	8450	8450	8450	8450	8450	8450	8450

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level.

Table O.5: Migrants' Composition and Average Labor Productivity - First stage - w/o bilateral controls

	2010	2000	1990	1980	1970	1960	1950
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$I_{o-r(d),1950} \times \frac{I_{-r(o)d,1950}}{I_{-r(o),1950}}$	-0.166*	-0.135	-0.192*	-0.178*	0.210***	0.328***	0.176***
	(0.101)	(0.124)	(0.105)	(0.092)	(0.075)	(0.048)	(0.023)
$I_{o-r(d),1960} \times \frac{I_{-r(o)d,1960}}{I_{-r(o),1960}}$	-0.125**	-0.158**	-0.233***	-0.262***	-0.252***	-0.061**	
	(0.057)	(0.070)	(0.060)	(0.053)	(0.044)	(0.027)	
$I_{o-r(d),1970} \times \frac{I_{-r(o)d,1970}}{I_{-r(o),1970}}$	0.110***	0.109***	0.178***	0.218***	0.208***		
	0.031	0.039	0.033	0.028	0.024		
$I_{o-r(d),1980} \times \frac{I_{-r(o)d,1980}}{I_{-r(o),1980}}$	0.062***	0.098***	0.132***	0.139***			
	(0.022)	(0.028)	(0.023)	(0.021)			
$I_{o-r(d),1990} \times \frac{I_{-r(o)d,1990}}{I_{-r(o),1990}}$	0.182***	0.330***	0.116***				
	(0.022)	(0.027)	(0.018)				
$I_{o-r(d),2000} \times \frac{I_{-r(o)d,2000}}{I_{-r(o),2000}}$	-0.211***	-0.352***					
	0.046	0.054					
$I_{o-r(d),2010} \times \frac{I_{-r(o)d,2010}}{I_{-r(o),2010}}$	0.035						
	(0.072)						
R^2	0.011	0.021	0.015	0.013	0.014	0.008	0.007
Obs	8450	8450	8450	8450	8450	8450	8450
Controls							
Dist	Y	Y	Y	Y	Y	Y	Y
Ag. sim.	Y	Y	Y	Y	Y	Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level.

Table O.6: Migrants' Composition and Average Labor Productivity - Alternative specifications

	OLS (1)	OLS (2)	OLS (3)	IV (4)	IV (5)	IV (6)	IV (7)	IV (8)	IV (9)
Panel a. Years after 2000									
Composition	0.205***	0.201***	0.231***	0.220	0.383**	0.326***	0.218	0.402***	0.349***
	(0.068)	(0.071)	(0.065)	(0.135)	(0.150)	(0.125)	(0.137)	(0.155)	(0.134)
R2 or K-P F	0.885	0.923	0.932	20.940	12.514	19.170	18.999	11.901	17.437
Obs	355	355	355	355	355	355	355	355	355
Panel b. Years after 1990									
Composition	0.243***	0.256***	0.336***	0.313**	0.513***	0.384***	0.323**	0.530***	0.410***
	(0.067)	(0.073)	(0.067)	(0.129)	(0.168)	(0.122)	(0.129)	(0.169)	(0.127)
R2 or K-P F	0.883	0.915	0.929	26.814	16.852	25.905	25.904	17.920	26.775
Obs	557	557	557	557	557	557	557	557	557
Panel c. Years after 1980									
Composition	0.205***	0.189***	0.275***	0.226**	0.344***	0.276***	0.231**	0.355***	0.294***
	(0.059)	(0.055)	(0.056)	(0.109)	(0.121)	(0.090)	(0.109)	(0.121)	(0.095)
R2 or K-P F	0.885	0.917	0.929	29.106	22.115	31.020	27.969	23.760	31.713
Obs	751	751	751	751	751	751	751	751	751
Panel d. Years after 1970									
Composition	0.142***	0.125***	0.199***	0.134	0.199**	0.137*	0.136	0.208**	0.150*
	0.052	0.042	0.050	0.101	0.096	0.077	0.100	0.097	0.080
R2 or K-P F	0.888	0.916	0.926	30.338	23.148	30.614	28.556	24.381	30.721
Obs	927	927	927	927	927	927	927	927	927
Control in OLS and IV estimates									
-Crop suitability	Y	Y	Y	Y	Y	Y	Y	Y	Y
-State-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
-Crop-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
-Broad Region-Crop FE	-	Y	Y	-	Y	Y	-	Y	Y
-Lagged prod.	-	-	Y	-	-	Y	-	-	Y
-Labor emp.	-	-	Y	-	-	Y	-	-	Y
-Share of mig.	-	-	Y	-	-	Y	-	-	Y
Controls in the construction of the IV									
-Dist	-	-	-	Y	Y	Y	-	-	-
-Ag. sim.	-	-	-	Y	Y	Y	-	-	-

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level.

Table O.7: Robustness - Dropping Crops

	Activities Dropped										
	banana (1)	cocoa (2)	coffee (3)	corn (4)	cotton (5)	cattle (6)	rice (7)	soy (8)	sugarcane (9)	tobacco (10)	capital-intensive (11)
Panel a. OLS											
Comp.	0.26***	0.30***	0.29***	0.26***	0.24***	0.27***	0.25***	0.27***	0.33***	0.31***	0.25***
	(0.06)	(0.08)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.07)	(0.06)
R2 / KP-F	0.93	0.92	0.94	0.94	0.93	0.93	0.94	0.93	0.91	0.93	0.91
Obs	648	719	669	648	729	674	648	694	648	682	569
Panel b. IV											
Comp.	0.27***	0.55	0.13	0.30***	0.28***	0.28***	0.31***	0.28***	0.33***	0.20**	0.25***
	(0.10)	(0.42)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)	(0.06)
R2 / KP-F	0.28	0.24	0.31	0.25	0.22	0.26	0.30	0.26	0.29	0.30	0.91
Obs	648	719	669	648	729	674	648	694	648	682	569

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level. Capital intensive include cotton, soy, and sugarcane.

Table O.8: Migrants's Composition and Average Labor Productivity - Alternative RHS

	DV: Log of Output per Worker					
	Quantity		Population		Revenues	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Composition	0.218***	0.242***	0.154**	0.314***	0.241***	0.242***
	(0.052)	(0.080)	(0.063)	(0.107)	(0.046)	(0.079)
R2 or K-P	0.927	52.408	0.923	38.562	0.928	49.491
Obs	751	751	751	751	751	751

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level. This table present results from estimates of equation (2) replacing $S_{i,k}$ by different measures of economic activity in the origin. Every regression control for lagged productivity, log of labor employment, log of share of migrants, and broad region/crop fixed effects. Regressions are weighted by states' initial population size.

Table O.9: Migrants's Composition Average Labor Productivity - Alternative LHS

	Dependent Variable					
	Rev per Worker		Price		Exports	
	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Composition	0.151***	0.216**	-0.102***	-0.082*	0.096	-0.439
	(0.048)	(0.094)	(0.038)	(0.046)	(0.170)	(0.369)
R2 or K-P	0.811	26.569	0.934	30.403	0.874	12.067
Obs	823	823	755	755	279	279

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level. This table present results from estimates of equation (2) replacing the dependent variable by alternative outcomes. Every regression control for lagged productivity, log of labor employment, log of share of migrants, and broad region/crop fixed effects. Regressions are weighted by states' initial population size.

Table O.10: The Impact of Migrants' Composition on Average Labor Productivity - Evaluating the role of Spatially Correlated Productivity

	DV: Log of Output per Worker			
			Baseline spec.	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
a. Baseline composition term				
Composition	0.247***	0.365***	0.275***	0.276***
	(0.048)	(0.090)	(0.056)	(0.090)
R2 or K-P	0.933	41.956	0.929	31.020
Obs	751	751	751	751
b. Dropping nearest neighbor from composition term				
Composition	0.251***	0.287***	0.274***	0.206**
	(0.051)	(0.083)	(0.058)	(0.085)
R2 or K-P	0.933	86.296	0.929	55.674
Obs	751	751	732	732
Controls in First and Second Stage				
-State/Year FE	Y	Y	Y	Y
-Crop/Year FE	Y	Y	Y	Y
-Broad Region/Crop FE	Y	Y	Y	Y
-Crop suit.	Y	Y	Y	Y
-Neighbor suit.	Y	Y	Y	Y
-Crop suit. trend	Y	Y	-	-
-Neighbor suit. trend	Y	Y	-	-
-Lagged productivity	Y	Y	Y	Y
-Total farmers	Y	Y	Y	Y
-Share of migrants	Y	Y	Y	Y
Controls in the Construction of Predicted Migration				
-No $-r(s)$ terms in Push-Pull	-	Y	-	Y
-Dist.	-	Y	-	Y
-Agsim.	-	Y	-	Y
-Dist. $\times$ crop	-	Y	-	-
-Agsim. $\times$ crop	-	Y	-	-
-Dist. in Suit.	-	Y	-	-

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state-crop level. Relative to the main Table 2, this table adds controls for average suitability of neighbors $\left(\frac{\sum_{s' \neq s} dist_{s's} S_{s'k}}{\sum_{s' \neq s} dist_{s'}}\right)$, we exclude the nearest state in Panel b, and we add controls for the difference in suitability S_{sk} between origin and destination states in the construction of predicted migration flows.

Table O.11: Migrants' Composition and Coffeee Specialization - Robusta vs Arabica

	DV: Share of Output in Arabica		
	OLS	OLS	OLS
	(1)	(2)	(3)
Composition	0.842***	0.848***	0.791***
	(0.269)	(0.261)	(0.150)
R2	0.352	0.666	0.685
Obs	93	93	93
Controls			
- Total employment in coffee	-	Y	Y
- Coffee suitability	-	Y	Y
- Region FE	-	-	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clusters at the state level. Every regression controls for the log of total number of coffee farmers in a region and broad region fixed effects. Regressions weighted by total number of farmers.

Table O.12: Migrants' Composition and Average Labor Productivity - Meso-Region level

	DV: Log of Output per Worker			
	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
Composition	0.853***	0.333**	1.263***	0.794*
	(0.143)	(0.153)	(0.344)	(0.407)
R2 or K-P	0.551	0.638	70.325	55.682
Obs	956	956	956	956
Controls				
-State/Year FE	Y	Y	Y	Y
-Crop/Year FE	Y	Y	Y	Y
-Lagged productivity	-	Y	-	Y
-Labor employment	-	Y	-	Y
-Share of migrants	-	Y	-	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. This table presents results for the impact of migrants' composition using meso-region level data. See Appendix OC for details.

Table O.13: Earnings and Farmers' Choice Regression - RAIS

	OLS	OLS	PPML	PPML	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a. Income (logs)							
Farmers in origin	0.010***	0.012***	0.010***	0.013***	0.025***	0.033***	0.027***
	(0.001)	(0.002)	(0.002)	(0.003)	(0.004)	(0.005)	(0.003)
R^2 or K-P	0.658	0.687	0.207	0.222	331.132	274.305	193.908
Overid. p							0.085
Obs	94472	70220	94472	70220	94472	94472	94472
b. Farmers in destination (logs)							
Farmers in origin	0.128***	0.159***	0.208***	0.232***	0.228***	0.254***	0.236***
	(0.005)	(0.007)	(0.008)	(0.011)	(0.009)	(0.016)	(0.010)
R^2 or K-P	0.815	0.842	0.863	0.873	331.132	274.305	193.908
Overid. p							0.108
Obs	94472	70220	94472	70220	94472	94472	94472
Dest/Act/Year FE	Y	Y	Y	Y	Y	Y	Y
Dest/Orig/Year FE	Y	Y	Y	Y	Y	Y	Y
Above Q1	-	Y	-	Y	-	-	-
Crop Suitability	-	-	-	-	Y	Y	Y
Harris's IV	-	-	-	-	Y	-	Y
Mig. Comp. IV	-	-	-	-	-	Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. This table replicates Table C.3 in the main body of the paper but using the administrative data on formal employment RAIS instead of the CENSUS.

Table O.14: Earnings and Farmers' Choice Regression - Alternative Specifications for the IV

	2SLS	2SLS	2SLS	2SLS
		Smaller	Include	Exclude
	Bline	Cost	Dist. < 1h	Dist. < 3h
	(1)	(2)	(3)	(4)
a. Income (logs)				
Farmers in origin	0.048***	0.050***	0.052***	0.045**
	(0.019)	(0.014)	(0.015)	(0.022)
R^2 or K-P	94.211	19.762	23.552	132.482
Obs	6778	6778	6778	6778
b. Farmers in destination (logs)				
Farmers in origin	0.075**	0.131***	0.134***	0.101***
	(0.032)	(0.032)	(0.035)	(0.030)
R^2 or K-P	94.211	19.762	23.552	132.482
Obs	6778	6778	6778	6778

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. This table replicates Appendix Table C.3 but using the administrative data on formal employment RAIS instead of the CENSUS.

Table O.15: Earnings Regression - Alternative RHS - OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.026**	0.021*	0.031*	0.031**	0.032**	0.032**
	(0.011)	(0.012)	(0.017)	(0.016)	(0.013)	(0.013)
Revenues		0.005				
		(0.006)				
Quantity			-0.005			
			(0.011)			
Land				-0.007		
				(0.012)		
Agr-group					-0.014	
					(0.018)	
Worker productivity in origin						-0.012
						(0.011)
R^2	0.699	0.699	0.699	0.699	0.699	0.699
Obs	6778	6778	6778	6778	6778	6778

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In column 5, we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane). Column 6 includes the

Table O.16: Farmers' Choice Regression - Alternative RHS - OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.084***	0.090***	0.087***	0.089***	0.105***	0.081***
	(0.016)	(0.019)	(0.020)	(0.021)	(0.020)	(0.017)
Revenues		-0.005				
		(0.008)				
Quantity			-0.003			
			(0.011)			
Land				-0.006		
				(0.016)		
Agr-group					-0.053**	
					(0.027)	
Worker productivity in origin						0.007
						(0.015)
R^2	0.755	0.755	0.755	0.755	0.756	0.755
Obs	6778	6778	6778	6778	6778	6778

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In columns 5 and 10 we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane).

Table O.17: Earnings Regression - Alternative RHS - OLS - RAIS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.010***	0.013***	0.016***	0.012***	0.011***	0.012***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Revenues		-0.003**				
		(0.001)				
Quantity			-0.005***			
			(0.001)			
Land				-0.003***		
				(0.001)		
Agr-group					-0.004	
					(0.003)	
Worker productivity in origin						-0.005***
						(0.001)
R^2	0.658	0.658	0.658	0.658	0.658	0.658
Obs	94472	94472	94472	94472	94472	94472

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In column 5, we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane). Column 6 includes the

Table O.18: Earnings Regression - Alternative RHS - PPML - RAIS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.010***	0.013***	0.016***	0.013***	0.011***	0.013***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Revenues		-0.002				
		(0.002)				
Quantity			-0.005***			
			(0.001)			
Land				-0.003*		
				(0.002)		
Agr-group					-0.005	
					(0.004)	
Worker productivity in origin						-0.005***
						(0.001)
R^2	0.207	0.207	0.207	0.207	0.207	0.207
Obs	94472	94472	94472	94472	94472	94472

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In column 5, we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane).

Table O.19: Farmers' Choice Regression - Alternative RHS - OLS - RAIS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.128***	0.124***	0.138***	0.128***	0.125***	0.131***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Revenues		0.003				
		(0.003)				
Quantity			-0.008***			
			(0.002)			
Land				-0.000		
				(0.003)		
Agr-group					0.013	
					(0.008)	
Worker productivity in origin						-0.006**
						(0.002)
R^2	0.815	0.815	0.815	0.815	0.815	0.815
Obs	94472	94472	94472	94472	94472	94472

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In columns 5 and 10 we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane).

Table O.20: Farmers' Choice Regression - Alternative RHS - PPML - RAIS

	(1)	(2)	(3)	(4)	(5)	(6)
Farmers in origin	0.208***	0.217***	0.219***	0.231***	0.205***	0.208***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)
Revenues		-0.008				
		(0.005)				
Quantity			-0.009***			
			(0.002)			
Land				-0.032***		
				(0.006)		
Agr-group					0.018	
					(0.019)	
Worker productivity in origin						0.001
						(0.002)
R^2	0.863	0.863	0.863	0.863	0.863	0.863
Obs	94472	94472	94472	94472	94472	94472

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. All specifications include destination-activity-year and destination-origin fixed effects. In columns 5 and 10 we aggregate crops into 4 groups: fruits (banana, cacao, and coffee), grains (corn, soy, and rice), and other (cotton, livestock, tobacco and sugarcane).

Table O.21: Earnings and Farmers' Choice Regression - State-level and Migration Definition

	OLS	OLS	PPML	PPML	2SLS	2SLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a. Income (logs)							
Farmers in origin	0.030***	0.039***	0.029*	0.038**	0.055	0.056**	0.056**
	(0.010)	(0.013)	(0.016)	(0.019)	(0.145)	(0.028)	(0.028)
R^2 or K-P	0.616	0.645	0.277	0.304	0.389	30.199	15.025
Overid. p							0.995
Obs	2672	2099	2672	2099	2672	2672	2672
b. Farmers in destination (logs)							
Farmers in origin	0.155***	0.231***	0.213***	0.247***	0.203	0.243***	0.245***
	(0.019)	(0.027)	(0.022)	(0.020)	(0.276)	(0.040)	(0.039)
R^2 or K-P	0.856	0.881	0.961	0.967	0.389	30.199	15.025
Overid. p							0.887
Obs	2672	2099	2672	2099	2672	2672	2672
Dest/Act/Year FE	Y	Y	Y	Y	Y	Y	Y
Dest/Orig/Year FE	Y	Y	Y	Y	Y	Y	Y
Above Q1	-	Y	-	Y	-	-	-
Crop Suitability	-	-	-	-	Y	Y	Y
Harris's IV	-	-	-	-	Y	-	Y
Mig. Comp. IV	-	-	-	-	-	Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. This table shows results using state-level aggregations instead of meso-regions. The definition of a migrant is based on the state of birth of an individual in the data set (instead of the previous meso-region, as in our baseline results).

Table O.22: Earnings and Farmers' Choice Regression - Dropping Crops

	Dep Var: Income					Dep Var: Farmers				
Dropped	PPML		2SLS			PPML		2SLS		
Actictivities	Coef	SE	Coef	SE	Obs	Coef	SE	Coef	SE	Obs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
a. RAIS										
none	0.010***	(0.002)	0.024***	(0.003)	94472	0.208***	(0.008)	0.228***	(0.009)	94472
cocoa	0.009***	(0.002)	0.021***	(0.003)	93649	0.213***	(0.008)	0.241***	(0.009)	93649
coffee	0.012***	(0.002)	0.024***	(0.004)	83581	0.227***	(0.008)	0.239***	(0.010)	83581
cotton	0.015***	(0.002)	0.027***	(0.003)	90070	0.223***	(0.008)	0.245***	(0.010)	90070
livestock	0.006**	(0.003)	0.021***	(0.005)	55865	0.172***	(0.009)	0.190***	(0.012)	55865
soy	0.015***	(0.003)	0.029***	(0.004)	78048	0.219***	(0.011)	0.211***	(0.013)	78048
sugarcane	0.005*	(0.003)	0.021***	(0.005)	71472	0.155***	(0.007)	0.212***	(0.010)	71472
tobacco	0.010***	(0.002)	0.024***	(0.003)	94147	0.209***	(0.008)	0.227***	(0.009)	94147
no cap-intensive	0.010**	(0.005)	0.035***	(0.008)	50646	0.099***	(0.012)	0.174***	(0.017)	50646
b. CENSUS										
none	0.047***	(0.016)	0.052***	(0.015)	6778	0.129***	(0.021)	0.133***	(0.034)	6778
banana	0.047***	(0.016)	0.052***	(0.015)	6504	0.132***	(0.022)	0.142***	(0.034)	6504
cocoa	0.040**	(0.017)	0.061***	(0.019)	6700	0.138***	(0.020)	0.144***	(0.045)	6700
coffee	0.057***	(0.018)	0.053***	(0.017)	5613	0.104***	(0.026)	0.154***	(0.041)	5613
corn	0.047***	(0.017)	0.051***	(0.017)	6424	0.141***	(0.022)	0.139***	(0.035)	6424
cotton	0.044***	(0.017)	0.052***	(0.015)	6654	0.129***	(0.022)	0.138***	(0.034)	6654
livestock	0.043**	(0.019)	0.066**	(0.028)	3918	0.161***	(0.032)	0.179***	(0.055)	3918
rice	0.050***	(0.017)	0.058***	(0.016)	6289	0.133***	(0.022)	0.128***	(0.035)	6289
soy	0.037**	(0.018)	0.037***	(0.013)	6538	0.124***	(0.023)	0.104***	(0.031)	6538
sugarcane	0.056***	(0.020)	0.054***	(0.018)	5753	0.094***	(0.022)	0.131***	(0.043)	5753
tobacco	0.045***	(0.017)	0.044***	(0.013)	6609	0.130***	(0.022)	0.110***	(0.029)	6609
no cap-intensive	0.039	(0.025)	0.037**	(0.016)	5389	0.081***	(0.025)	0.097**	(0.038)	5389

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level.

Table O.23: Earnings and Farmers' Choice Regression - Alternative explanations

	PPML (1)	Previous Network PPML (2)	Controls for SES PPML (3)	2SLS (4)	Previous Network 2SLS (5)	Controls for SES 2SLS (6)
a. RAIS						
Farmers in origin	0.010***	0.010***		0.024***	0.024***	
	(0.002)	(0.002)		(0.003)	(0.003)	
R^2 or K-P	0.207	0.207		308.511	309.660	
Obs	94472	94472		94472	94472	
b. CENSUS						
Farmers in origin	0.047***	0.045***	0.049***	0.052***	0.048***	0.046***
	(0.016)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)
R^2 or K-P	0.338	0.338	0.346	23.683	25.003	24.089
Obs	6778	6778	6778	6778	6778	6778
Dest/Act/Year FE	Y	Y	Y	Y	Y	Y
Dest/Orig/Year FE	Y	Y	Y	Y	Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. Our dataset from RAIS does not include socio-economic status variables, we therefore only include such controls for the regressions using CENSUS data.

Table O.24: Earnings Regression - Evaluating experience - RAIS

	No Experience		With Experience	
	PPML (1)	2SLS (2)	PPML (3)	2SLS (4)
a. Income (logs)				
Farmers in origin	0.006*	0.011***	0.006**	0.025***
	(0.003)	(0.004)	(0.002)	(0.003)
R^2 or K-P	0.198	170.173	0.208	181.591
Obs	53967	53967	78735	78735
b. Farmers in destination (logs)				
Farmers in origin	0.093***	0.097***	0.246***	0.272***
	(0.007)	(0.010)	(0.009)	(0.011)
R^2 or K-P	0.685	170.173	0.855	181.591
Obs	53967	53967	78735	78735
Dest/Act/Year FE	Y	Y	Y	Y
Dest/Orig/Year FE	Y	Y	Y	Y
Suitability in Origin		Y		Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. Instrumental variable results use both instruments, that is, the access to workers from other regions and the migrants' composition.

Table O.25: Earnings and Workers' Choice Regression - Manufacturing

	Dep Variable	
	Income PPML (1)	Number of Workers PPML (2)
Workers in origin	0.059***	0.156***
	(0.020)	(0.018)
R^2	0.354	0.727
Obs	7266	7266
Dest/Act/Year FE	Y	Y
Dest/Orig/Year FE	Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level. The manufacturing activities included are: automotive, leather, furniture, processed tobacco, oil, paper, clothing, textile, perfume, and wood.

Table O.26: Earnings and Farmers' Choice Regression - Workers' Composition in the Origin

	OLS (1)	OLS (2)	PPML (3)	PPML (4)	2SLS (5)	2SLS (6)	2SLS (7)
a. Income (logs)							
Farmers in origin	0.032**	0.049**	0.047**	0.078***	0.076***	0.155**	0.081***
	(0.013)	(0.020)	(0.020)	(0.029)	(0.024)	(0.077)	(0.024)
R^2 or K-P	0.699	0.724	0.338	0.257	23.064	15.715	11.589
Overid. p							0.209
Obs	6778	5262	6778	5262	5241	6778	5241
b. Farmers in destination (logs)							
Farmers in origin	0.074***	0.080***	0.119***	0.113***	0.156***	0.265*	0.162***
	(0.016)	(0.025)	(0.023)	(0.038)	(0.047)	(0.146)	(0.047)
R^2 or K-P	0.757	0.778	0.784	0.802	23.064	15.715	11.589
Overid. p							0.394
Obs	6778	5262	6778	5262	5241	6778	5241
Dest/Act/Year FE	Y	Y	Y	Y	Y	Y	Y
Dest/Orig/Year FE	Y	Y	Y	Y	Y	Y	Y
Above Q1	-	Y	-	Y	-	-	-
Crop Suitability	-	-	-	-	Y	Y	Y
Harris's IV	-	-	-	-	Y	-	Y
Mig. Comp. IV	-	-	-	-	-	Y	Y
Orig - Workers' Q Prod	Y	Y	Y	Y	Y	Y	Y
Orig - Workers' R Prod	Y	Y	Y	Y	Y	Y	Y
Orig - Harris's Q Prod	-	-	-	-	Y	-	Y
Orig - Harris's R Prod	-	-	-	-	Y	-	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Multiway clustered standard errors clustered at destination-activity-year and origin-year level.

Table O.27: Calibration of the Lineage Model

	Specification					
	(1)	(2)	(3)	(4)	(5)	(6)
a. Coefficients						
β	0.050	0.034	0.003	0.080	0.064	0.032
ς	1.000	1.500	2.000	1.000	1.500	2.000
RF coef	0.050	0.050	0.050	0.080	0.080	0.080
b. Intergenerational Occupation Mobility						
No switch between activities	0.599	0.781	0.862	0.599	0.780	0.861

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. This table shows different calibrations of the "lineage" model, in which parents pass along to their children knowledge about their own sector. Columns 1 to 6 calibrate the model and the knowledge externality parameter β under different values of ς, the premium for staying in the same activity as your parents, so that the model generates a reduced-form coefficient of log of income with respect to the labor force in the origin of 0.05 and of 0.08.

Table O.28: Migration Costs and Travel Distance

	(1)	(2)	(3)
Log of travel time	0.267***	0.366***	0.475***
	(0.016)	(0.025)	(0.028)
R2 or K-P	0.434	0.903	375.855
within R2		0.499	
Obs	2028	2028	2028
Origin-Destination FE		Y	Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Dependent variable is the estimated symmetric component of the migration costs ($\tilde{\mu}_{ss',t}$) implied by the estimation of the migration gravity equation (28). Travel distance between states is measured based on the infrastructure of highways available at time t (see Figure (O.5)). Column 3 uses the radial lines constructed to connect Brasília to the rest of Brazil as in Morten and Oliveira (2016).

Table O.29: Crop Choices over Time across Age Groups for 2017 and 2010

	Age Group			
	16 to 25	25 to 40	40 to 65	65+
	(1)	(2)	(3)	(4)
Panel a. Farmers in 2017				
% of years working on the same crop	96.1	91.7	92.5	95.2
Average number of years in RAIS	2.45	5.1	6.6	7.6
Observations	335363	948287	900762	25668
Panel b. Farmers in 2010				
% of years working on the same crop	95.1	89.1	87.7	86.7
Average number of years in RAIS	2.37	5.1	6.4	7.6
Observations	438519	1071505	822869	16739

Notes: The table displays the percentage of time that farmers in the 2017 and 2010 RAIS dataset spent working on the years on the same crop, based on the crop that they spent most time working, averaged over the preceding ten years in which they were found in the RAIS data. The table reports the average number of years during which the worker appeared in the RAIS database ten years prior to 2017 and 2010.

Table O.30: Knowledge Portability - Effects on the West

	Portability to the West		
	25%	Crop-specific	
	Δ RBE	Δ RBE	$\bar{\phi}_k$
Crop	(1)	(2)	(3)
soy	-33.5	-65.1	0.90
corn	-15.5	-3.5	0.68
beef	-18.9	-1.2	0.24
sugarcane	-23.7	-16.9	0.17
cotton	-24.0	-44.3	0.87
coffee	-22.9	-34.6	1.00
rest of agriculture	-10.3	-3.5	1.00
cacao	-22.5	2.2	0.64
rice	-11.4	-0.6	0.95
banana	-12.9	-17.2	0.91
tobacco	2.1	1.2	0.82

Notes: This table shows the aggregate impact on RBE of limiting the portability of knowledge from the East to the West of the country between 1950 and 2010. See discussion in Section (6.4).

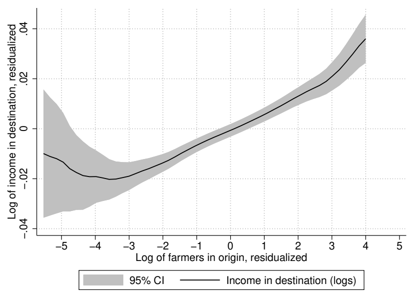
Table O.31: Migration to the West in different Counterfactuals

	No roads to W	No INCRA	No Incra & No roads to W	No road expansion	No mig cost reduction
	(1)	(2)	(3)	(4)	(5)
Share of baseline	0.306	0.030	0.328	0.515	0.595
Share of no change in mig cost	0.515	0.050	0.552	0.865	1.000

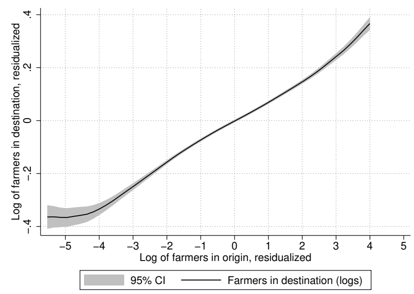
Notes: This table shows, on the first row, the share of the total migration to the West between 1950 and 2010 that can be attributed to the counterfactual change in fundamental described in the column. Specifically, for the first row, we measure the p.p. increase in migration from 1950 and 2010 (=0.0854), and we divide the p.p. increase in migration under the counterfactual by that number (for example, for column 1 first row, we compute $0.306 = 0.0261/0.085$). For the second row, we divide all values by the share of migration that is explained by the reductions in migration costs in column 5 (for example, for column 1 second row, we compute 0.515=0.306/0.595).

Figure O.1: Local Polynomial Regressions of the Influence of the Region of Origin on Crop Choice and Income of Farmers in their Destination Region

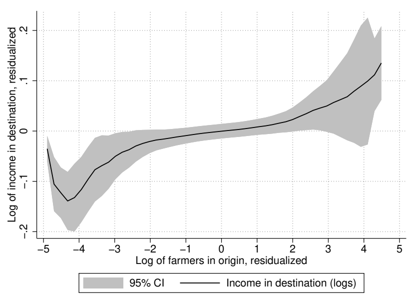
(a) Income (RAIS)



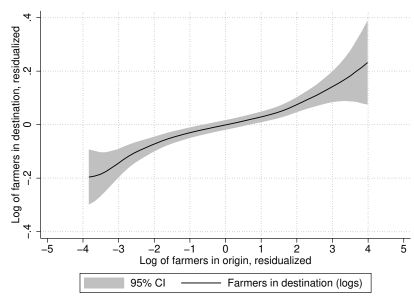
(b) Farmers (RAIS)



(c) Income (CENSUS)



(d) Farmers (CENSUS)

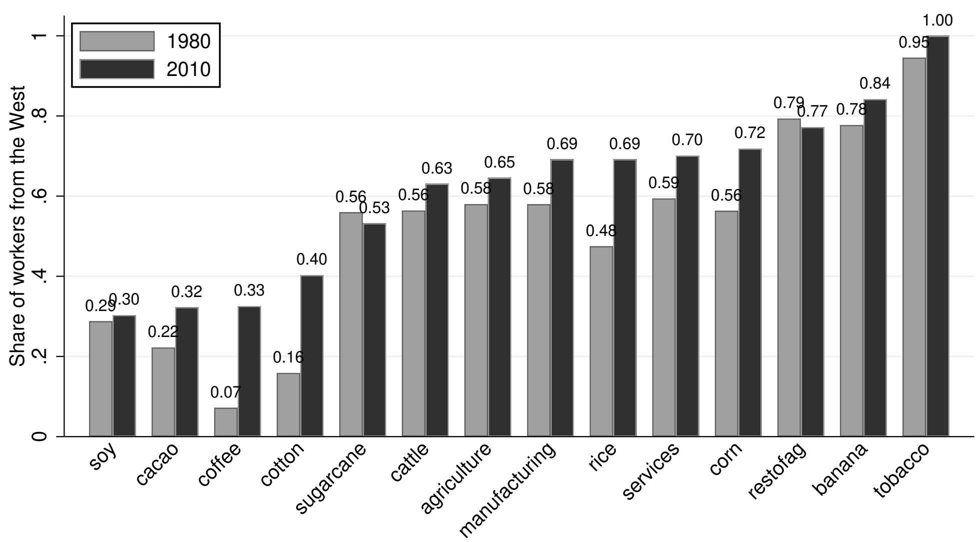


Notes: To construct this figure, we first absorb origin-destination-year and destination-crop-year fixed effects from dependent and independent variables of interest in equations 2 and 3. Using the residuals from these variables, we run a non-parametric regression using a local polynomial smooth.

Figure O.2: Federal Government Propaganda about the March to the West, 1940s

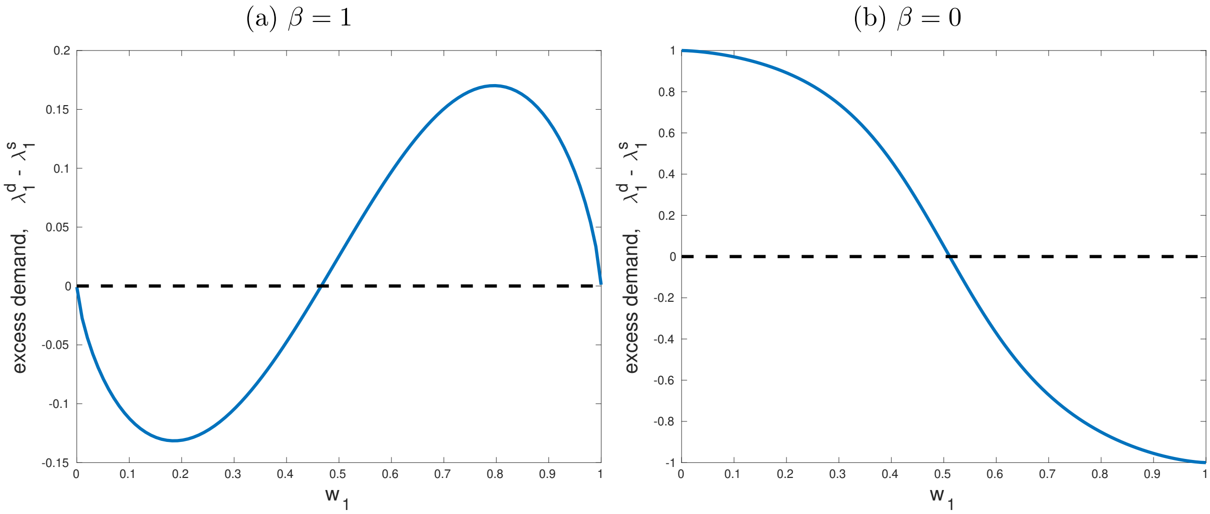
Notes: Poster features Getulio Vargas, (president 1930-1945 and 1951-1954). The quote in the bottom translates to "The true meaning of Brazilianness is the March to the West". This quote comes from one of his famous speeches, later named the "The speech at midnight" ("*O discurso da meia noite*") given at midnight on December 31st 1937 from Guanabara Palace - Getulio's official residence - and transmitted via the national radio (Vargas, 1938).

Figure O.3: The Share of Workers in the West who were born there, by Economic Activity



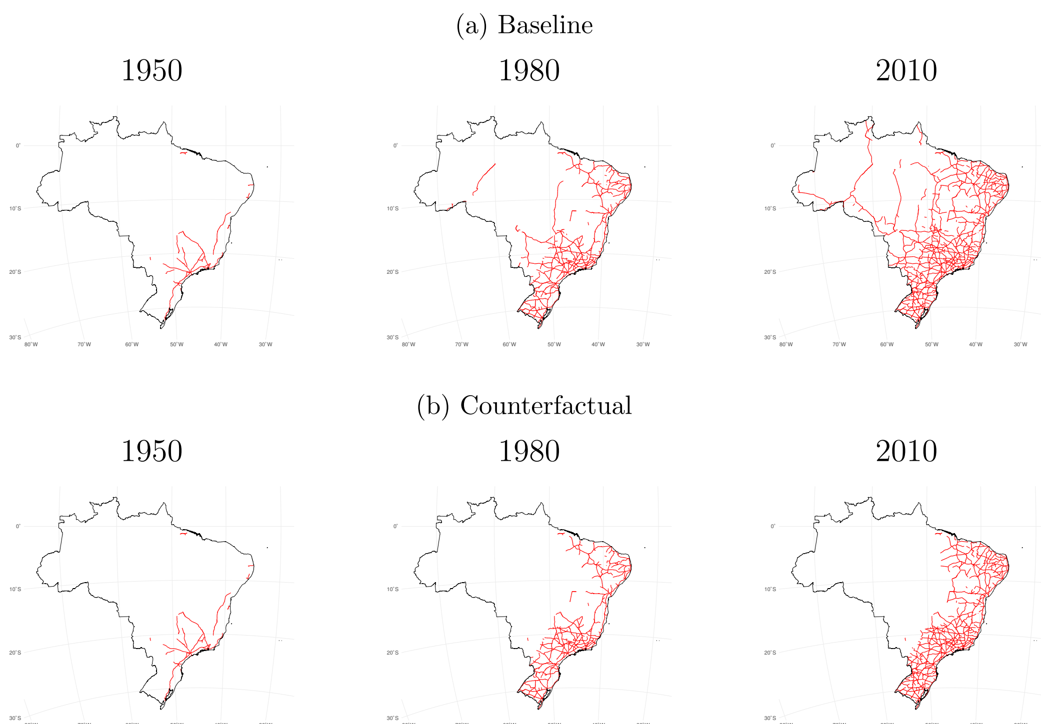
Notes: The figure presents the fraction of the total employment in the West, in each year and economic activity, comprised by workers born in the West.

Figure O.4: An example of multiplicity



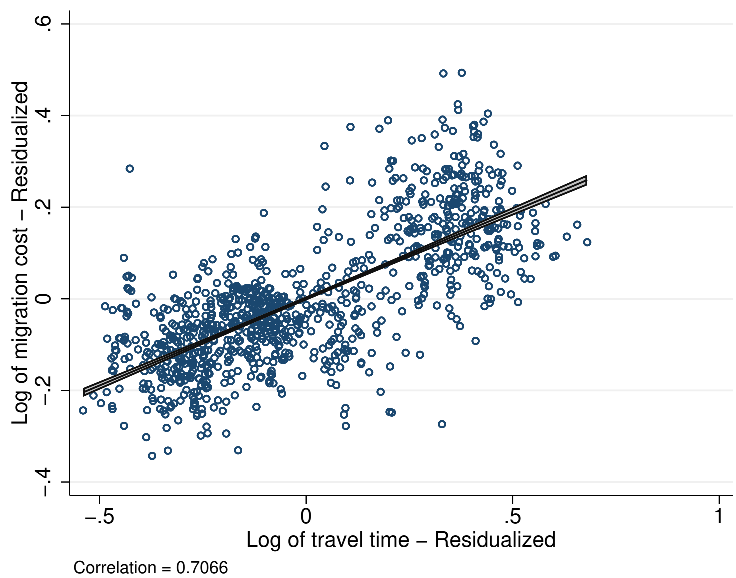
Notes: The figure presents the fraction of the total employment in the West, in each year and economic activity, comprised by workers born in the West. It shows that an increase in the knowledge externality can generate multiplicity of equilibria. In the case above, one of the equilibria is superior to the other, indicating the possibility of poverty traps in the economy. See discussion in Online Appendix (OF.2).

Figure O.5: Highway Expansion



Notes: Panel (a) shows the expansion of highways in Brazil between 1950 and 2010. Panel (b) shows the expansion of highways used for the counterfactual analysis in which we assume to expansion of highways towards the West of Brazil. Using this highway infrastructure, we apply the Fast Marching Method (FMM) to measure the travel distance between tthe centroid of any two meso-regions in our dataset.

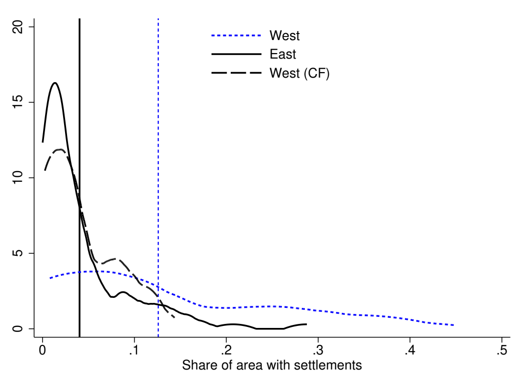
Figure O.6: Migration Costs and Travel Distance



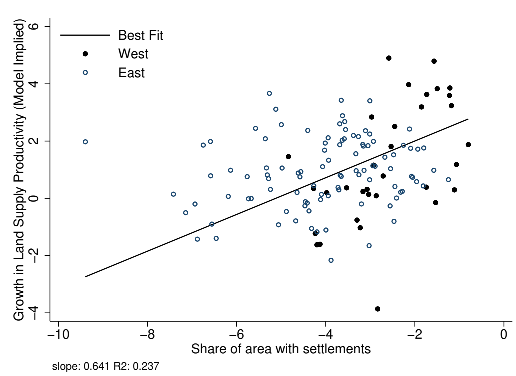
Notes: This figure shows estimates of the symmetric component of migration cost ($\tilde{\mu}_{ss',t}$) estimated based on (28) against the travel distance between states — the average across meso-regions. We residualize both the dependent variable and the explanatory variable based on destination and origin fixed effects. The figure therefore captures the correlation between changes in migration costs against changes in the travel time.

Figure O.7: Land Productivity and Land Settlement

(a) Land Share Distribution



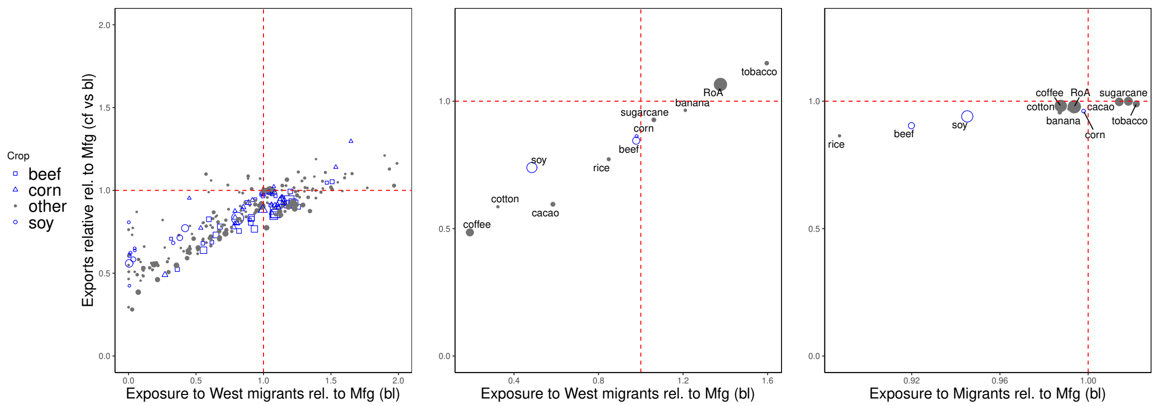
(b) Correlation



Notes: Panel (a) shows the distribution of land settlements across regions in the West and in the East of Brazil. Additionally, it shows the counterfactual distribution of land settlements, if the average number had been the same between the East and the West. Panel (b) shows the correlation between land settlements and the growth in the productivity of land supply (for 2010). The two vertical lines mark the average of the distribution.

Figure O.8: Counterfactual Change in Exports relative to Manufacturing (1980), for the West and Brazil as a whole

(a) Meso-regions - West (b) West (c) Brazil



Notes: This figure replicates the results from Figure 3 in the main body of the text for the year of 1980. See notes from 3 for details.

Figure O.9: The Impact of Integration Policies on the Comparative Advantage of the West

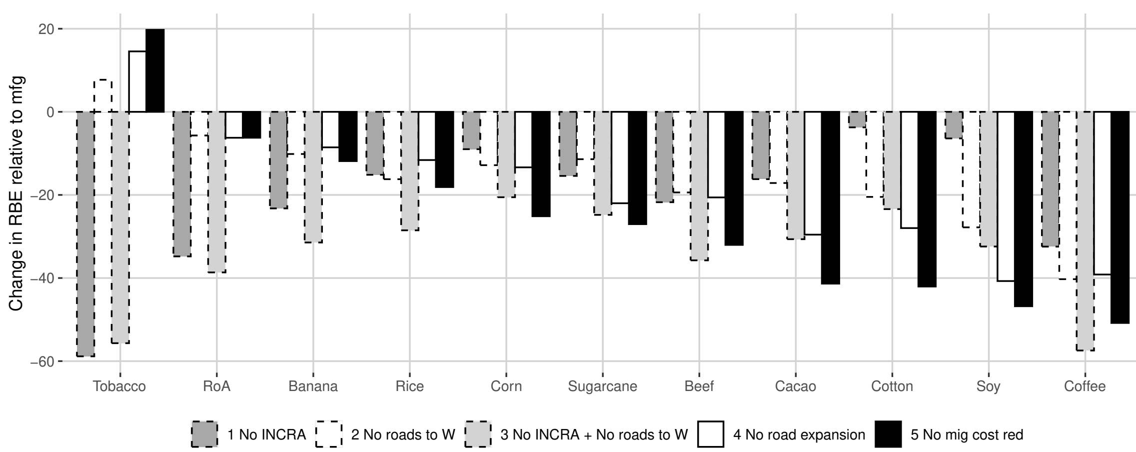
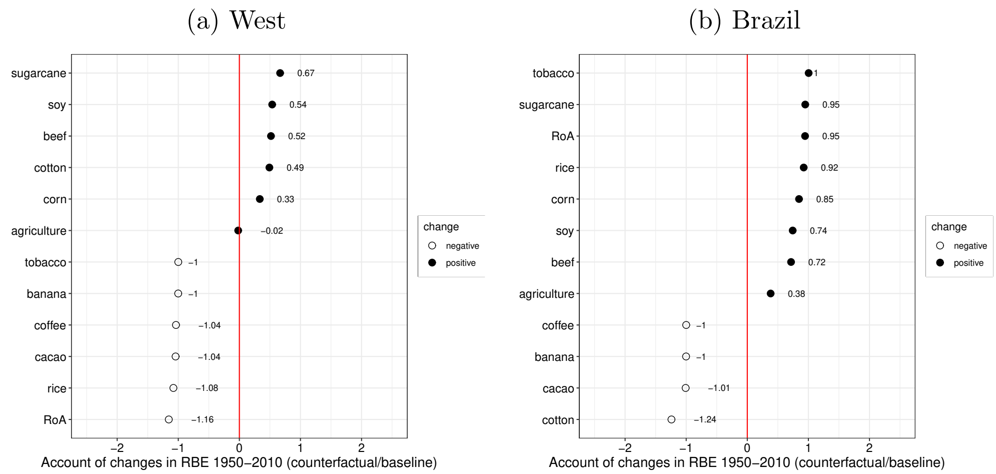


Figure O.10: Accounting for Observed Changes in Specialization (2010)

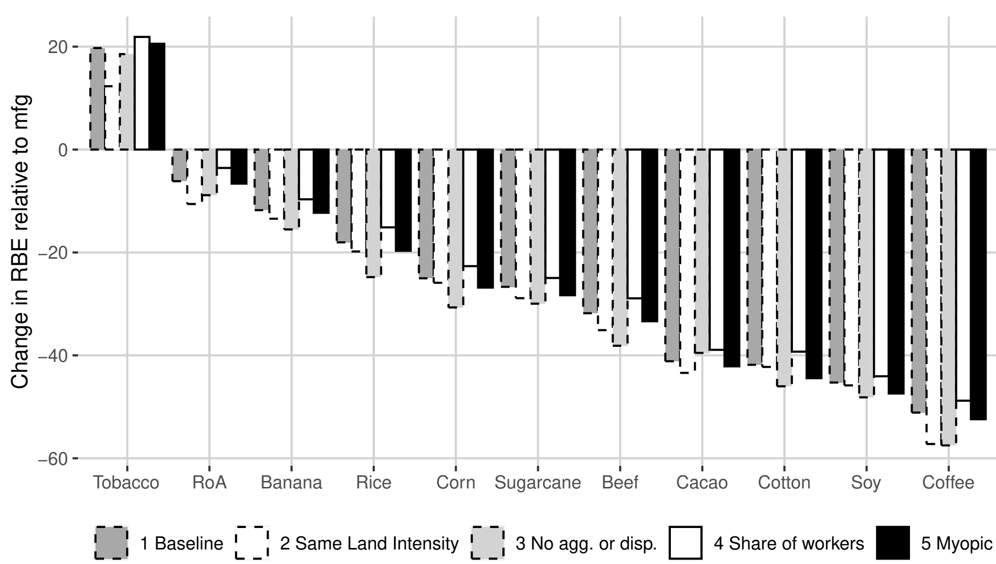


Notes: Each panel plots, for each activity k

$$\frac{RBE^{\text{counterfactual}}_{BF,kMfg,2010}/RBE^{\text{baseline}}_{BF,kMfg,1950}-1/}{|RBE^{\text{baseline}}_{BF,kMfg,2010}/RBE^{\text{baseline}}_{BF,kMfg,1950}-1|}.$$

Activities in black circles are those in which RBE grew over the period, those in white circles shrank. These figures capture the share of the total change in relative bilateral export accounted by the counterfactual, in which we keep migration costs between the East and the West at 1950s level.

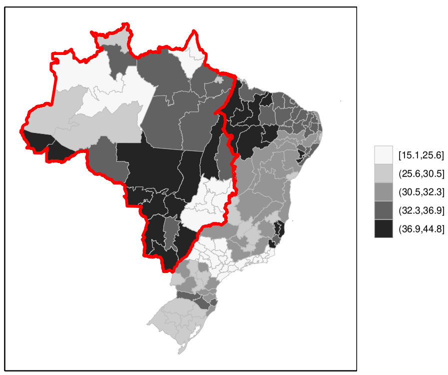
Figure O.11: Alternative Parametrizations of the Model - Impact of No Migration to the West on the West's specialization



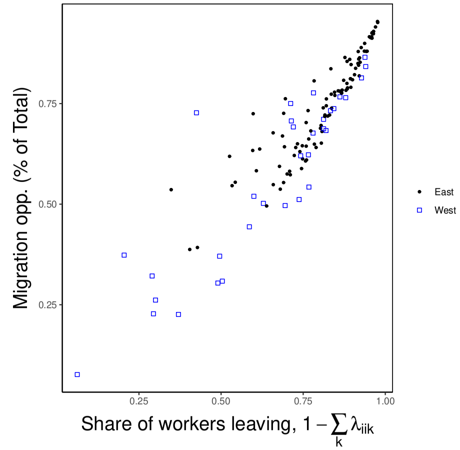
Notes: This figure shows the impact on the West's specialization of no reductions in migration costs between the East and the West of Brazil under alternative parametrizations of the model. It shows results for (1) our baseline calibration and calibrations in which (2) the land-intensity is the same between manufacturing, services and agricultural activities ($\gamma_k = \gamma$), (3) there are no agglomeration or congestion forces ($\xi = \chi = 0$), (4) knowledge externality is based on the share of workers instead of the scale ($s_{ik,t} = \bar{s}_k L_{ik,t-1}/L^{\beta}_{i,t-1}$), and (5) agents are myopic ($\delta = 0$).

Figure O.12: Comparative Advantage, Migration, and the Losses from Autarky (2010)

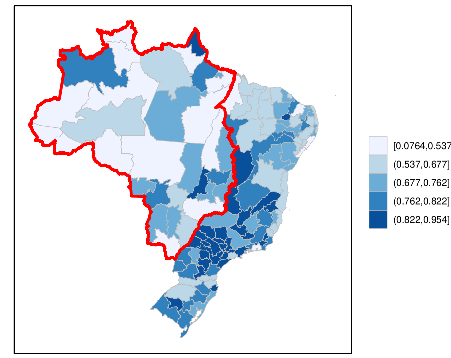
(a) Full Gains from Trade



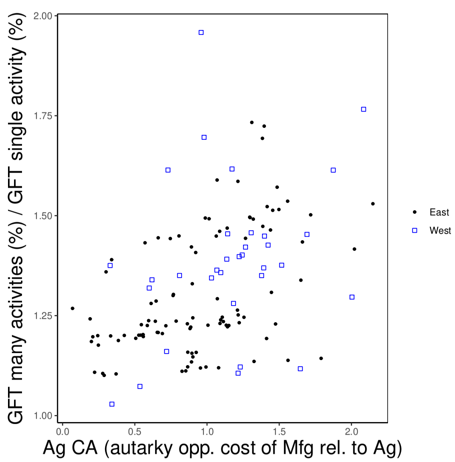
(b) Contribution of Mig. Opp.



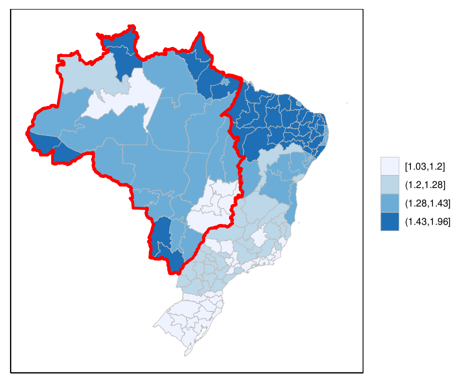
(c) Contribution of Mig. Opp. (Map)



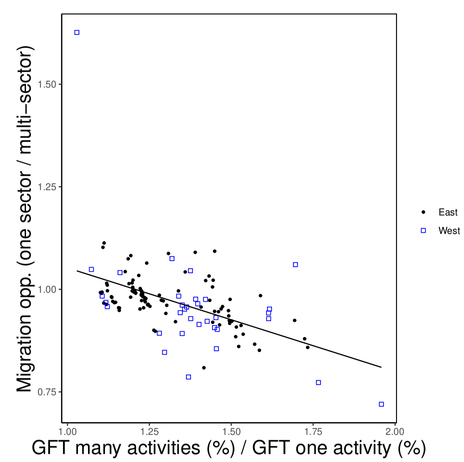
(d) CA and the GFT



(e) Contribution of CA (Map)



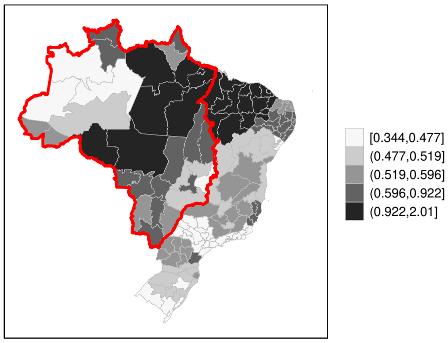
(f) The Interaction of Mig. and CA



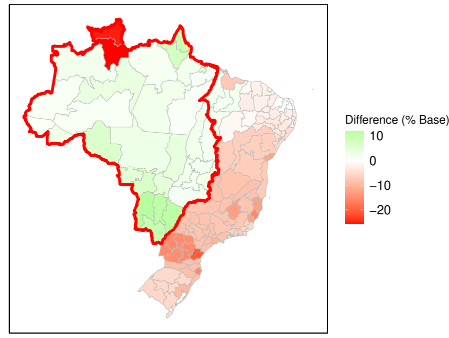
Notes: All simulations are for 2010 and each observation is a region. Panel (a) shows the welfare losses from letting each region go to full trade autarky, as a percent of baseline welfare. Panel (b): The horizontal axis is the fraction of workers leaving that region in the baseline simulation. The vertical axis measures the ratio of the welfare cost that results solely from migration opportunities (i.e. setting the domestic contribution to zero) to the total costs and taking the ratio of $\hat{\Xi}_i$. Panel (c): shows the vertical axis of Panel (b). Panel (d): The vertical axis plots the ratio of the GFT with many activities to the GFT with one activity, both as percent of baseline. The horizontal axis shows the autarky opportunity cost of manufacturing relative to agriculture (measured as relative price indexes in autarky). Panel (e) shows the horizontal axis in Panel (d). Panel (f): The vertical axis is the same as in in Panel (d). The horizontal axis shows the change in the contribution of migration opportunities in going from a one-activity to a many-activity version of our model.

Figure O.13: The March to the West and the International Gains from Trade (2010)

(a) The Gains from International Trade



(b) Counterfactual Changes in International GFT



Notes: Panel (a) shows for each region the gains from trade with the rest world, defined as the welfare cost of prohibiting foreign trade only (but allowing domestic trade). Panel (b) subtracts the baseline gains from trade from the gains from trade in the counterfactual scenario (no East-West migration).

Appendix References

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

Allen, T. and D. Donaldson (2022). Persistence and path dependence in the spatial economy.

Alvarez, F. and R. J. Lucas (2007, September). General equilibrium analysis of the eaton-kortum model of international trade. *Journal of Monetary Economics 54*(6), 1726–1768.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012, February). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019). Migrants, Ancestors, and Foreign Investments. *Review of Economic Studies 86*(4), 1448–1486.

Bustos, P., B. Caprettini, and J. Ponticelli (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review 106*(6), 1320–65.

Dix-Carneiro, R. and B. K. Kovak (2017, October). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–46.

Farrokhi, F. and H. S. Pellegrina (2023). Trade, technology, and agricultural productivity. *Journal of Political Economy - Forthcoming*.

Galle, S., A. RodrÃguez-Clare, and M. Yi (2023). Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade. *Review of Economic Studies 90*(1), 331–375.

Ireland, C. T. and S. Kullback (1968). Contingency tables with given marginals. *Biometrika 55*(1), 179–188.

Levchenko, A. A. and J. Zhang (2016). The evolution of comparative advantage: Measurement and welfare implications. *Journal of Monetary Economics 78*(C), 96–111.

Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in great britain and the united states since 1850. *American Economic Review 103*(4), 1109–1137.

Morten, M. and J. Oliveira (2016). The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper 22158*.

Ramondo, N., A. Rodriguez-Clare, and M. Saborío-Rodríguez (2016). Trade, domestic frictions, and scale effects. *American Economic Review 106*(10), 3159–84.

Vargas, G. (1938). *A nova politica do Brasil - Vol V*.

Vasconcelos, J. R. d. (2001). Matriz do fluxo de comércio interestadual de bens e serviços no brasil-1999. Technical report.