Migration, Specialization, and Trade: Evidence from the Brazilian March to the West

Heitor S. Pellegrina
NYU Abu Dhabi

Sebastian Sotelo
University of Michigan*

April 2020

Please click here for most recent version

Abstract

Exploiting a large migration of farmers to the West of Brazil between 1950 and 2010, we study how internal migration shapes aggregate and regional comparative advantage. We document that farmers emigrating from regions with high employment in a given crop are more likely to grow that crop and have higher earnings than other farmers doing so. We incorporate this heterogeneity into a quantitative model of trade and migration. By reshaping Ricardian and Heckscher-Ohlin comparative advantage, the migration cost decline we observe contributed substantially to Brazil's rise as a leading commodity exporter. A large part of this effect comes from the reallocation of knowledge carried by migrants.

Keywords: International Trade, Migration, Comparative Advantage

*We thank Jonathan Eaton and Andrei Levchenko for helpful comments, as well as Walter Steingress and João Pessoa for useful discussions. We also thank participants in seminars at Purdue, Michigan, NYU Abu Dhabi, St. Louis Fed, IADB, TIGN-LACEA, Middlebury, RMET, UEA, FGV/EESP, Banco Central de Colombia, Insper, UFABC, Banco Central de Chile, UCLA, Berkeley ARE and the World Bank. Natalia Martinez Camelo and Brian Cevallos Fujiy provided excellent research assistance. Financial support from the Michigan Institute for Teaching and Research in Economics (MITRE) is gratefully acknowledged. E-mail: ssotelo@umich.edu and heitor.pellegrina@nyu.edu. All errors are our own.

1 Introduction

A central task in international trade, and in spatial economics more generally, is to understand the impact of trade on welfare and the patterns of specialization across locations. Understanding the impact of trade, moreover, goes beyond purely academic interest, as policymakers often attempt to influence their own countries' export and import activity, against the backdrop of increasing globalization. Seeking to quantify the impact of trade, a recent literature has incorporated comparative advantage—the notion that differences in relative costs across locations drive trade—into quantitative models and established it as a major determinant of trade flows.[1] But while recent work has documented that comparative advantage itself evolves over time, there has been comparatively less progress in quantifying the mechanisms that drive that evolution.

In this paper, we start by noting that large migrations within countries are common, and that they are often accompanied by shifts in specialization and trade patterns. Consider, for example, the U.S. westward expansion and, more recently, the large migration of Chinese workers to export clusters. The question we ask, based on this observation, is: Does the migration of workers within a country shape regional and aggregate comparative advantage? We consider three mechanisms. First, migration determines the allocation of labor across regions that differ in their natural advantage across goods. Second, migration alters the relative abundance of land and labor across regions. Third, heterogeneous workers sort across regions and goods according to their own comparative advantage, which affects the labor supply composition across sectors and regions.[2]

To answer our question, we extend a dynamic model of trade and migration to incorporate these three mechanisms. In our model, workers are endowed with good-specific knowledge, which they acquire through exposure to economic activity in their origin region, and choose the location and activity that is best suited to that knowledge. In equilibrium, regional and aggregate comparative advantage reflect a combination of natural advantage (such as land quality), the abundance of labor, and the knowledge of the labor force.

We quantify the mechanisms in our model by applying it to Brazilian agriculture, which

[1]See Eaton and Kortum (2012) and Costinot and Rodriguez-Clare (2015) for a review of the literature on quantitative trade models.

[2]For example, Sabel, Fernandez-Arias, Hausmann, Rodriguez-Clare, and Stein (2012) describe how migrants have used their knowledge to form new export sectors in Latin America. Opala (1987) documents that plantation owners in South Carolina and Georgia were willing to pay higher prices for slaves from this Sierra Leone and Liberia, denominated the "Rice Coast", due to their knowledge about the production of rice. Other historical examples include the diffusion of crops during the Columbian Exchange (Crosby, 1973), the introduction of new varieties of wheat in the northeast of the US in the 19th century (Olmstead, Rhode, et al., 2008), the introduction of wheat in North Africa during the diffusion of Islam (Watson, 1983), and the production of flowers by Dutch refugees in England in the late 16th century (Scoville, 1951).

provides us with a rich setting to explore the relation between migration and trade. In the second half of the 20th century, a large number of Brazilian workers migrated to the low-density areas in the Cerrado and the Amazon; this episode, often called the "March to the West", coincided with Brazil's transformation into an important global exporter of crops such as soy, corn, and livestock. Following a series of public initiatives to integrate the country's West to regions in the East between the 1950s and the 1970s, farmers from all parts of Brazil migrated to the West, raising the share of Brazil's population living in the West from 6 percent in 1950 to 15 percent in 2010.[3]

For our analysis, we assemble a detailed data combining several waves of Brazil's demographic census with disaggregate information on trade and internal migration since the 1950s, and use it to document four facts that guide our modeling approach. First, Brazil's external trade shifted dramatically since the 1950s: new agricultural goods such as soy, corn and livestock captured a large share of Brazil's exports (compared to the rest of the world), reflecting the intensity with which Brazil's West specializes in these goods. Second, the West's shares in population, land, and agricultural value added roughly doubled during the March to the West. Third, there is a strong link between a migrant's origin and her activity choices and earnings. Comparing farmers who, upon migrating, produce the same agricultural goods in the same destination, but who emigrate from different regions, we find that a 1 percent increase in the number of farmers producing an agricultural good in the origin region is associated with a 0.06-0.11 percent increase in the number of emigrants from that region producing that same good, as well as a 0.02-0.07 percent increase in the earnings of that region's emigrants. Fourth, for a given agricultural good, regional output and revenues increase when the regional mix of workers favors workers from regions that employ more farmers in that same good, controlling for the number of farmers.

Motivated by these facts, we introduce heterogeneous workers whose knowledge differs across sectors, into a dynamic model of trade and migration with overlapping generations. We characterize analytically the impact of migration costs on comparative advantage and find that, from one period to the next, the allocation of workers and their knowledge amplify regional and national comparative advantage with respect to the rest of the world. If workers are heterogeneous, moreover, migration has an additional, ambiguous effect on comparative advantage: By spreading knowledge across regions, migration undoes the ex-ante allocation of worker knowledge, which may strengthen or weaken comparative advantage.

We next bring our model to the data. We show that the two key parameters controlling

[3] The magnitude of Brazilian march to the west is comparable to the US Great Migration during the 20th century, a period in which roughly 6 million Americans emigrated from the south of the US. See Bazzi, Gaduh, Rothenberg, and Wong (2016) for an analysis of the Transmigration Program in Indonesia, which involved 2 million migrants.

worker heterogeneity and migration decisions, (i) the elasticity of a farmer's productivity to crop employment in her region of origin and (ii) the elasticity of occupational choice to income can be transparently identified from reduced form elasticities. We calibrate the rest of parameters by combining the model with state and meso-region level data on Brazilian cropping patterns, internal migration, employment and exports for 1950, 1980 and 2010, effectively setting a generation in our model as 30 years.

Having calibrated the model and estimated its key parameters, we conduct an experiment to gauge the impact of migration cost reductions on specialization and trade. Specifically, we shut down the evolution of migration costs since the 1950s, while allowing all other exogenous factors to evolve over time, and compare the evolution of this counterfactual economy to our baseline simulation. Because this scenario limits migration to the West, the resulting increase in the share of workers living there between 1950 and 2010 is half of that in the data.

Migration substantially reshapes regional comparative advantage in this experiment, especially for the West. Across Western regions, a measure of comparative advantage consistent with our theory —export specialization relative to manufacturing and to RoW— decreases by as much as 30% in the counterfactual, compared to our baseline. The baseline share of the workforce in each region and activity that comes from the East is strongly associated with these changes in specialization, since it relates to the direct impact of limiting migration on relative marginal costs. In Eastern regions, in turn, which were already densely populated in 1950, the differential impact of migrants from the West across sectors is more muted.

For Brazil as a whole, reductions in migration costs account for a sizable part of observed changes in comparative advantage. Observed changes in migration costs account for about 70% of the increase in Brazil's revealed comparative advantage in agriculture between 1950 and 2010. For soy, livestock, and corn, Brazil's new flagship commodities, we find that aggregate comparative advantage would have been between 30 and 40 percent smaller had migration costs not dropped.

Our quantification also allows us to assess each of our margins of comparative advantage. Land being abundant in the West, Heckscher-Ohlin forces act as a buffer for land-intensive crops when labor becomes scarce due to limited migration. This dampening is especially true for livestock, a particularly land intensive crop. Second, we re-evaluate the impact of migration after shutting down worker heterogeneity. We find that workers heterogeneity accounts for 5 to 15 percent of our aggregate results, especially among new export goods.

While migration enabled large swings in specialization within the agricultural sector, on aggregate these reallocations allowed for a small increase in import shares of non-agricultural goods, and therefore its impact on the gains from trade with the rest of the world are small.[4]

[4] The gains from allowing for internal migration, in turn, are about half as large as the gains from trade.

But this aggregate result masks large interactions between migration and the gains from trade across regions, which largely offset each other in the aggregate. Compared to the low-migration counterfactual, in the baseline some regions grow as workers migrate there and, in consequence, their gains from trade decrease as they become relatively closed. We find, however, that regions whose comparative advantage in agricultural goods was strengthened by migration strongly deviate from this relation.

Our paper relates to three strands of literature. The first is a recent trade literature that quantitatively studies the evolution of productivity differences across countries. Levchenko and Zhang (2016) and Hanson, Lind, and Muendler (2015) document substantial changes in Ricardian comparative advantage over time and cross countries. Buera and Oberfield (2016) and Cai, Li, and Santacreu (2019), among others, study the diffusion of ideas in an open economy and how it drives trade across countries (see Lind and Ramondo, 2018 for a summary of this literature). Arkolakis, Lee, and Peters (2018) study the impact of migrants on the technological frontier in the United States in the 19th century. More broadly, we relate to a quantitative literature studying the interaction of goods and people mobility, which includes Allen and Arkolakis (2014), Redding (2016), Bryan and Morten (2015), Tombe and Zhu (2019) and Morten and Oliveira (2016). We add to this literature by examining how migration shapes aggregate and regional comparative advantage, which we do building on the recent dynamic approaches of Caliendo, Dvorkin, and Parro (2015) and Allen and Donaldson (2018).

Second, we relate to a growing literature quantitatively examining the determinants and implications of international trade in agriculture, including Costinot and Donaldson (2014), Costinot, Donaldson, and Smith (2016), Fajgelbaum and Redding (2014), Allen and Atkin (2016), Pellegrina (2019), Porteous (2019), Tombe (2015), Porteous (2020), Farrokhi and Pellegrina (2019), Gouel and Laborde (2018) and Sotelo (2020). Most of this literature treats comparative advantage as exogenous, arising from quality of land, factor proportions, or both. We contribute to this literature, first, by studying how migration shapes these traditional forces of comparative advantage. Second, we show that migration introduces a new source of comparative advantage: the domestic allocation of knowledge.

Third, our paper relates to research documenting the consequences of migration for migrants, as well as origin and destination regions, when workers differ in their ability across activities. Using exogenous variation from a government-led population resettlement program in Indonesia, Bazzi, Gaduh, Rothenberg, and Wong (2016) show that differences in land suitability between origin and destination condition migrants' performance. In the agricultural context, Olmstead and Rhode (2011) have documented the role of geography and migration in the expansion of different crops in the US. More broadly, recent work has estimated the effects of a workers' past environment on his learning and productivity, including Sviatschi et al. (2018) and de la Roca and Puga (2017). Closely related to our empirical findings, Bahar and Rapoport (2016) provide evidence that international migration can affect comparative advantage across countries. We contribute to this literature in two ways. First, we provide a new set of measurements of migrants' productivity heterogeneity and how it relates to their origin. Second, different from these papers, we also embed this mechanism in a quantitative model to measure how the heterogeneity of migrants reshapes regional and country-wide trade and to compare this effect with others classical forces driving comparative advantage.

The rest of the paper is organized as follows. Section 2 provides a brief account of the March to the West in Brazil. Section 3 describes our data and documents the facts that form the empirical basis of the paper. Section 4 introduces our model of trade and the spatial reallocation of knowledge due to migration. Section 5 and 6 uses the model to quantify the strength of our mechanism. Section 7 concludes the paper.

2 The March to the West

The West of Brazil is nowadays one of the World's major agricultural powerhouses. If the region were a country, it would be among the 15th largest agricultural exporters in the world.[5] This status, however, came rather recently. Despite the fact that the region accounts for 60% of Brazil's territory, in 1950, less than 5% of Brazil's value added in agriculture came from the West and approximately 95% of Brazil's population lived in the East. The geographic concentration of the economy in the East reflected the historical development of the Brazilian economy: with the exception of the gold extraction in the interior of Brazil during the 18th century and the exploitation of rubber in the Amazon forest in the late 19th century, the Brazilian economy was largely based on export-oriented crops such as sugarcane, coffee and cotton that required access to ports located along the Atlantic coast in the East of Brazil.

The Onset of the March

The rise of the West began in the 1950s, when urbanization and demographic transition took off in Brazil. Concerned with food security and the population pressure in urban centers of the southeast, the president at the time, Getulio Vargas, initiated a large-scale project to

[5] The Brazilian States are officially divided in five broad regions based on socio-economic and geographic features: Central-West, North, Northeast, Southeast and South. Our analysis focuses on the occupation of the Central-West and the North, shown in Figure 1 inside the red contour. For simplicity, we label "West" the broader region comprising the Central-West and the North and "East" the rest of Brazil.

promote the migration of families to the Central-West. He named the project "March to the West" and, as stated by the government propaganda, the goal was to construct a nation that was free from the "vices of the coast". Part of the project consisted in changing Brazilian's perceptions about the West via propaganda.[6] Another part of the project consisted in the creation expeditions to the West, investments and improvements of railroads and the creation of institutions to establish new agricultural colonies in the West.[7]

While Getulio Vargas' government set the initial stage of the occupation of the Brazilian west, it was not until the 1960s, when successive governments undertook larger investments in infrastructure to occupy the interior of Brazil, that the migration to the West consolidated as a large scale phenomenon.

The Progress of the March

After Getulio Vargas' initial steps towards the occupation of the West, the next major step occurred during the 1960s, when the president Juscelino Kubitschek moved the Brazilian capital from Rio de Janeiro, in the coast of Brazil, to a newly constructed city, Brasília, in the Central-West region of the country. The new capital was inaugurated in 1964 and a series of new roads were constructed to connect Brasília to peripheral regions of Brazil. These investment had a large effect on the West's economy. Besides the spatial relocation of the capital, which led to the creation of a metropolitan area that currently accounts for 15% of the population in the West, many of the key roads connecting the region to ports and cities along the coast were constructed during this period.[8]

Between the 1960s and the 1980s, the military dictatorship expanded the projects from Juscelino Kubitschek to further integrate the Amazon region in the North of Brazil. The military government invested in new roads under a new national transportation plan (*Plano Nacional de Viação*), granted land to agricultural colonization companies and created a free economic zone in Manaus (a city located along the Amazon river). In addition, during this period the Brazilian agricultural research institute, Embrapa, expanded its research on the adaptation of crops to regions closer to the tropics, with soybeans being one of the flagship

[6] Figure 7 in the appendix, for example, shows a poster of the government's propaganda at the time, with a quotation claiming that "the true sense of Brazilianness is the March to the West".

[7] More specifically, in 1941, Getulio Vargas launched the first expedition to the west called "Roncador-Xingu". The goal of this expedition was to discover unpopulated regions in the interior of Brazil that were amenable to the construction of new cities. In 1943, the government created an institution called Fundação Brasil Central that had as its goal the colonization of the Western and Central region of Brazil. This initiative inaugurated a series of cities in the region such as Aragarças and Nova Xarantina. This institution also took charge of the railroad of Tocantins, which crossed the center of Brazil in the region close to Brasilia.

[8] See Morten and Oliveira (2016) and Bird and Straub (2020) for a detailed description of the roads that were constructed during this period to connect Brasília to peripheral regions.

cases of adaptation of crops to tropical areas.[9] The March to the West lost momentum in the mid-1980s, when Brazil entered a decade long period of economic depression and hyperinflation that interrupted the cycle of large-scale investments in infrastructure in Brazil that started in the 1960s.

In summary, different factors promoted the migration of workers to the West of Brazil: changes in migration costs associated with the construction of roads and the federal government's propaganda about the West,[10] reductions in the price of land associated with land grants, and productivity shocks associated with the adaptation of new crops. The structure of the spatial economy model that we formulate later allows for the influence of these different shocks on the evolution of the economy. To close this section, we briefly discuss the relationship between migration and the expansion of specific economic activities in the West.

Crop Diffusion during the March

Research in economic history has underscored the contribution of migrants' knowledge from their origin region to the expansion of new economic activities in their destination region in several historical episodes (Olmstead, Rhode, et al., 2008; Scoville, 1951; Watson, 1983). This contribution has been discussed in the particular case of the Brazilian march to the west. The following passage from Sabel, Fernandez-Arias, Hausmann, Rodriguez-Clare, and Stein (2012), p.181, for example, highlights the role of knowledge of migrants from the south of Brazil - called *gauchos*-, in the expansion of soybeans in the West:

> *The first movers had some experience with these crops in the southern part of Brazil, a region with a favorable climate and adequate conditions for soybean agriculture[...] Such experience and technical capabilities allowed them to experiment with soybean cultivation in other regions of the country at a time when international markets started to demand higher volumes of soybeans.*

As a first inspection of this link between a migrants' activity choice in their destination region and the dominant activities in their region of origin, we carry the following accounting exercise. For each crop being produced in the West in 2010, we compute the share of migrants from the East, and the share of migrants coming from low producing regions in the East of that same crop (as defined by the regions in the bottom quartile in the employment of workers). First, we find that the share of migrants from the East differs substantially across crops. For example, more than 10% of soybean or sugarcane producers in the East are

[9] See Pellegrina (2019) for an analysis of the expansion of soybeans to tropical areas during the 1970s using quantitative economic geography models.

[10] The approach that we take here for migration costs follows the classical formulation in Sjaastad (1962), who considers both the material and the psychological costs of moving to a new region.

migrants from the West, but in the case of livestock, only 3% of producers from the East come from the West (see Figure 9 in the appendix). Second, we find that migrants in the West producing a given agricultural commodity rarely come from regions within the East that do not produce that same commodity. Of course, if we were to remove the migrants from the East producing in the West, general equilibrium effects would change the incentives for non-migrants to produce that same crops. The quantitatvie model that we develop later accounts for these general equilibrium effects.

3 Data and Motivating Facts

This section describes the data and four facts about comparative advantage, specialization and migration in Brazil. The first two facts describe aggregate changes in Brazil's revealed comparative advantage and spatial allocation of workers. The following two facts examine how farmers' earnings and choices relate to their region of origin and motivate the specific structure that we adopt in our model.

3.1 Data

We collected and digitized data from various sources to construct a panel data with information about employment, migration, gross output and domestic and international trade for Brazil between 1950 and 2010. Our final data contains 133 meso-regions,[11] 26 States, two countries (Brazil and the rest of the world), and 14 economic activities (12 agricultural activities, manufacturing and services).[12] Next, we provide a brief overview of our data, leaving a thorough description to the appendix.

First, migration and employment variables come from decadal editions of the Brazilian demographic censuses from 1950 until 2010. For 1970 onwards, we have micro-data at the workers' level. Before the 1970s, we only have state-level aggregates. Our micro-level data is unique in the sense that we observe, along with migration variables, a worker's economic activity and earnings (as measured by their income). We interchangeably refer to agricultural workers as farmers.

Second, gross output in agriculture comes from PAM (*Produção Agrícola Municipal*), which gives meso-region level data by crop since 1974 and state level aggregates since the

[11] Meso-region is a formal political boundary defined by the Brazilian statistica bureau, IBGE, that combines a few municipalities according to similarities in their economic activity and labor markets. The original data contains 137 meso-regions and 27 states but it is not balanced over time. We therefore merged new regions and the state associated with Brasília to make boundaries that are consistent over time.

[12] The 12 agricultural activities are: banana, chicken, cacao, coffee, cotton, corn, livestock, rice, soy, sugarcane, tobacco and a residual agricultural activity which we call "rest of agriculture".

1930s.[13] For non-agricultural activities, we construct gross output based on value added data from IPEA, the Institute of Applied Economic Research.[14] We convert all nominal values from Brazilian datasets according to dollar values and match our aggregates to value added in dollars by sector given by UN National Accounts. For the rest of the world, we construct revenues and value added using data from FAO-STAT and UN National Accounts.

Third, our data on international trade flows comes from FAO. For domestic trade flows between Brazilian states, we compiled information from the Brazilian statistical yearbooks for the 1950s and 1970s.[15] For recent years, we use data from Vasconcelos (2001).

Fourth, we gather data on land use and total labor employment for Brazil from the Brazilian economic research institute (IPEA)[16] and for the rest of the world from FAO. For Brazil, we bring data on land use disaggregated by meso-region also from IPEA.

Lastly, we construct a matrix of nodes and arc as in Sotelo (2020) and Donaldson (2015) using information about the Brazilian road network to calculate the minimum travel distance between meso-regions.

3.2 Facts about Migration and Comparative Advantage

Fact 1: Since the 1950s, Brazil has gained comparative advantages in crops exported by the West.

In what follows, we use a common measure of revealed comparative advantage (hereafter, RCA)

$$RCA_k = \frac{X_k^{BR} \Big/ \sum_{k' \in \mathcal{K}} X_{k'}^{BR}}{X_k^{W} \Big/ \sum_{k' \in \mathcal{K}} X_{k'}^{W}}, \tag{1}$$

where k is an index for the goods from a sector, $\mathcal{K}$ the set of sectors and contains all crops and a non-agricultural activity, X_k^{BR} the exports of Brazil and X_k^W the global exports. The RCA measures the specialization of Brazilian exports in crop k, relative to the world's specialization in the same crop. A number above one suggests Brazil has a comparative advantage in crop k.[17]

[13] For livestock and chicken, we complement our data with gross-output data from agricultural censuses.

[14] *Instituto Brasileiro de Geografia e Estatistica.*

[15] The federal government measured trade flows between states and reported these flows in the statistical yearbooks before the 1980s. For recent years we do not have this information in statistical yearbooks.

[16] *Instituto de Pesquisa Econômica Aplicada.*

[17] Since Balassa (1965), the revealed comparative advantages has been used extensively in the literature to examine the patterns of specialization. For example, Goldberg, Khandelwal, Pavcnik, and Topalova (2010) and Menezes-Filho and Muendler (2011). See French (2017) and Costinot, Donaldson, and Komunjer (2012) for a detailed discussion of different measures of revealed comparative advantage and how they map into trade models.

Table 1 shows that, in 2010, Brazil exports 5-10 times more coffee, livestock and tobacco than the rest of the world, and 15-20 times more soybeans and sugarcane. Brazil has changed substantially its comparative advantage relative to the world. There was a large expansion in the RCA of cassava, chicken, livestock, soy, sugarcane and tobacco. For recent years, we can disaggregate the RCA by region. The West has substantially larger RCA in soy and cattle, which are key agricultural goods in the export basket of Brazil. These agricultural goods account for roughly 40% of Brazil's agricultural exports and 16% of its total exports.

Fact 2: Since the 1950s, the fraction of Brazil's agricultural land and labor employed in the West has increased substantially.

The 1950s marked an inflection in the evolution of the spatial allocation of economic activity in Brazil. Figure 2 shows that, from the 1870s and until 1950s, the share of Brazil's population living in the west gravitated around 7 percent. After then, this share rose by 7 percentage points between 1950 and 2010. The figure also shows that the share of agricultural land employed in the West increased by 12 percentage points, and the share of value added in agriculture generated in the West increased by 20 percentage points.

The composition of migrants among agricultural workers in the West mirrors the region's increasing population. In 2010, the overall share of migrants in the west was 15 percent, with some regions having more than 40 percent of migrants. Importantly, migrants came from all parts of the East. As shown in Figure 10 in the appendix, ten percent of migrants in the West came from Bahia, a state in the northeast that is a large producer of cacao, whereas 13 percent of them came from Rio Grande do Sul in the farther south of Brazil in the frontier with Argentina, a region that specializes in the production of soybeans and cattle.

Fact 3: Upon migration, farmers originating in regions with high crop employment are more likely to grow that same crop and earn higher incomes than other farmers doing so.

To investigate how a migrant farmer's origin region relates to her farming choices and her earnings, we estimate

$$\log\left(workers_{ij,kt}\right) = \iota^{W}_{j,kt} + \iota^{W}_{ij,t} + \alpha^{W}\log(workers_{i,kt-1}) + \epsilon^{W}_{ij,kt}, \tag{2}$$

$$\log\left(earnings_{ij,kt}\right) = \iota^{E}_{j,kt} + \iota^{E}_{ij,t} + \alpha^{E}\log(workers_{i,kt-1}) + \epsilon^{E}_{ij,kt}, \tag{3}$$

where i indexes the origin region, j the destination region, k the agricultural activity and t the year. The parameters α^W and α^E capture the elasticity of the number of workers ($workers_{ij,kt}$) or earnings ($earnings_{ij,kt}$) with respect to the number of workers in the origin

i producing k in period $t-1$ ($workers_{i,kt-1}$). To gain precision, we stack data for the years $t = 2000$ and $t = 2010$. We use the number of farmers lagged by thirty years as our measure of $workers_{i,kt-1}$. $\epsilon^W_{ij,kt}$ and $\epsilon^E_{ij,kt}$ are the error terms. To avoid the inclusion of the same farmer in both sides of the equation, we exclude non-migrants from our sample.[18]

In equations (2) and (3), α^W and α^E capture the link between farmers' comparative advantage and the size of the workforce in the origin region. To control for correlations between $workers_{i,kt-1}$ and factors that are destination and crop specific, we include destination and crop fixed effects ($\iota^W_{j,kt}$ and $\iota^E_{j,kt}$) in our specifications. This terms controls for any factor that is common across workers in destination j and activity k, including natural advantages, local institutions or price shocks. In addition, to control for correlations between $workers_{i,kt-1}$ and factors that are origin but not activity specific, we include origin and destination fixed effects ($\iota^W_{ij,t}$ and $\iota^E_{ij,t}$) in our specifications. This term controls for any factor that is common among farmers from region i producing in j, such as human capital and bilateral migration costs.

Panel A in Table 2 shows estimates of equation (2). Column 1 shows that an increase in the number of agricultural workers in the region of origin in a given activity of 1% increases the number of agricultural workers in the destination in this same activity by 0.06%. This elasticity drops to 0.053% when we include destination-origin fixed effects. To address a common problem of attenuation bias related to statistical zeros in the estimation of gravity type of regressions (Silva and Tenreyro, 2006),[19] we drop from our sample observations in the bottom quartile in the distribution of $workers_{ij,kt}$, and estimate the equations via pseudo-poisson maximum likelihood (PPML), and find larger point estimates. Lastly, we include zeros in the dependent variable in our PPML estimation, which gives a larger coefficient of 0.16. Panel B shows that the elasticity of migrants' earnings with respect to the number of farmers in the origin is between 0.02%-0.07% across the same specifications. In addition to the specification presented in table 2, table 8 to 11 included in the appendix present a series of robustness tests. We experiment with different lags for $workers_{ij,kt-1}$, we control for observable socio-economic characteristics, we use finer levels of geographic disaggregation, and we run individual level regressions. Our results are largely robust across specifications.

There are two main causal mechanisms that can generate correlations between farmers' comparative advantage and the size of the workforce in their region of origin. Previous

[18] Figure (11) in the appendix shows non-parametric regressions of equations (2) and (3). They indicate that a log-log relationship provides a good approximation of the relationship between our dependent and explanatory variables.

[19] The Brazilian census is divided into two questionnaires: the universe and the sample one. The universe one has a restricted number of questions and is applied to every citizen. The sample questionnaire contains a more detailed set of questions, but is applied only to a sample of the population. The sample includes approximately 20% of the total population.

research suggests that one of these mechanisms is learning spillovers. de la Roca and Puga (2017), for example, show causal evidence of the impact of living in big cities on workers' earnings. A second possibility is that farmers acquire their knowledge in a third place, and that the correlation that we observe in the data reflects instead sorting patterns across regions of different workforce size. To examine this possibility, in the appendix, we run regressions where we control for the sorting patterns of workers, whether they are return migrants and whether they have lived in more than two states, and where we keep only migrants whose previous meso-region coincide with the state of birth.[20] Our estimated coefficients remain largely unchanged in these empirical exercises, suggesting that sorting is unlikely to generate our results. In the next section, we will therefore connect our model to data based on learning externalities. We noticed, however, that part of our results will only depend on the existence of a correlation between workers' productivity and their region of origin, and not on the specific causal mechanism driving this correlation.

Fact 4: Agricultural revenues increase with immigration from regions employing many farmers in that crop, after controlling for employment.

In Fact 3, we showed that migrants choices and earnings in their destination are heterogeneous and associated with the economic activities of their region of origin. We now turn to an aggregate implication of Fact 3. Specifically, we examine how the aggregate production in a region depends on farmers' composition in terms of their region of origin. To do so, we estimate

$$\log\left(y_{j,kt}\right) = \iota_{j,t} + \iota_{k,t} + \alpha^A \underbrace{\log workers_{j,kt}}_{Abundance} + \alpha^C \underbrace{\log \sum_i \omega_{ij,kt} workers_{i,kt-1}}_{Composition} + \epsilon_{j,kt}, \quad (4)$$

where $workers_{j,kt}$ is the aggregate number of workers producing activity k in destination j and $\omega_{ij,kt}$ is the share of workers in destination j producing activity k who come from origin i. We estimate equation (4) using two dependent variables: quantity and revenues. The first set of fixed effect on the right hand side captures any level effect such as the size of a region or the overall demand for agricultural goods and the second one captures any crop specific characteristic such as the land intensity. When the composition term is larger for a given destination and crop, then farmers come from origins that are more specialized in the production of this crop.

Table 4 shows, first, that a 1 percent increase in the abundance of farmers in a region is

[20] Unfortunately, we observe whether a worker lives in his meso-region of birth or not, but not which meso-region is her meso-region of birth.

associated with a 0.9% increase in revenues. Second, it shows that a 1 percent increase in the average number of farmers in the origin is associated with a 0.15% increase in revenues, even controlling for the abundance of farmers, which indicates that the composition of farmers is strongly associated with total size of the sector in a region. To test the robustness of our results, we add controls for the composition of workers in terms of their socio-economic characteristics and the share of migrants in a region. To address potential endogeneity in migration patterns, we follow the literature on migration and use predicted values from migration gravity equations where we include only the euclidean distance between regions as a predictor of migration. In the appendix, we also run the same specification using only migrants and obtain similar results. As shown in Table 4, our results are robust to these different specifications.

4 A Model of Migration and Comparative Advantage

In this section, we develop a quantitative dynamic model of trade and migration in which comparative advantage is driven by land productivity, labor supplies, and the good-specific knowledge of migrants.

4.1 Environment

Geography and Commodities.

We focus attention on a Home country, which we divide into $j = 1, \ldots, I$ regions, and a rest of the world composite, denoted by F. There are $o = 1, ..., O$ sectors, which corresponds to manufacturing, services and agriculture. Within each sector we have $k = 1, \ldots, K_o$ goods (or economic activities) and each region produces an unique variety of each good. Time is discrete, and indexed by t. Iceberg trade and migration costs deter the flux of agents and goods across space. In each time, the geography of the economy is given by a set of natural advantages, a matrix of bilateral trade costs and a matrix of bilateral migration costs: $\{A_{j,kt}, \tau_{ij,kt}, \mu_{ij,kt}\}$. We omit time indexes whenever unnecessary for our presentation.

Technology.

A continuum of agricultural workers produce in sector k, region i. An agricultural worker with knowledge s rents land and produces according to

$$q_{j,k}\left(s\right) = A_{j,k} s^{\gamma_k} l^{1-\gamma_k},$$

where γ_k measures the knowledge intensity of crop k. For non-agricultural activities, we set $\gamma_k = 1$. To substantially simplify notation in this section, we assume a simple Cobb-Douglas production function. When we take the model to the data, we assume a CES with an elasticity of substitution between knowledge and land of ρ, a shifter of land intensity of ν_k, and a constant share of value added of α_k where the price index of intermediate inputs correspond to the price index of consumers' final good.[21]

Workers.

People live two periods, young and old. An adult at time t, upon observing her knowledge, decides where to live, what sector to work on and spawns a child. To simplify matters, only adults consume, and they ignore their child's utility. Adults have constant elasticity of substitution (CES) preferences between O sectors, with elasticity of substitution σ, a CES preference between K_o goods within sectors, with elasticity of substitution σ_k, and a CES preferences between varieties of each good, with an elasticity of substitution η_k. Let L_{jt} denote the adult population at time t in j.

Adult workers maximize welfare by choosing where to live and in which sector to work at time t:

$$\max_{j,k} W_{ij,kt}\varepsilon_{i,kt},$$

where preference shocks are drawn i.i.d from $G(\varepsilon) = \exp(-\varepsilon^{-\kappa})$ and W_{ijkt} is the systematic component of welfare. This systematic component is given by

$$W_{ij,kt} = \frac{w_{j,kt}s_{i,kt}}{\mu_{ij,kt}P_{j,t}}, \tag{5}$$

where $w_{j,kt}$ is the wage per efficiency unit of labor (i.e., the return to a unit knowledge), $\mu_{ij,kt}$ represents iceberg migration costs that reduce utility directly, and $P_{j,t}$ is the CES price index of aggregate consumption in destination region j. Reflecting our empirical findings, $s_{i,kt}$ is a farmer's knowledge to produce in sector k, which depends on the region she comes from. The CES price indexes are given by $P_{j,t}^{\frac{\sigma-1}{\sigma}} = \sum_{o\in O} P_{j,ot}^{\frac{\sigma-1}{\sigma}}$, $P_{j,ot}^{\frac{\sigma_k-1}{\sigma_k}} = \sum_{k\in K_o} a_{k,t}^{\frac{1}{\sigma_k}} P_{j,kt}^{\frac{\sigma_k-1}{\sigma_k}}$ and $P_{j,kt}^{\frac{\eta_k-1}{\eta_k}} = \sum_{i\in\{I+F\}} p_{i,kt}^{\frac{\eta_k-1}{\eta_k}}$, where $a_{k,t}$ is a preference shifter.[22]

A child born in i at time $t-1$ is characterized by a vector of sector-specific productivities,

[21] Specifically, the production function that we take to the data is given by

$$q_{j,k}(s) = A_{j,k}\left(\nu_k s^{1-\rho} + (1-\nu_k)l^{1-\rho}\right)^{\frac{\alpha_k}{1-\rho}} (C_j)^{1-\alpha_k}.$$

[22] Workers can migrate within the Home country, but not between Home and Foreign.

$s_{i,kt}$, which depends on the employment structure in the region where he is born.

Knowledge Endowment.

A worker's knowledge to grow each crop depends on her origin region: $s_{i,kt} \equiv s(L_{i,kt-1})$. In particular, motivated by Fact 3 in Section 3, we assume that knowledge depends on good-specific employment in the origin region through the following functional form:[23]

$$s_{i,kt} = \bar{s}_k L_{i,kt-1}^{\beta}. \tag{6}$$

Land supply.

As discussed in Sections 2 and 3, land use adjusts during the migration process, and farmers purchased land at subsidized rates. To capture these margins parsimoniously, we introduce a government that develops farmland ($H_{j,t}$) using the following technology:

$$H_{j,t} = b_{j,t} x_{j,t}^{1/\varsigma}. \tag{7}$$

where $b_{j,t}$ is the productivity of the land technology and $x_{j,t}$ is a final output requirement. . The government prices land competitively and rebates land rents to farmers proportionately to their land use. With this formulation, the elasticity of the land supply with respect to land rent is $\zeta = 1/(1+\varsigma)$.

4.2 Equilibrium

To define the equilibrium, it is useful to write the unit cost of producing a unit of the good. Defining efficiency wages as $w_{j,k} \equiv \pi_{j,k}(s)/s$, i.e. as earnings per unit of knowledge, equilibrium unit costs are given by

$$\frac{c_{j,kt}}{A_{j,k}},$$

where we also define $c_{j,kt} \equiv \kappa_{\pi}^{-1} w_{j,k}^{1-\gamma_k} r_j^{\gamma_k}$.

As a result of utility maximization, the share of region j's expenditure in sector k goods

[23] As we discuss in the appendix, because our regressions in Fact 3 include origin-destination fixed effects, when we take the model to the data it is unclear whether productivity $s_{i,k}$ should depend on the level or the share of agricultural employment in the origin region. Section 6.3 verifies that our results do not depend substantially on this assumption, by considering an alternative formulation in which knowledge depends on labor shares, $s_{i,k} = (L_{i,kt-1}/L_{i,t-1})^{\beta}$.

produced in region i is given by:

$$\pi_{ij,kt} = \frac{\left(c_{j,kt}\tau_{ij,kt}/A_{i,kt}\right)^{1-\eta}}{\sum_{i'}\left(c_{i',kt}\tau_{i'j,kt}/A_{i',kt}\right)^{1-\eta}}.$$

Next, using the definition of the observable component of welfare (5), optimal worker sorting gives the share of workers from i choosing to work in region j and sector k, $\lambda_{ij,kt}$

$$\lambda_{ij,kt} = \frac{W^{\kappa}_{ij,kt}}{\Xi^{\kappa}_{i,t}} \tag{8}$$

where and $\Xi^{\kappa}_{i,t} \equiv \sum_j \sum_k \left[w_{j,kt}s_{ikt}/\left(\mu_{ij,kt}P_{j,t}\right)\right]^{\kappa}$. It follows that the flow of workers from i to region j, sector k is $L_{ij,kt} = \lambda_{ij,kt}L_{i,t-1}$. We define the effective units of labor migrating from i to region j, sector k as

$$E_{ij,kt} \equiv s_{i,kt}\lambda_{ij,kt}L_{i,t-1}. \tag{9}$$

Finally, reflecting the land grants program in Brazil, we assume that workers become landowners when they move to a region.

To close the model, we note that total expenditure in region j reflects payments to factors there

$$X_{j,t} = \sum_k w_{j,kt}E_{j,kt} + r_{j,t}H_{j,t},$$

and sectoral expenditure, $X_{j,kt}$ reflects the preferences described above.[24]

We are now ready to define an equilibrium for this economy. We break down the equilibrium in two parts, as in Caliendo, Dvorkin, and Parro (2015): a goods market equilibrium, which takes migration flows as given, and then the migration equilibrium. A competitive equilibrium is a sequence of allocations that satisfies both the goods market and the migration equilibrium. Finally, we also provide a definition of a steady state equilibrium.

Goods market equilibrium in period t.

Given the geography at time t, migration flows $\{L_{ij,kt}\}_{ijk}$, and past labor allocations, $\{L_{i,kt-1}\}_{ik}$, at time t, a goods market equilibrium is a set of factor prices and allocations of efficiency units of labor $\{r_{j,t}\left(\{L_{ij,kt}\}\right), w_{j,kt}\left(\{L_{ij,kt}\}\right), E_{j,kt}\left(\{L_{ij,kt}\}\right)\}_{jk}$ such that:

[24] In taking the model to the data, we allow for trade imbalances, which we omit here to simplify the exposition.

1. The market for efficiency units of labor clears in region j and sector k:

$$w_{j,kt}E_{j,kt} = (1-\gamma_k)\sum_j \pi_{ij,kt}X_{j,kt}.$$

2. Land markets clear in region j:

$$r_{j,t}H_{j,t} = \sum_k \gamma_k \sum_j \pi_{ij,kt}X_{j,kt},$$

3. Total immigration into region j, sector k determine the effective supply of labor there:

$$E_{j,kt} = \sum_i s_{i,kt}\left(L_{i,kt-1}\right)L_{ij,kt},$$

   where the function $s_{i,kt}$ is defined in equation (6).

In the definition above, prices and allocations in the goods market equilibrium at time t depend on migration flows and past allocations, which introduce dynamics into the system.

Migration equilibrium in period t.

Given geography at time t and labor allocations in period $t-1$, $\{L_{i,kt-1}\}$, a migration equilibrium at time t is a set of migration flows, labor allocations and prices:$\{L_{ij,kt}$, $w_{j,kt}\left(L_{ij,kt}\right)$, $r_{j,t}\left(L_{ij,kt}\right)$, $E_{j,kt}\left(L_{ij,kt}\right)\}$ for regions i and j, and sector k, such that migration flows evolve according to optimal sorting of workers for each pair of regions i and j, and each sector k:

$$L_{ij,kt} = \lambda_{ij,kt}L_{i,t-1}, \tag{10}$$

where $\lambda_{ij,kt}$ is given by equation 8.

Finally we define a competitive equilibrium for this dynamic model.

Competitive equilibrium.

Given a geography for $t = 1,\ldots,\infty$, and initial labor allocations in period 0, $\{L_{i,k0}\}_{i,k}$, a competitive equilibrium is a sequence of migration flows, efficient labor allocations, and prices, $\{L_{ijkt}$,$E_{i,kt}$, $w_{i,kt}$, $r_{i,t}$, $\}_{t=1}^{\infty}$ that satisfy the goods market and migration equilibria in each period t.

To study the long-run behavior of this economy, we define a steady state equilibrium next.

Steady State Equilibrium.

Given a constant geography for $t = 1, \ldots, \infty$, a steady state equilibrium is a competitive equilibrium in which migration flows, labor allocations, and prices, are unchanged: $L_{ijkt} = \bar{L}_{ijk}$, $w_{i,kt} = \bar{w}_{i,k}$, $r_{i,t} = \bar{r}_i$, and $E_{i,kt} = \bar{E}_{i,k}$, $\forall t = 1, \ldots, \infty$.

Discussion of the equilibrium.

The properties of the equilibrium are shaped by the interaction of agglomeration and dispersion forces. First, the idiosyncratic draws are a force towards populating all region-crop cells. The strength of this force is governed by the dispersion in preference shocks κ: as κ decreases, individuals have stronger idiosyncratic tastes for working in different regions and activities. Second, the external sector has a downward sloping demand for the goods in Brazil; this acts as a force against full agglomeration in a given crop, within regions. The strength of this force is governed by η: as η grows, terms of trade turn against Brazil faster as output in a given crop increases. Third, our assumptions on technology yield high marginal values of labor when $L_{i,kt} = 0$, which provides an incentive for workers to be employed in each region-crop combination.

The opposing, agglomeration force is given by the spatial allocation of knowledge: if there is a large number of workers populating a region-crop cell, workers want to locate there because their productivity is larger. The strength of the agglomeration force is governed by β. Note that this force only operates in steady state, since in each period past allocations are taken as given. In other words, at any given time, conditional on past labor allocations, ours is a standard model of migration and trade in which there are no agglomeration forces. Relatedly, there is a dynamic externality in the way we model knowledge diffusion, since workers do not internalize their impact on the productivity of the next generations.

4.3 How Migration Costs Shape Comparative Advantage

To guide our quantitative assessment of the impact of migration on comparative advantage, in this section we define a suitable measure of comparative advantage, and use it to show analytically that migration as an ambiguous effect on comparative advantage.

4.3.1 A theoretically consistent measure of comparative advantage.

Through the rest of the paper, we measure comparative advantage using a measure we term the "direction of trade." We say that region i has a comparative advantage in k, relative to

i' and k', if

$$\frac{X_{ij,k}}{X_{ij,k'}} > \frac{X_{i'j,k}}{X_{i'j,k'}}. \tag{11}$$

Besides being easily measurable, this definition has two attractive features relative to Balassa's RCA, which we used for illustration earlier. First, by comparing their relative market penetration in region j', this measure focuses only on the contribution of supply-driven differences in exports —precisely the ones that our paper studies. Thus, the conditions under which 11 holds coincide with Haberler's classic definition based on relative costs. Second, as in standard definitions of comparative advantage, it is defined for activity pairs and region pairs. Since we seek to understand how regions within Brazil interact with the rest of the world, we study 11 for $i' = j = F$. See French (2017) and Costinot, Donaldson, and Komunjer (2012) for further discussion on the virtues of this measure and the pitfalls of measuring comparative advantage using Balassa's RCA.

4.3.2 Analytical Characterization of the Effects of Migration on Comparative Advantage

The general equilibrium interactions in our model are, unfortunately, too complex to admit a simple characterization of 11. Therefore, in this section only, we derive analytical results using a pared down version of our model. We let labor to be the only productive factor ($\alpha_k, \gamma_k = 0$) and Armington elasticities to be the same, $\eta_k = \eta$.[25] In addition, we assume that workers are born of a type k and can only choose where to live. Echoing earlier work on the determinants of comparative advantage we do the following exercise. We study 11 under free trade ($\tau = 1$), under perfect migration and under prohibitive migration costs.

A common theme in our results is that low migration costs reallocate workers and their knowledge acrossregions. Migration thus introduces a tension: Whether it strengthens natural differences in productivity depends on the allocation of workers in the previous period, $L_{i,t-1}$.

Migration costs and regional comparative advantage.

Suppose migration costs are prohibitive, so $\mu_{ijk} \to \infty$ for $i \neq j$. Then region i in Home has a comparative advantage in k relative to Foreign (and to k') if

$$\left(\frac{s_{i,k}}{s_{i,k'}}\right)^{\xi} \left(\frac{L^0_{i,k} A_{i,k}}{L^0_{i,k'} A_{i,k'}}\right)^{\xi} > \left(\frac{s_{F,k} L^0_{i,k} A_{F,k}}{s_{F,k'} L^0_{i,k'} A_{F,k'}}\right)^{\xi}. \tag{12}$$

[25]Allowing for a Heckscher-Ohlin force precludes an analytic characterization such as the one we present. But in Section 4 we discuss how it operates and measure its quantitative mpact.

If instead there is free mobility, so $\mu_{ijk} = 1$ for any i and j, then region i in Home has a comparative advantage in k relative to Foreign (and to k') if

$$\left(\frac{\Upsilon_{H,k}}{\Upsilon_{H,k'}}\right)^{\xi}\left(\frac{\mathcal{L}_{i,k}A_{i,k}}{\mathcal{L}_{i,k'}A_{i,k'}}\right)^{\xi} > \left(\frac{s_{F,k}L^{0}_{F,k}A_{F,k}}{s_{F,k'}L^{0}_{F,k'}A_{F,k'}}\right)^{\xi}, \tag{13}$$

where $\xi = (\eta - 1)/\eta$, $\Upsilon_{H,k} \equiv \sum_j s_{j,k}L^0_{j,k}$, $L^0_{j,k}$ is population in j of type k in the previous period and $\mathcal{L}_{i,k} \equiv A_{i,k}^{\kappa\frac{\eta-1}{\eta+\kappa}} / \sum_j A_{j,k}^{\kappa\frac{\eta-1}{\eta+\kappa}}$.

Comparing expressions (12) and (13), we obtain two results. First, migration spreads workers' knowledge. When migration costs are prohibitive, local knowledge acts as a productivity shifter, just like natural advantage, reflecting the fact that the only knowledge available is that of workers born in that location. With free mobility, knowledge from all the country is available to each region, as reflected by the term $\Upsilon_{H,k}$, which measures the aggregate effective supply of workers at Home, for activity k. Since migration erases any exogenous correlation between local worker productivity and local natural productivity, its impact on regional comparative advantage is ambiguous: it will strengthen regional comparative advantage if high knowledge workers were born in low productivity locations, but it will weaken it if the opposite is true. Second, migration magnifies the role of natural advantage through $\mathcal{L}_{i,k}$, which gives the share of workers of type k absorbed by region i. With free mobility, regions with higher productivity absorb a larger mass of raw labor relative to regions with lower productivity.

Migration costs and aggregate comparative advantage.

An advantage of expression 11 is that it provides a well defined notion of aggregate comparative advantage and marginal costs for a country composed of many regions. With prohibitive migration costs, saying that H has a comparative advantage in activity k, relative to F and relative to k', is equivalent to

$$\frac{\sum_i \left(s_{i,k}L^0_{i,k}A_{i,k}\right)^{\xi}}{\sum_i \left(s_{i,k'}L^0_{i,k'}A_{i,k'}\right)^{\xi}} > \frac{\left(s_{F,k}L^0_{F,k}A_{F,k}\right)^{\xi}}{\left(s_{F,k'}L^0_{F,k}A_{F,k'}\right)^{\xi}}. \tag{14}$$

With free mobility, we obtain

$$\left(\frac{\Upsilon_{H,k}}{\Upsilon_{H,k'}}\right)^{\xi}\frac{\sum_i \left(\mathcal{L}_{i,k}A_{i,k}\right)^{\xi}}{\sum_i \left(\mathcal{L}_{i,k'}A_{i,k'}\right)^{\xi}} > \frac{\left(s_{F,k}L^0_{F,k}A_{F,k}\right)^{\xi}}{\left(s_{F,k'}L^0_{F,k}A_{F,k'}\right)^{\xi}}. \tag{15}$$

Expression (15) shows that, with free mobility, aggregate productivity, measured by

$\Upsilon_{H,k}/\Upsilon_{H,k'}$, also drives Home's aggregate comparative advantage. This is an aggregation result that follows from the fact that every region within Home has equal access to knowledge. In contrast, with prohibitive migration costs, aggregate marginal costs depend on the correlation between regional populations, $L^0_{i,k}$, and local worker and natural productivity, captured by $s_{i,k}A_{i,k}$. Migration again has an ambiguous impact on comparative advantage, now in the aggregate. If, on average, labor was exogenously allocated to regions where labor productivity is low, migration will undo that correlation, and amplify relative productivity differences.

The results of this section provide a sharp characterization of the role of migration costs in shaping comparative advantage. In applications however, changes in migration costs are finite and policies cannot reduce them to zero. Sectors have different trade elasticities, while land and intermediate-input intensities are additional drivers of comparative advantage. In the following two sections, we take our model to the data and evaluate the impact of observed changes in migration costs in a calibrated model that encompasses all of these margins.

5 Taking the Model to the Data

This section describes the quantification of the exogenous parameters of the model. We map the model's goods-market equilibrium to the years of 1950, 1980 and 2010. Thus, we effectively set a time period to 30 years. We choose these periods to strike a balance between the quality and availability of data, and a time period early enough that we can observe the transformation of the Brazilian economy due to the March to the West.

We begin by mapping our model to the elasticities presented in Fact 3, which provides direct measures of the parameters controlling farmers' knowledge heterogeneity and incentives to migrate. With these elasticities in hand, we discuss the rest of the calibration, including the data we match and the parameters we choose.

5.1 Measuring κ and β using Reduced Form Evidence

Using the elasticities estimated in Fact 3, we uncover two key parameters from our model: the preference dispersion parameter, κ and the elasticity of knowledge with respect to the size of the workforce in the region of birth β (which is new to our theory). First, using equation (6), and letting $\mu_{ij,kt} = \mu_{ij,t}\epsilon^{\mu}_{ij,kt}$, and $\bar{s}_k = \tilde{s}_k\tilde{s}_{ij,t}\epsilon^{s}_{ij,kt}$, our model relates the earnings of migrants from i into j in activity k to the employment in that activity at the migrants' origin:

$$\begin{aligned}\log\left(earnings_{ij,kt}\right) &= \log\left(w_{j,kt}s_{i,kt}\right) \\ &= \iota_{j,kt} + \iota_{ij,t} + \beta \log L_{i,kt-1} + u_{ij,kt}^{earnings},\end{aligned} \tag{16}$$

where we define $\iota_{j,kt} \equiv \log w_{j,kt} + \log \tilde{s}_k$, $\iota_{ij,t} \equiv \tilde{s}_{ij,t}$ and $u_{ij,kt}^{earnings} \equiv \log \epsilon_{ij,kt}^{s}$. By setting $L_{i,kt-1}$ to $workers_{i,kt-1}$, the expression above is equivalent to equation 3 in Fact 3, giving a structural interpretation to each of its terms. In particular, our estimates of α^E in equation 3 identify directly β in our model.

Second, we examine how employment shares within destination-crop relate to migrant origin. In our model these employment shares are linked to migration flows, $L_{ij,kt}$. Substituting equation (6) and (8) into (10), we obtain our econometric specification:

$$\log L_{ij,kt} = \iota_{j,kt} + \iota_{ij,t} + \kappa\beta \log L_{i,kt-1} + u_{ij,kt}^{migration}, \tag{17}$$

where $\iota_{ij,t} = \log L_{i,t-1} - \kappa\left(\log \mu_{ij,t} + \log \tilde{s}_{ij,t} - \log \Xi_{it}\right)$, $\iota_{j,kt} = \kappa\left(\log w_{j,kt} + \log \tilde{s}_k\right) - \log P_{j,t}$, and $u_{ij,kt}^{migration} = \kappa \log \epsilon_{ij,kt}^{\mu} - \kappa \log \epsilon_{ij,kt}^{s}$. As in the earnings equations, by setting $L_{i,kt-1}$ to $workers_{i,kt-1}$, the expression above becomes equivalent to equation 2. Here, our estimate of α^W captures the combination of two structural parameters, κ and β. The identification of β in the earnings equation 3 allows us to disentangle κ from β in the workers equation 2. Specifically, the ratio of α^W/α^E identifies κ.

From the elasticities presented in Table 2, our values for β range between 0.02 and 0.075 and for κ between 1.1 and 3.5. Since β is new to our theory, there is no direct benchmark in the literature. The closest parallel that we could trace is to estimates of the effect of city size on productivity. de la Roca and Puga (2017), for example, found a medium run elasticity of earnings with respect to city size of 0.05, which is within the range of values obtained in our specification for β. We can however compare κ with several recent papers to check our estimates. Using meso-region data from Brazil, Morten and Oliveira (2016) estimate an elasticity of migration with respect to wages of 1.9. Examining more granular data on commuting within metropolitan regions in the US, Monte, Redding, and Rossi-Hansberg (2015) estimate an elasticity of 3.3. Using province level data from China, Tombe and Zhu (2019) find preference dispersions between 1.2 and 1.6. Therefore, our estimates of κ are well within the range found in the literature. In what follows, we pick $\beta = \{0.035, 0.07\}$ and $\kappa = 1.5$.

5.2 Calibrating other elasticities and recovering wedges

In addition to β and κ, we need to calibrate productivities ($A_{j,kt}$), land supply productivity ($b_{j,t}$), trade costs ($\tau_{ijk,t}$), migration costs (μ_{ij}), preferences (η_k, σ and $a_{k,t}$), and production technology (ρ and ν_k). The calibration of these exogenous parameters follows more standard procedures in economic geography models.

Productivities ($A_{j,kt}$ and $b_{j,t}$).

We calibrate $A_{j,kt}$ in the model to match gross output in the data. As discussed in Allen and Arkolakis (2014), we can only identify $A_{j,kt}$ up to a normalization, so we set $A_{F,kt} = 1$. We adjust $b_{j,t}$ to match the data on total agricultural land use.

Preferences (σ, σ_k, η_k and $a_{k,t}$)

We set σ to 1.5, σ_k to 2.5 according to Sotelo (2020), η_k in agriculture to 9.5 and η_k for other sectors to 5.5 following Caliendo and Parro (2015). We adjust preference shifters $a_{k,t}$ in the model to match the aggregate apparent consumption in each activity.

Technology (ρ, ν_k and α_k)

We set the share of value added (α_k) to 0.25 for manufacturing, 0.55 for agriculture and 0.6 for services according to the share of value added for Brazil in the World Input-Output Database (WIOD) of 2010. For non-agricultural activities, we set the cost share of land to $\nu_k = 0$. For agricultural activities, we calibrate ρ and ν_k according to the literature, but we have different approaches here. While Costinot, Donaldson, and Smith (2016) assume perfect complementarity between land and labor, which effectively sets ρ to 0, agricultural economists have traditionally adopted Cobb-Douglas functions, which gives $\rho = 1$. We therefore set $\rho = 0.5$, which is in between these two approaches. We set ν_k according to the cost share of land in Pellegrina (2019), and study how our results change if we used a common $\nu_k = 0.3$ across activities based on Mundlak (2001).

Trade Costs ($\tau_{ijk,t}$).

We need to calibrate the whole matrix of trade costs between regions and sectors in each period, which gives $134 \times 134 \times 14 \times 3$ parameters. Since we do not have trade data between any two regions, we follow the literature and impose more structure on trade costs. We assume, for $i = j$, that $\tau_{ij,kt} = 1$ and, for $i \neq j$, that

$$\tau_{ij,kt} = \left[\delta_t \left(dist_{ij}\right)^{\delta_t^T}\right]^{\iota_{ij}^T} \left[\delta_t \delta_{ij,kt} \left(dport_i \times dport_j\right)^{\delta_t^T}\right]^{1-\iota_{ij}^T}, \tag{18}$$

where ι_{ij}^T is a dummy variable that equals one if i and j belong to the same country and zero otherwise, $dist_{ij}$ is the travel distance between i and j and $dport_i$ is the travel distance to the nearest port.

We choose τ_t to match the share of intra-regional trade in total domestic trade in Brazil as in Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016). In particular, we construct $\sum_{s \in H} X_{ss}/X_{HH}$, where X_{ss} are sales of a state to itself and X_{HH} are sales of Brazil to itself. We target a domestic trade flow of 0.7 in 1950 and 0.6 for 1980 and 2010, which gives δ_t of 0.08, 0.42 and 0.58.[26] Together with δ_t, we search for values of the distance parameter δ_t^T so that the simulated model matches the elasticity of trade flows between states with respect to distance in the data. [27]For 2010, the OLS estimate of the trade elasticity equals 1.05, for 1980 it equals 1.25, and for 2010 it equals 2.5, which gives δ_t^T of 0.22, 0.25 and 0.52 respectively.

We assume that $\delta_{ij,kt}$ is symmetric for manufacturing, asymmetric for the remaining economic activities and infinity for services. This gives 25 parameters to be estimated. We adjust $\tau_{ij,kt}$ in the model to match aggregate Brazilian exports, as well as exports and imports in each agricultural activity. We are forced to assume symmetric trade costs for one of the sectors due to identification issues discussed in Waugh (2010).

Migration Costs ($\mu_{ijk,t}$).

As with trade costs, we have a large matrix of migration costs to calibrate but limited data on migration flows. We therefore impose additional structure on this matrix. We assume, for $i = j$, that $\mu_{ij,kt} = 1$ and, for $i \neq j$, that

$$\mu_{ij,kt} = \left[\mu_t \left(dist_{ij}\right)^{\delta^M}\right]^{\iota_{ij}^M} \left[\mu_{ss',t} \mu_{ss',kt} \left(dcap_i \times dcap_j\right)^{\delta^M}\right]^{1-\iota_{ij}^M} \tag{19}$$

where ι_{ij}^M is a dummy variable that equals one if i and j belong to the same state $s$and zero otherwise, $dcap_i$ is the travel distance to the state capital. We assume that the inter-state migration cost contains a symmetric component $\mu_{ss',t} = \mu_{s's,t}$. Our parametrization of migration costs assumes that workers have to pass by the state capital to reach other

[26]For comparison, using data from the Commodity Flow Survey from the US on the manufacturing sector only, Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016) find equivalent shares between 0.35 and 0.45 depending on the level of geographic aggregation.

[27]Since the model is defined at the sectoral level, it does not deliver a log-linear gravity equations that maps directly the parameters from reduced form estimates of the gravity equation given at the aggregate level to structural parameters from the model.

states. As such, we have a hub-and-spoke type of migration cost, which is similar to the structure discussed in Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016) for trade cost between countries containing sub-regions.

Given our parametrization of migration costs, the model delivers the following gravity equation for migration

$$\log \left(L_{ss',kt}\right) = \alpha_{s,kt} + \alpha_{s',kt} + \tilde{\mu}_{ss',t} + \epsilon_{ss',kt}. \tag{20}$$

where $\alpha_{s,kt}$ and $\alpha_{s',kt}$ are destination, $\tilde{\mu}_{ss',t} \equiv -\kappa \log \left(\mu_{ss',t}\right)$ and origin fixed effects and $\epsilon_{ss',kt}$ is an error term. We estimate the equation above via OLS and recover the implicit $\mu_{ss',t}$. In the appendix, we show that our estimates of $\mu_{ss',t}$ are strongly correlated with the Head and Ries index, which is a common measure of trade cost based on gravity models, but applied to aggregate migration flows between states ($L_{ss',t} = \sum_k L_{ss',kt}$). Estimating $\mu_{ss',t}$ using equation 20 comes with two main advantages relative to using the Head and Ries index. First, it provides measures of $\mu_{ss',t}$ that are theoretically consistent with our model, which is given at the economic activity level, given our parametrization of trade costs. Second, it does not force us to assume symmetric migration costs as in the Head and Ries index. We recover the symmetric component $\mu_{ss't}$, which is common across activities k, while allowing asymmetric terms to become part of the error.

Using our estimates of $\mu_{ss',t}$, we calibrate the rest of the migration costs as follows. We adjust μ_t in the model to match the share of workers living in their meso-region of birth and $\mu_{ss',kt}$ to match the migration of workers between states and activities. Later, in the simulation of our model, we treat $\mu_{ss',kt}$ as a residual and keep it fixed across different counterfactuals. Finally, we set δ_t^M, which controls migration costs within states, to 0.3 according to values obtained in Morten and Oliveira (2016).

5.3 The March to the West as seen Through our Model

Table 4 presents selected summary statistics of our calibration. Each panel highlights a different type of exogenous variation in our model that can generate migration of workers across Brazilian regions. Specifically, our model captures exogenous variations coming from four sources: (i) migration costs, (ii) productivity and (iii) trade costs.

Panel A shows that domestic migration costs—measured by $\mu_{ij,t}$ in the model— declined sharply over this period, in line with what one would expect from the policies buttressing the March to the West. Migration costs from the East to the West declined the most.[28]

[28] The magnitudes of these migration costs, although large, are in line with those measured, for example, by Tombe and Zhu (2019) for China.

Migration costs within Wests and, particularly, the East, are much lower. Part of these trends are explained by a decline in the elasticity of migration costs with respect to distance. The level of our migration costs are consistent with the ones obtained in Tombe and Zhu (2019), who find an overall migration cost between provinces in China of 25.

Panel B shows the evolution of productivity, $A_{i,k}$, in the West relative to the East. The West's agricultural productivity caught up and surpassed that of the East, which gives rise to the mismatch between labor and productivity that we discussed in Section 4.3. For soybeans, specifically, the upward trend in relative productivity is in line with what has been documented by Bustos, Caprettini, and Ponticelli (2016). Panel B also indicates that the productivity of the land supply sector, as captured by $b_{j,t}$, increased from 1950 to 2010 in the West relative to the East, which is in line with the various policies that the federal government adopted to facilitate the acquisition of land in the West.

Panel C shows that the costs of trading internationally in Brazil have also declined, which is what one would expect given Brazil's increasing trade openness (Dix-Carneiro, 2014). Domestic trade costs, moreover, also declined and became less sensitive to distance, in line with the transportation policies that fostered East-West trade integration.

6 The Aggregate and Regional Effects of Migration

In this section, we evaluate the quantitative importance of migration as a driver of comparative advantage. Using our model, we ask how Brazil's external trade would have evolved between 1950 and 2010, had migration costs not declined as we saw in the data. Our choice of counterfactual is guided by policy. We interpret our counterfactual results as asking: What would have happened had the Brazilian government's domestic integration policies never been enacted?

6.1 Constructing a Counterfactual Economy

We simulate a counterfactual economy in which migration costs remain at their 1950 levels, while all other exogenous shocks evolve as in the data. We compare the evolution of the direction of trade and the gains from trade with those in the data. To gauge the role of different theoretical mechanisms in the model, we present our in two steps. First, we present our results assuming $\beta = 0.035$. Second, we shut down the role of knowledge by setting $\beta = 0$ and the role of Heckscher-Ohlin forces by setting $\nu_k = 0$. In each parametrization, we re-adjust all shocks in the model.

Before moving to the analysis of our counterfactual on trade, we note that the integration policies that generated the large drop in migration costs between 1950-1980 are key for the occupation of the West. Table 5 shows the evolution of the share of Brazil's population living in the West according to different scenario. It shows that changes in migration costs account for 55% of the increase in Brazil's share of population living in the West.

6.2 The Impact of Migration Costs on the Direction of Trade

Figures 3 (a) and (b) plots, on the vertical axis, changes in the direction of trade for each region i in the West relative to the rest of the world, using manufacturing as a reference. In the horizontal axis, we plot the relative share of workers born in the West in the total supply of effective labor for each region and activity. Intuitively, the smaller this share, the more exposed is the region-activity pair to a reduction of immigration from the East, compared to manufacturing.[29]

We highlight three patterns from this figure. First, both in 1980 and 2010, most region-activity pairs fall below zero, meaning that migration cost reductions shift trade towards agriculture. In region-activity pairs that are highly exposed to migration, these changes can be substantial, shifting the direction of trade by more than 50%. Second, large producers of the new commodities (soybean, livestock, and corn) are among the most affected, indicating that, within agriculture, migration cost reductions shifted the direction of trade toward these goods. Third, the relationship with migration exposure is tighter in 1980. As other shocks hit the economy between 1980 and 2010, exposure to migrants becomes less important in explaining total changes in relative exports. In sum, the quantification shows that migration costs are a key driver of regional comparative advantage in Brazil, especially as it relates to new export commodities.

We aggregate these regional changes in Table 6 to study the impact of migration costs the direction of trade for Brazil as a whole. The first three columns of the table display the evolution of the direction of trade in the data; the last two, the fraction of the changes in the direction of trade between 1950-1980 and 1950-2010 that the observed decline in migration costs accounts for. The table reveals two central messages. First, migration costs alone account for more than the observed gain in agricultural comparative advantage up to 1980, and three-quarters of it up to 2010. Second, breaking down the changes in agriculture across crops, migration costs account for a sizable fraction of the evolution of soybeans, livestock, and corn. These are the results we expect given the prominent role of the West in the exports

[29] In the case in which a region i goes to migration autarky relative to all other regions, the share of domestic workers in total labor supply is a sufficient statistic for the effects of going to migration autarky. This is analogous to the well-known result by Arkolakis, Costinot, and Rodriguez-Clare (2012) on the gains from trade. Appendix C provides a proof.

of these crops by 2010, and given the fact that the West was a receiver of net migrants in the previous six decades.

6.3 Quantifying the Margins of Comparative Advantage

We now turn to measure the contribution of each of the adjustment margins we have included in the model. To do so, we recalibrate the model under several specifications and perform the same simulations we described in Section 6.1.

Factor Proportions.

The model captures Heckscher-Ohlin forces by allowing land intensity to vary across activities, so we recompute our results after eliminating that heterogeneity. In the horizontal axis of Figure 5 (a), we plot the counterfactual reduction in the direction of trade in a calibration in which land is not an input in production, or $\nu_k = 0$. In the vertical axis we plot the additional impact implied by our baseline, with ν_k given by the data. Allowing for factor proportions forces reduces the negative effect on livestock substantially because it is particularly land intensive. Two mechanisms explain this result: the direct negative impact of the labor shortfall on marginal costs is smaller and, because the supply of land is not perfectly elastic, in equilibrium the price of land declines—which also affects land-intensive crops the most. The relative magnitude of additional impact by the reference lines in the Figure: the impact in livestock comparative advantage would have been 11 percentage points lower, or 30 percent of the baseline.

Worker Heterogeneity.

A natural approach is to examine how our results depend on β, since we parameterize worker heterogeneity as $s_{i,kt} = L^{\beta}_{i,kt-1}$. Figure 5 (b) plots, on the horizontal axis, counterfactual changes in the direction of trade in a calibration in which $\beta = 0$. The vertical axis plots the additional impact of our baseline value of $\beta = 0.035$. We can see that worker heterogeneity magnifies the impact of migration costs on new crops such as soybeans and livestock by approximately 10% of the what one would obtain with $\beta = 0$. In Figure 14 in the appendix, we repeat this analysis with our upper bound value of $\beta = 0.07$ instead and find that worker heterogeneity can account for up to 15% of migration's impact on these crops.

6.4 Gains from Trade and from Migration

An alternative way of evaluating the impact of migration on Brazil's trade is to assess how it affects the gains from trade. In our model, a measure of utility is, Ξ_i, which gives the ex-ante utility attained by a person born in region i. We average this measure using population weights to compute aggregate welfare, and we define the gains from trade as the welfare lost from going to autarky (see Costinot and Rodriguez-Clare, 2015). In this section, we return to our baseline parameterization.

A key message from the paper is that aggregate gains from trade almost do not interact with migration. Table 7 shows that the gains evolve over time, but that they are almost the same in the baseline and the counterfactual scenario of high migration costs. To understand this result, we turn to analyzing the very large heterogeneity across regions. Figure 6 (a) shows in a map how the gains from trade change in going from the baseline to the counterfactual. It reveals a clear pattern: on average, Western regions gain more from trade in the baseline, while the opposite is true for Eastern regions. Moreover, Figure 6 (b) shows that net changes in employment and changes in the gains from trade between scenarios are negatively correlated.

The key to these results are domestic trade costs and the changes in the labor force induced by migration. To see why, consider a region i in the West. Compared to the baseline, without migration region i has a small population and it is therefore relatively open to the East and to Foreign. In contrast, in the baseline, region i has larger employment, and it is therefore relatively closed to both the East and Foreign. Domestic trade costs play a dual role. First, they allow for welfare to be different across regions, enabling differences in the gains from trade. Second, with domestic trade costs, population growth increases region i's market access, which mitigates the resulting drop in wages.

A second pattern emerges from Figure 6, in which the sizes of the dots indicate the changes in agricultural comparative advantage experienced by those regions due to migration 6. The regions that depart from the regularity we have just established are those whose comparative advantage change substantially. For these regions, migration reshapes comparative advantage and, in fact, boosts the gains from trade. In some cases —e.g., Mato Grosso, the heart of the soybean revolution— the change in comparative advantage dominates, and the GFT increase with migration. In fact, regressing changes in the gains from trade on net changes in the effective employment and on changes in comparative advantage accounts for 35% of the variation in the data, about one-seventh of which is due to changes in comparative advantage.

7 Conclusions

We study how the allocation of labor within a country shapes comparative advantage at the regional and aggregate level. We show that the decline in East-West migration costs since 1950 that drove the March to the West also encouraged the development of new comparative advantage industries, as soy, livestock, and corn. Key to these new developments was the reallocation from regions in the East, where these crops were already being produced, to those in the West, were productivity grew and matched with the inflow of workers.

Our research contains lessons for research on trade and migration, by highlighting that, over long periods of time, migration can amplify natural differences across locations and shape comparative advantage, which previous research—including ours—often considers an unchanging feature of the world. This complementarity also holds policy lessons on the formation of new exporting sectors. Migrants can substantially boost the expansion of new exporting activities, specially if they bring knowledge about these activities.

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. Atkin (2016): "Volatility and the Gains from Trade," Working Paper 22276, National Bureau of Economic Research.

Allen, T., and D. Donaldson (2018): "The Geography of Path Dependence," .

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

Arkolakis, C., S. K. Lee, and M. Peters (2018): "European Immigrants and the United States' Rise to theTechnological Frontier in the 19th Century," .

Bahar, D., and H. Rapoport (2016): "Migration, knowledge diffusion and the comparative advantage of nations," *The Economic Journal*.

Balassa, B. (1965): "Trade liberalisation and revealed comparative advantage 1," *The manchester school*, 33(2), 99–123.

Bazzi, S., A. Gaduh, A. D. Rothenberg, and M. Wong (2016): "Skill Transferability, Migration, and Development: Evidence from Population Resettlement in Indonesia," *American Economic Review*, 106(9), 2658–2698.

Bird, J., and S. Straub (2020): "The Brasília experiment: The heterogeneous impact of road access on spatial development in Brazil," *World Development*, 127, 104739.

Bryan, G., and M. Morten (2015): "Economic development and the spatial allocation of labor: Evidence from Indonesia," *Manuscript, London School of Economics and Stanford University*, pp. 1671–1748.

Buera, F. J., and E. Oberfield (2016): "The Global Diffusion of Ideas," Working Paper 21844, National Bureau of Economic Research.

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

Cai, J., N. Li, and A. M. Santacreu (2019): "Knowledge Diffusion, Trade, and Innovation across Countries and Sectors," .

Caliendo, L., M. Dvorkin, and F. Parro (2015): "The impact of trade on labor market dynamics," Discussion paper, National Bureau of Economic Research.

Caliendo, L., and F. Parro (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82(1), 1–44.

Costinot, A., and D. Donaldson (2014): "How Large are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-1997," .

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79(2), 581–608.

Costinot, A., D. Donaldson, and C. Smith (2016): "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields around the World," *Journal of Political Economy*, 124, 205–248.

Costinot, A., and A. Rodriguez-Clare (2015): "Chapter 4 - Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, ed. by K. R. Elhanan Helpman, and G. Gopinath, vol. 4 of *Handbook of International Economics*, pp. 197 – 261. Elsevier.

Crosby, A. W. (1973): *The Columbian Exchange. Biological and Cultural Consequences of 1492*, no. 2 in Contributions in American Studies. Greenwood.

de la Roca, J., and D. Puga (2017): "Learning by working in big cities," *The Review of Economic Studies*, 84(1), 106–142.

Dix-Carneiro, R. (2014): "Trade liberalization and labor market dynamics," *Econometrica*, 82(3), 825–885.

Donaldson, D. (2015): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, Forthcoming.

Eaton, J., and S. Kortum (2012): "Putting Ricardo to Work," *Journal of Economic Perspectives*, 26(2), 65–90.

Fajgelbaum, P., and S. Redding (2014): "External Integration and Internal Development: Evidence from Argentina, 1870-1914," NBER Working Papers 20217, National Bureau of Economic Research, Inc.

Farrokhi, F., and H. S. Pellegrina (2019): "Global Trade and Margins of Productivity in Agriculture," *Working Paper*.

French, S. (2017): "Revealed comparative advantage: What is it good for?," *Journal of International Economics*, 106(C), 83–103.

Goldberg, P. K., A. K. Khandelwal, N. Pavcnik, and P. Topalova (2010): "Imported intermediate inputs and domestic product growth: Evidence from India," *The Quarterly journal of economics*, 125(4), 1727–1767.

Gouel, C., and D. Laborde (2018): "The Crucial Role of International Trade in Adaptation to Climate Change," Working Paper 25221, National Bureau of Economic Research.

Hanson, G. H., N. Lind, and M.-A. Muendler (2015): "The Dynamics of Comparative Advantage," Working Paper 21753, National Bureau of Economic Research.

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.

Lind, N., and N. Ramondo (2018): "Innovation, Knowledge Diffusion, and Globalization," Working Paper 25071, National Bureau of Economic Research.

Menezes-Filho, N. A., and M.-A. Muendler (2011): "Labor reallocation in response to trade reform," Discussion paper, National Bureau of Economic Research.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2015): "Commuting, migration and local employment elasticities," Discussion paper, National Bureau of Economic Research.

Morten, M., and J. Oliveira (2016): "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," *NBER Working Paper*, 22158.

Mundlak, Y. (2001): "Production and supply," *Handbook of agricultural economics*, 1, 3–85.

Olmstead, A. L., and P. W. Rhode (2011): "Adapting North American wheat production to climatic challenges, 1839–2009," *Proceedings of the National Academy of Sciences*, 108(2), 480–485.

Olmstead, A. L., P. W. Rhode, et al. (2008): “Creating Abundance,” *Cambridge Books*.

Opala, J. A. (1987): *The Gullah: rice, slavery and the Sierra Leone-American connection.* USIS.

Pellegrina, H. S. (2019): “Trade and the Spatial Organization of Agriculture: Evidence from Brazil,” *Working paper*.

Porteous, O. (2019): “High Trade Costs and Their Consequences: An Estimated Dynamic Model of African Agricultural Storage and Trade,” *American Economic Journal: Applied Economics*, 11(4), 327–66.

——— (2020): “Trade and agricultural technology adoption: Evidence from Africa,” *Journal of Development Economics*.

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

Redding, S. J. (2016): “Goods trade, factor mobility and welfare,” *Journal of International Economics*, 101(C), 148–167.

Sabel, C., E. Fernandez-Arias, R. Hausmann, A. Rodriguez-Clare, and E. Stein (2012): “Export Pioneers in Latin America,” .

Scoville, W. C. (1951): “Minority migrations and the diffusion of technology,” *The Journal of Economic History*, 11(4), 347–360.

Silva, J. S., and S. Tenreyro (2006): “The log of gravity,” *The Review of Economics and statistics*, 88(4), 641–658.

Sjaastad, L. A. (1962): “The costs and returns of human migration,” *Journal of political Economy*, 70(5, Part 2), 80–93.

Sotelo, S. (2020): “Domestic Trade Frictions and Agriculture,” *Forthcoming Journal of Political Economy*.

Sviatschi, M. M., et al. (2018): “Making a narco: Childhood exposure to illegal labor markets and criminal life paths,” *Manuscript, Department of Economics, Princeton University. https://rpds. princeton. edu/sites/rpds/files/sviatschi making-a-narco march2018. pdf*.

TOMBE, T. (2015): "The Missing Food Problem," *American Economic Journal: Macroeconomics*, 7(3), 1–33.

TOMBE, T., AND X. ZHU (2019): "Trade, migration, and productivity: A quantitative analysis of china," *American Economic Review*, 109(5), 1843–72.

VASCONCELOS, J. R. D. (2001): "Matriz do fluxo de comércio interestadual de bens e serviços no Brasil-1999," Discussion paper.

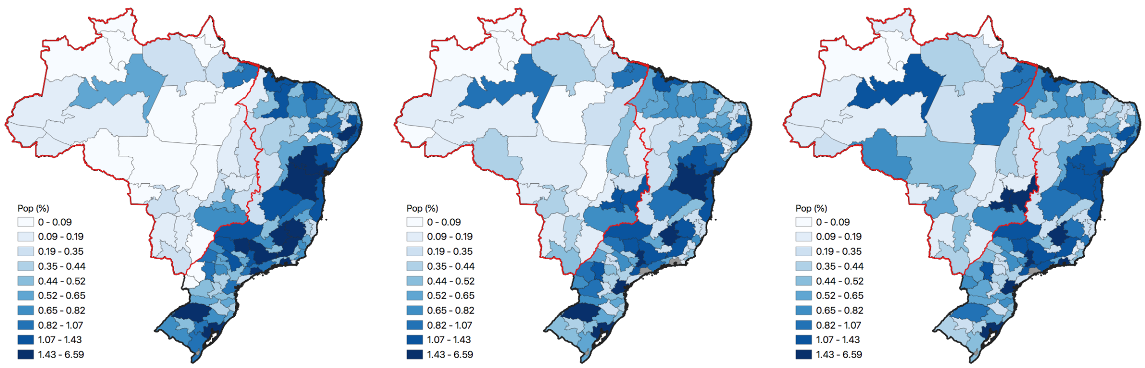
WATSON, A. M. (1983): "Agricultural innovation in the early Islamic world; the diffusion of crops and farming techniques, 700-1100," .

WAUGH, M. E. (2010): "International Trade and Income Differences," *American Economic Review*, 100(5), 2093–2124.

8 Tables and Figures

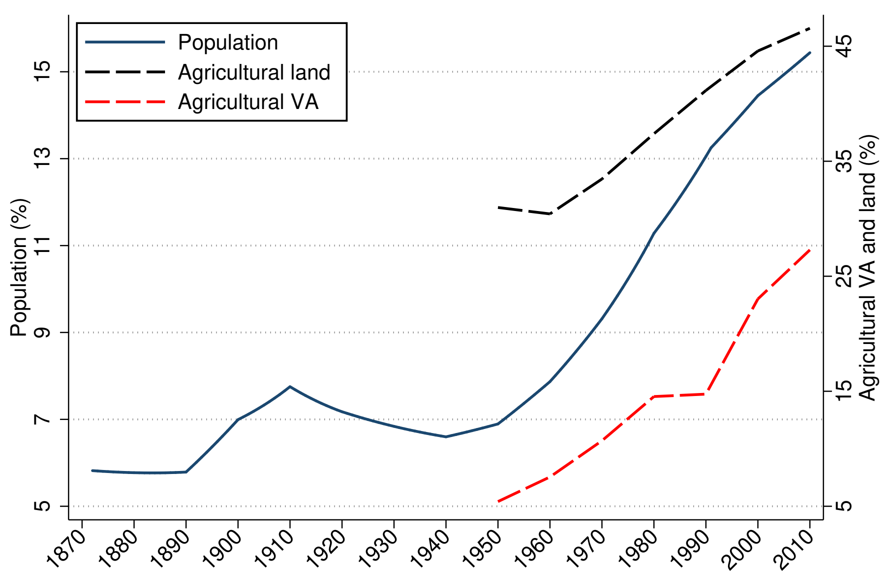
Figure 1: Spatial Distribution of the Brazilian Population between 1950 and 2010

(a) 1950 (b) 1980 (c) 2010



Notes: The red line shows the definition of the West used in our analysis. The figure shows the division of Brazil into meso-regions, which is the geographic unit used in our analysis.

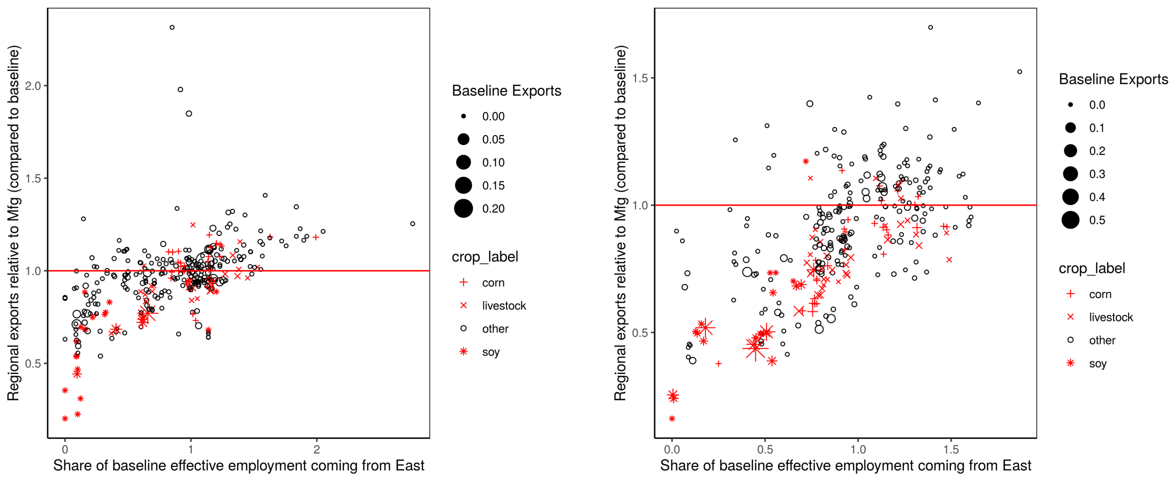
Figure 2: Migration, Agricultural Land Use and Agricultural Value Added in the West of Brazil



Notes: This figure highlights the increase in the share of the Brazilian population living in the west, the share of total land use coming from the west as well as the relative importance of the west for the total value added in agriculture.

Figure 3: Exposure to Migrants and Changes in Export in the West

(a) 1980 (b) 2010



Notes: This figure shows, for meso-regions in the West, the relationship between changes in exports in a region-activity pair relative to that region's manufacturing exports on the baseline share of effective employment in that region and activity pair from the East relative to the effective employment. The size of the circles represent the magnitude of exports.

Figure 4: Changes in Exports relative to Manufacturing by Region between 1950 and 2010 for Selected Activities

(a) Soybeans (b) Livestock

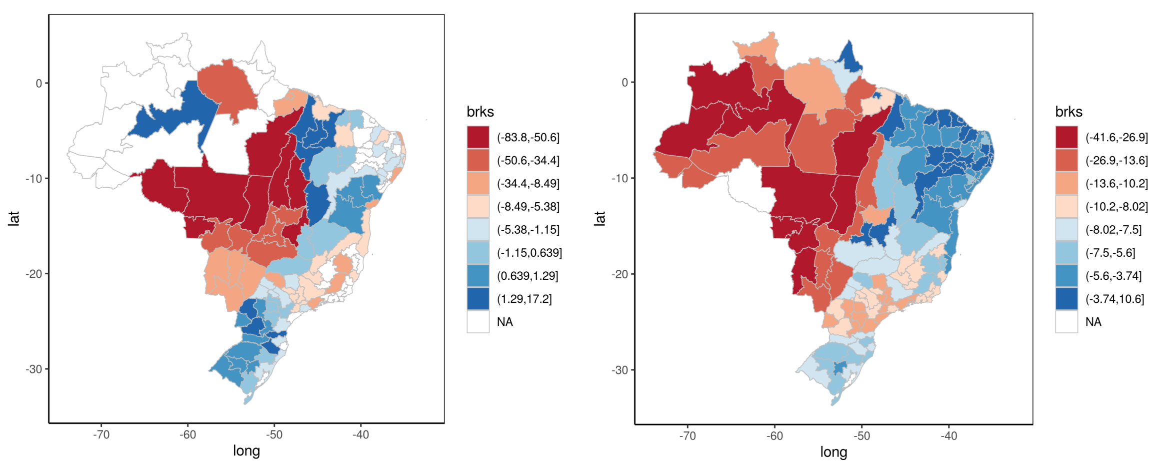
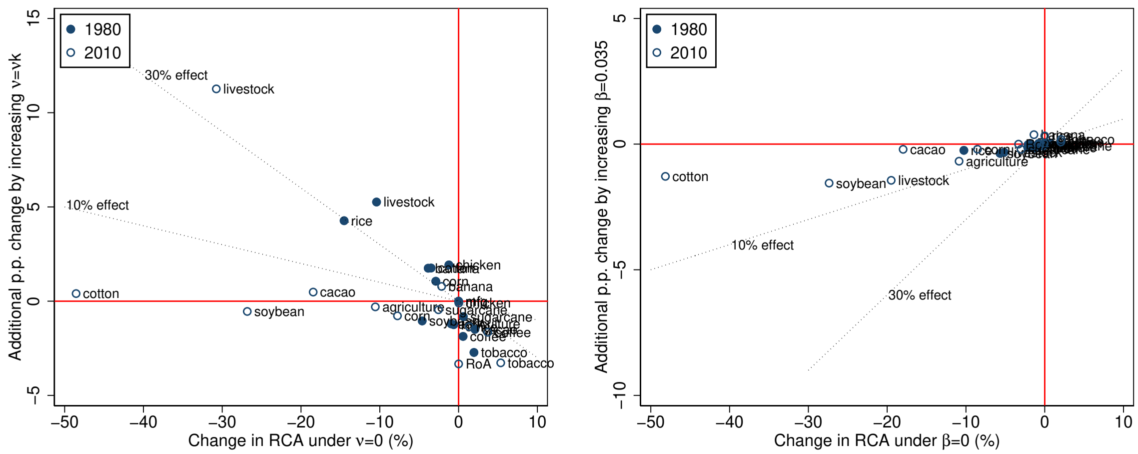


Figure 5: Examining the Effects of Knowledge and Land

(a) Heckscher-Ohlin forces (b) Knowledge reallocation channel



Notes: This figure shows the changes in Brazil's theoretically consistent RCA when we add different margins to the model. The reference dotted lines show the percent effect of 10% and 30%.

Figure 6: Effects of Migration on Gains from Trade across Regions in 2010

(a) Effects across regions (b) Drivers of Changes

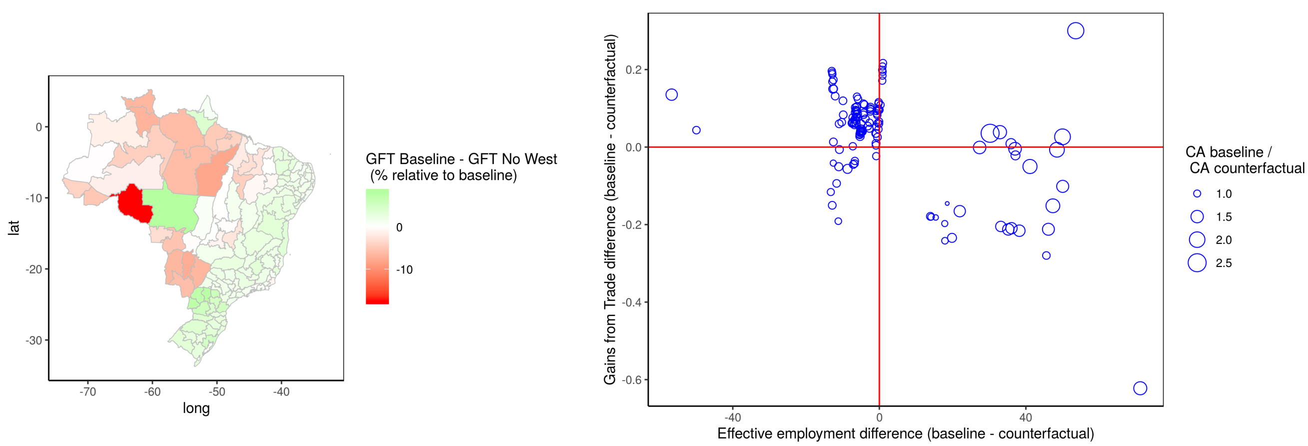


Table 1: The Evolution of Brazil's Revealed Comparative Advantage (1950-2010)

	Brazil			East	West
	1950	1980	2010	2010	2010
	(1)	(2)	(3)	(4)	(5)
mfg	0.26	0.63	0.73	0.76	0.59
agriculture	1.00	1.00	1.00	1.00	1.00
- RoA	0.78	1.55	1.11	1.15	0.87
- banana	1.90	1.29	0.36	0.43	0.00
- chicken	0.01	6.32	16.11	19.06	0.80
- cacao	16.93	14.98	1.26	1.50	0.01
- coffee	35.11	17.58	15.34	18.25	0.29
- corn	1.52	0.34	6.70	2.21	29.95
- cotton	2.93	0.50	5.72	3.91	15.13
- livestock	2.41	3.17	9.53	5.92	28.28
- rice	0.59	0.27	1.35	1.59	0.16
- soy	0.00	15.33	22.56	15.06	61.46
- sugarcane	3.32	6.08	29.03	33.05	8.17
- tobacco	1.47	4.07	6.37	7.59	0.00

Notes: This table computes Brazil's revealed comparative advantage in columns 1 to 3 using data from FAO and COMTRADE. In columns 4 and 5, we use data on exports by region coming from Comexstat.

Table 2: The Relationship between Farmers' Choices and Earnings and their Region of Origin

	OLS	OLS	OLS	PPML	PPML	PPML
	(1)	(2)	(3)	(4)	(5)	(6)
a. Farmers in destination (logs)						
Farmers in origin	0.070***	0.058***	0.082***	0.105***	0.118***	0.164***
	(0.006)	(0.012)	(0.018)	(0.012)	(0.019)	(0.010)
R^2	0.189	0.731	0.753	-	-	-
Obs	8265	8265	6316	8265	6316	156735
b: Earnings (logs)						
Farmers in origin	0.020***	0.023**	0.048***	0.030***	0.072***	-
	(0.004)	(0.009)	(0.014)	(0.011)	(0.016)	-
R^2	0.345	0.692	0.714	-	-	-
Obs	7520	7520	5757	7520	5757	
Dest-Act-Year FE	Y	Y	Y	Y	Y	Y
Dest-Orig-Year FE		Y	Y	Y	Y	Y
Above Q1			Y		Y	
Include zeros						Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clustered at the destination-crop-year level in parenthesis. The unit of observation is a given at the destination-activity-origin-year. Column 3 and 5 drop values in the bottom quartile in the distribution of the dependent variable. Explanatory variable is the log of workers in the same activity in the region of origin lagged by thirty years. Our sample exclude return and non-migrants.

Table 3: Relationship between the Composition of Farmers and Agricultural Output

	Revenues			Quantity		
Explanatory Variable	(1)	(2)	(3)	(4)	(5)	(6)
Abundance	0.907***	0.865***	0.884***	0.979***	0.904***	0.919***
	(0.037)	(0.057)	(0.056)	(0.041)	(0.066)	(0.065)
Composition	0.206***	0.199***	0.171***	0.237***	0.229***	0.220***
	(0.042)	(0.043)	(0.043)	(0.050)	(0.051)	(0.052)
R^2	0.873	0.874	0.873	0.833	0.834	0.833
Obs	1679	1679	1679	1460	1460	1460
Region-Year	Y	Y	Y	Y	Y	Y
Activity-Year	Y	Y	Y	Y	Y	Y
Controls: total migration		Y	Y		Y	Y
Controls: SES		Y	Y		Y	Y
Predicted composition			Y			Y

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clustered at the mesoregion level in parenthesis. The composition of farmers is the log of the number of farmers in the origin as weighted by the share of farmers in a destination. Regressions include the years of 2000 and 2010.

Table 4: Results from the Quantification of the Model

	Year		
	1950	1980	2010
	(1)	(2)	(3)
a. Migration costs			
Avg migration costs	38.55	25.03	27.21
Migration costs within states	2.77	2.47	2.88
Migration costs between states: East versus West	111.49	51.57	51.28
Migration costs between states: within the West	55.43	24.50	24.23
Migration costs between states: within the East	51.23	31.11	31.70
Elast. of migration costs w.r.t. travel distance	1.13	0.94	0.83
b. Productivity			
Productivity in man in the West relative to the East	0.52	0.78	0.99
Productivity in agr in the West relative to the East	0.45	0.63	0.89
- Soybeans	0.11	0.96	1.57
- Livestock	0.80	0.93	1.22
- Cotton	0.46	0.92	3.45
- Corn	0.58	0.73	1.94
Productivity of land supply in the West relative to the East	1.31	1.47	1.77
c. Trade costs			
Trade cost between Brazil and RoW - manufacturing	7.28	4.68	4.03
Trade cost between Brazil and RoW - agriculture	11.65	5.69	4.26

Notes: This table show results from the calibration of the model. Migration costs are presented in terms of its harmonic average. Productivity is averaged according to the mass of of workers in each activity. International trade costs in agriculture are averaged according to total exports in each sector. The elasticity of interstate migration cost w.r.t. travel distance is the slope of a regression of the log of estimated migration costs between states against the log of travel distance.

Table 5: Effects of Migration Costs on Migration to the West

		No evolution of					
		Migration Costs		Productivities		Other factors	
	Baseline	Level	Shares	Level	Shares	Level	Shares
Year	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1950	0.068	0.068	-	0.068	-	0.068	-
1980	0.113	0.090	52.33	0.099	31.36	0.094	43.24
2010	0.155	0.113	49.15	0.125	34.99	0.117	44.01

Notes: This table show the effects of the evolution of different exogenous factors on the overall migration of workers to the West. Shares represent the proportion of the change in the proportion of the population living in the West that is explained by the evolution of the respective set of shocks.

Table 6: Effects of Migration on RCA

	RCAs in the baseline			% of change in RCA accounted by Migration	
	1950	1980	2010	1980	2010
	(1)	(2)	(3)	(4)	(5)
mfg	1	1	1	-	-
agriculture	6.04	6.14	7.30	143.85	67.07
- RoA	1.37	2.48	2.07	4.11	9.80
- banana	6.10	2.03	0.42	-1.06	-0.07
- chicken	0.00	3.07	9.09	-0.73	0.04
- cacao	327.68	361.41	17.10	-6.31	-1.00
- coffee	786.47	419.55	195.24	-1.66	0.70
- corn	2.54	0.58	5.15	-0.54	17.25
- cotton	11.28	0.77	3.86	-0.13	-25.80
- livestock	1.34	2.02	6.52	16.45	26.37
- rice	0.22	0.07	0.47	-5.49	-0.62
- soybean	0.00	74.30	84.93	6.04	28.92
- sugarcane	25.98	52.49	274.32	0.58	3.55
- tobacco	6.98	15.50	17.19	1.36	-3.81

Notes: This table shows the percentage of the total change in Brazil's RCA that can be attributed to the migration of workers to the West.

Table 7: The Gains from Trade and from Migration

	Gains from Trade				Gains from Migration		
	All	All	Agric	Agric			
	Baseline	No Mig	Baseline	No Mig	Baseline	No Trade	No Ag Trade
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1950	4.70	4.70	3.63	3.63	0.00	0.00	0.00
1980	2.58	2.58	1.18	1.18	1.56	1.55	1.54
2010	3.18	3.16	0.70	0.75	1.71	1.67	1.65

Notes: This table shows our measures of gains from trade under different counterfactual scenarios.

A Data

Employment and Migration.

Our data on migration and employment comes from the decadal demographic and economic census organized by the Brazilian statistical institute IBGE (*Instituto Brasileiro de Geografia e Estatistica*). We use the micro-data from the editions of 1970, 1980, 1991, 2000 and 2010. The information from the census is divided in two questionnaires, an universal one with basic questions about education and the family structure, and a sampled one with detailed information on migration and employment. In 1970 and 1980, 25% of the population was sampled for the detailed questionnaire. For 1990, 2000 and 2010, about 25% of the population was sampled in smaller municipalities and 10% for the larger ones. The municipality thresholds defining the sample size depend on the year of the census. To illustrate the final sample size, in 2000, the census included about 12 million individuals in the sampled questionnaire.

Since the census of 1980, we have information about the current and the previous municipality of residence of each individual in the case of migration if they have migrated less within the previous 10 years. 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. Since less than 0.1% of the population was born abroad, we ignore international migration. For 1970, we have micro-data with information on the state of birth and the state of residence. For 1950 and 1960, we digitized historical records from the census to obtain information on the total population in each state and their state of birth.

For 1980 and 2010, we can use our micro-data to directly construct the migration flows of workers who were born in state s and live in state and activity pair s' and k, which we denote by $L_{ss',kt}$. For 1950, we observe directly only the migration flows from a state s to a state s', which is given by $L_{ss',t}$. We therefore use entropy methods typically used in the construction of input-output matrices to obtain $L_{ss',kt}$. Specifically, we apply the algorithm developed in Ireland and Kullback (1968). The method consists in searching for values of $L_{ss',kt}$ based on a guessed value $\tilde{L}_{ss',kt}$ until they are consistent with the aggregate data on $L_{s',kt}$ and $L_{s,t}$. We use the values from 1980 as guesses for 1950.

The census contains a specific module of labor employment with questions about the sector of employment of the worker. In each edition, there are more than 150 sectors. About 25 of them can be classified as agricultural activities. We identified 12 agricultural activities that have definitions that are consistent over time and that can also be found in other datasets used in the paper: rest of agriculture, banana, chicken, cacao, coffee, corn, cotton, livestock, rice, soy, sugarcane and tobacco. In addition, the census ask questions about total revenues of a worker. Since many agricultural workers do not receive their earning in the form of wages, which is the case of managers who are about 30% of our sample, we use data on total income instead of their information on wages, which is the measure that is theoretically consistent with our model.

Gross Output.

For agriculture, we used data on gross output by agricultural activity from PAM (*Produção Agrícola Municipal*), which is organized by the Brazilian census bureau. It contains municipality level data since 1974 for more than 20 crops and state level data since 1930s for a subset of these crops. We combine this data with information from the agricultural census on agricultural revenues from the agricultural census of 1960, 1970, 1996 and 2006, to obtain gross output in livestock and chicken. We converted the data on gross output into value added and computed the share of value added coming from each agricultural activity within each meso-region. We then multiplied this share of value added by the share of value added in agriculture produced by each meso-region according to IPEA. We multiplied these shares by the total value in levels given by the value added data from United Nations. For manufacturing and services, we brought data on value added by meso-region from IPEA. With data on value added, we constructed gross output using the share of value added in the World Input-Output Database for Brazil in 2010. For 1950, we do not directly observe value added by economic activity at the meso-region level. We therefore use the entropy method from Ireland and Kullback (1968) to adjust our values. In particular, using the algorithm proposed by Ireland and Kullback (1968), we search for values of value added $VA_{j,kt}$ based on guesses of value added $\tilde{VA}_{j,kt}$ until they are consistent with observed value added by state and activity $VA_{s,kt}$ and value added by meso-region $VA_{j,t}$. We use the values from 1980 as guesses of $\tilde{VA}_{j,kt}$ for 1950.

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 our 12 agricultural activities. 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. In our structural estimation, to obtain the global imports of each good, we combined our trade flow data from Comexstat with trade data from FAO-STAT. Since the FAO-STAT does not contain a category for fish, we assumed that 10% of the global trade in agricultural commodities come from trade in fish and fishery products, which is consistent with reports on the fishery industry from FAO.

For 2010, we use 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.

B Calibration Details

This section described the algorithm that we set up for the calibration of the model. The algorithm can be divided in three steps as described below.

Step 1: Absolute Advantages ($T_{j,k,t}$), trade costs ($\tau_{ij,kt}$) and preferences (σ, σ_k, η_k and $a_{k,t}$)

In the first step of our calibration, we set up an algorithm that searches for values of absolute advantages, which we define as $\overline{T}_{j,kt} \equiv (c_{j,kt}/A_{j,kt})^{1-\eta_k}$, and preference shifters ($a_{k,t}$) until the model achieves a perfect fit with data on aggregate apparent consumption by economic activity and the gross output by region and economic activity. In particular, we adjust absolute advantages and preference shifters using

$$X_{j,kt} = \sum_i \frac{\overline{T}_{j,kt}\tau_{ji,kt}^{1-\eta_k}}{P_{j,kt}^{1-\eta_k}} \frac{P_{j,kt}^{1-\sigma_k}}{P_{o,jt}^{1-\sigma_k}} \frac{P_{o,jt}^{1-\sigma}}{P_{j,t}^{1-\sigma}} X_{j,t}$$

where

$$P_{j,t}^{1-\sigma} = \sum_o P_{j,ot}^{1-\sigma}$$

and

$$P_{j,ot}^{1-\sigma_k} = \sum_{k \in K_o} a_{k,t} P_{j,kt}^{1-\sigma_k}$$

and

$$P_{j,kt}^{1-\eta_k} = \sum_i \overline{T}_{j,kt}\tau_{ji,kt}^{1-\eta_k}.$$

Within each iteration of this algorithm, we adjust trade costs based on the parametrization described in the paper until the model matches the elasticity of trade between Brazilian states with respect to distance, the exports and imports of international trade flows with respect to the rest of the world, and the total exports of Brazil in manufacturing. Once the model converges, it spits out the price index of each region, $P_{j,t}$, which we use to construct real income in the next step.

Step 2: Migration costs ($\mu_{ij,kt}$) and land intensity (ν_k)

In the second step, we set up an algorithm that contains an outer and an inner-loop. In the outer-loop, we search for values of the residual migration costs between states and activities ($\tilde{\mu}_{ss',kt}$) and the intercept of migration cost (μ_t) until the model achieves a perfect fit with the migration data between states and activities ($L_{ss',kt}$) and the share of workers living in their region of birth. In the inner-loop, we search for values of ν_k that makes the model achieve a perfect fit with the data in terms of the targeted aggregate cost share of land in Brazil.

In particular, in the inner-loop, given a guess of ν_k, we construct wages using

$$\sum_i \frac{\left(\mu_{ij,kt} s_{i,kt} w_{j,kt}/P_{j,kt}\right)^{\kappa}}{\sum_{k'} \sum_h \left(\mu_{ih,k't} s_{i,k't} w_{h,k't}/P_{h,t}\right)^{\kappa}} s_{i,kt} L_{i,t-1} = \bar{\nu}_{j,kt} \alpha_{k,t} X_{j,kt},$$

where $X_{j,kt}$ is observed gross output, and land rents using

$$r_{j,kt} = (1 - \bar{\nu}_{j,kt}) \alpha_{k,t} X_{j,kt},$$

where $\bar{\nu}_{j,kt}$ is given by

$$\bar{\nu}_{j,kt} = \frac{\nu_{k,t} \left(w_{j,kt}\right)^{1-\rho}}{\nu_{k,t} \left(w_{j,kt}\right)^{1-\rho} + (1 - \nu_{k,t}) \left(r_{j,t}\right)^{1-\rho}}.$$

With these values for implied wages and land rents, we construct aggregate cost share of land implied by the model given our guess of ν_k, which we adjust until we match the targeted values. Noticed that the inner-loop also spits out the labor flows between regions, which is given by

$$L_{ss',kt} = \sum_{i \in s} \frac{\left(\mu_{ij,kt} s_{i,kt} w_{j,kt}/P_{j,kt}\right)^{\kappa}}{\sum_{k'} \sum_h \left(\mu_{ih,k't} s_{i,k't} w_{h,k't}/P_{h,t}\right)^{\kappa}} L_{i,t-1}.$$

Using these values for the migration flow, we search for values of migration costs in the outer-loop until the model matches $L_{ss',kt}$.

Step 3: Natural advantage ($A_{j,kt}$) and land supply productivity ($b_{j,t}$)

Once with values for implied wages and land rents, we can construct the unit cost of production $c_{j,kt}$ and recover the natural advantage of a region from $\overline{T}_{j,kt} = (c_{j,kt}/A_{j,kt})^{1-\eta_k}$. For the land supply, we recover land productivity using $b_{j,t} = (\varsigma P_{j,t}/r_{j,t})^{1/\varsigma}(H_{j,t})^{(\varsigma-1)/\varsigma}$.

C Migration and Comparative Advantage

In this section, we explain how we derive the conditions for comparative advantage used in Section 4. Following French (2017), we focus on conditions that characterize the relative autarky costs of production between two regions and two crops.

C.1 Regional Comparative Advantage

We are interested in understanding when, under free trade ($\tau_{ij,k} = 1$), a region has a comparative advantage in good k relative to k' and relative to Foreign if

$$\frac{X_{iF,k}}{X_{iF,k'}} > \frac{X_{FF,k}}{X_{FF,k'}}.$$

Under the assumptions laid out in the paper, this is equivalent to

$$\begin{aligned} \frac{(w_{i,k}/A_{i,k})^{1-\eta} P_k^{\eta-1} X_{kF}}{(w_{i,k'}/A_{i,k'})^{1-\eta} P_{k'}^{\eta-1} X_{k'F}} &> \frac{(w_{F,k}/A_{F,k})^{1-\eta} P_k^{\eta-1} X_{kF}}{(w_{F,k'}/A_{F,k'})^{1-\eta} P_{k'}^{\eta-1} X_{k'F}} \\ \frac{w_{i,k}^{1-\eta} A_{i,k}^{\eta-1}}{w_{i,k'}^{1-\eta} A_{i,k'}^{\eta-1}} &> \frac{w_{F,k}^{1-\eta} A_{F,k}^{\eta-1}}{w_{F,k'}^{1-\eta} A_{F,k'}^{\eta-1}}. \end{aligned} \tag{21}$$

where X_{kF} is total expenditure of F on goods k and P_k is the price index of goods k, which is the same across regions since there are no trade costs. To obtain conditions relating to exogenous forces in the model, we now solve for wages using labor market clearing. When migration costs are prohibitive, so $\mu_{ijk} \to \infty$ for $i \neq j$, then the labor market clearing condition is

$$w_{i,k} s_{i,k} L_{i,k}^0 = \sum_j (w_{i,k}/A_{i,k})^{1-\eta} P_k^{\eta-1} X_{kj}$$

where X_{kj} is expenditure of region j on goods from k. Isolating $w_{i,k}$ gives

$$w_{i,k} = \left(\frac{1}{A_{i,k}^{1-\eta} s_{i,k} L_{i,k}^0} \sum_j \frac{X_{kj}}{P} \right)^{\frac{1}{\eta}}. \tag{22}$$

Inserting the expression above into 21, we get

$$\frac{\left(A_{i,k}^{1-\eta}s_{i,k}L_{i,k}^{0}\right)^{\frac{\eta-1}{\eta}}A_{i,k}^{\eta-1}}{\left(A_{i,k'}s_{i,k'}L_{i,k'}^{0}\right)^{\frac{\eta-1}{\eta}}A_{i,k}^{\eta-1}}>\frac{\left(A_{F,k}s_{F,k}L_{F,k}^{0}\right)^{\frac{\eta-1}{\eta}}A_{F,k}^{\eta-1}}{\left(A_{F,k'}s_{F,k'}L_{F,k'}^{0}\right)^{\frac{\eta-1}{\eta}}A_{F,k}^{\eta-1}}$$

which simplifies to

$$\left(\frac{s_{i,k}L_{i,k}^{0}A_{i,k}}{s_{i,k'}L_{i,k'}^{0}A_{i,k'}}\right)^{\xi}>\left(\frac{s_{F,k}L_{i,k}^{0}A_{F,k}}{s_{F,k'}L_{i,k'}^{0}A_{F,k'}}\right)^{\xi}$$

where $\xi\equiv(\eta-1)/\eta$.

With free migration, so $\mu_{ijk}=1$ for $i\neq j$, the labor market condition is given by

$$\begin{aligned}\sum_{i'}w_{i,k}\frac{\left(w_{i,k}s_{i',k}\right)^{\kappa}}{\sum_{h}\left(w_{h,k}s_{i',k}\right)^{\kappa}}s_{i',k}L_{i',k}^{0}&=\sum_{j}\left(w_{i,k}/A_{i,k}\right)^{1-\eta}P_{k}^{\eta-1}X_{kj}\\ w_{i,k}^{\eta+\kappa}\left(\sum_{h}w_{h,k}^{\kappa}\right)^{-1}\sum_{i'}s_{i',k}L_{i',k}^{0}&=\sum_{j}\left(1/A_{i,k}\right)^{1-\eta}P_{k}^{\eta-1}X_{kj}\\ w_{i,k}^{\eta+\kappa}\left(\sum_{h}w_{h,k}^{\kappa}\right)^{-1}\Upsilon_{H,k}&=\sum_{j}\left(1/A_{i,k}\right)^{1-\eta}P_{k}^{\eta-1}X_{kj}\end{aligned}$$

where $\Upsilon_{H,k}\equiv\sum_{i'}s_{i',k}L_{i',k}^{0}$. The expression above can be written as

$$w_{i,k}=\left(\frac{1}{A_{i,k}^{1-\eta}}\frac{\sum_{h}w_{h,k}^{\kappa}}{\Upsilon_{H,k}}\sum_{j}\frac{X_{kj}}{P}\right)^{\frac{1}{\eta+\kappa}}.$$

Take the ratio of wages in region h with respect to region i to get

$$w_{h,k}=w_{i,k}\left(\frac{A_{h,k}}{A_{i,k}}\right)^{\frac{\eta-1}{\eta+\kappa}}$$

Using the expression above, we can write wages as

$$\begin{aligned}w_{i,k}&=\left(\frac{1}{A_{i,k}^{1-\eta}}\frac{\sum_{h}\left(w_{i,k}\left(A_{h,k}/A_{i,k}\right)^{\frac{\eta-1}{\eta+\kappa}}\right)^{\kappa}}{\Upsilon_{H,k}}\sum_{j}\frac{X_{kj}}{P}\right)^{\frac{1}{\eta+\kappa}}\\ w_{i,k}&=\left(\frac{1}{A_{i,k}^{1-\eta}}\frac{1}{\mathcal{L}_{i,k}}\frac{1}{\Upsilon_{H,k}}\sum_{j}\frac{X_{kj}}{P}\right)^{\frac{1}{\eta}}\end{aligned}\tag{23}$$

where we defined $\mathcal{L}_{i,k} \equiv A_{i,k}^{\kappa\frac{\eta-1}{\eta+\kappa}} / \sum_h A_{h,k}^{\kappa\frac{\eta-1}{\eta+\kappa}}$. Substitute the expression above into 21 to obtain

$$\left(\frac{\Upsilon_{H,k}}{\Upsilon_{H,k'}}\right)^{\xi}\left(\frac{\mathcal{L}_{i,k}A_{i,k}}{\mathcal{L}_{i,k'}A_{i,k'}}\right)^{\xi} > \left(\frac{s_{F,k}L^0_{F,k}A_{F,k}}{s_{F,k'}L^0_{F,k'}A_{F,k'}}\right)^{\xi}.$$

This complete the proof for the regional comparative advantage

C.2 Aggregate Comparative Advantage

$$\frac{X_{HF,k}}{X_{HF,k'}} > \frac{X_{FF,k}}{X_{FF,k'}}.$$

Under the assumptions laid out in the paper, this is equivalent to

$$\frac{\sum_i (w_{i,k}/A_{i,k})^{1-\eta} P_{F,k}^{\eta-1}\alpha_k X_F}{\sum_i (w_{i,l}/A_{i,l})^{1-\eta} P_{F,l}^{\eta-1}\alpha_l X_F} > \frac{(w_{F,k}/A_{F,k})^{1-\eta} P_{F,k}^{\eta-1}\alpha_k X_F}{(w_{F,l}/A_{F,l})^{1-\eta} P_{F,l}^{\eta-1}\alpha_l X_F}$$

$$\frac{\sum_i w_{i,k}^{1-\eta}A_{i,k}^{\eta-1}}{\sum_i w_{i,l}^{1-\eta}A_{i,l}^{\eta-1}} > \frac{w_{F,k}^{1-\eta}A_{F,k}^{\eta-1}}{w_{F,l}^{1-\eta}A_{F,l}^{\eta-1}}. \tag{24}$$

If we substitute equation 22 into the expression above, we get

$$\frac{\sum_i \left(s_{i,k}L^0_{i,k}A_{i,k}\right)^{\xi}}{\sum_i \left(s_{i,k'}L^0_{i,k'}A_{i,k'}\right)^{\xi}} > \frac{\left(s_{F,k}L^0_{F,k}A_{F,k}\right)^{\xi}}{\left(s_{F,k'}L^0_{F,k}A_{F,k'}\right)^{\xi}}.$$

If we substitute the expression for wages with free mobility 23 into 24, we get

$$\left(\frac{\Upsilon_{H,k}}{\Upsilon_{H,k'}}\right)^{\xi}\frac{\sum_i (\mathcal{L}_{i,k}A_{i,k})^{\xi}}{\sum_i (\mathcal{L}_{i,k'}A_{i,k'})^{\xi}} > \frac{\left(s_{F,k}L^0_{F,k}A_{F,k}\right)^{\xi}}{\left(s_{F,k'}L^0_{F,k}A_{F,k'}\right)^{\xi}}.$$

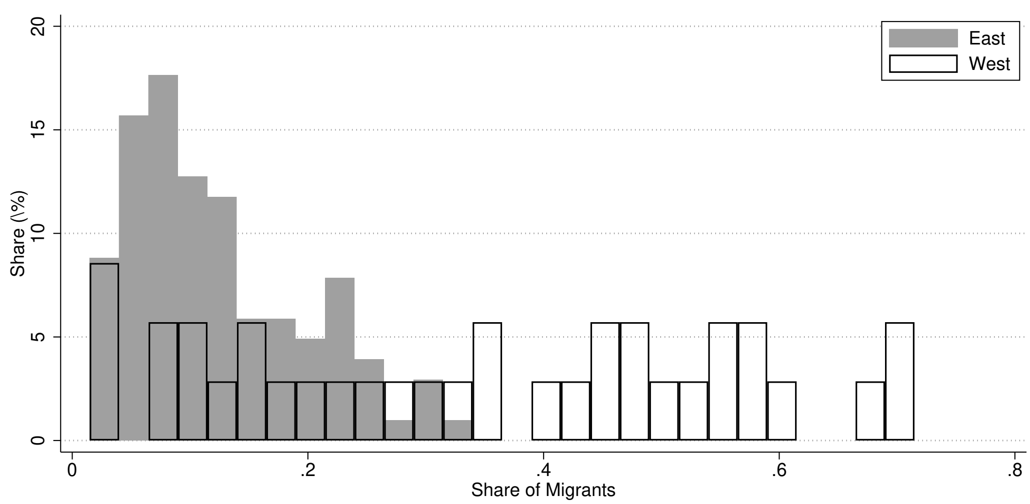
This complete the proof for the aggregate comparative advantage.

D Additional Figures and Tables

Figure 7: A Poster from the Federal Government about the March to the West in the 1940s

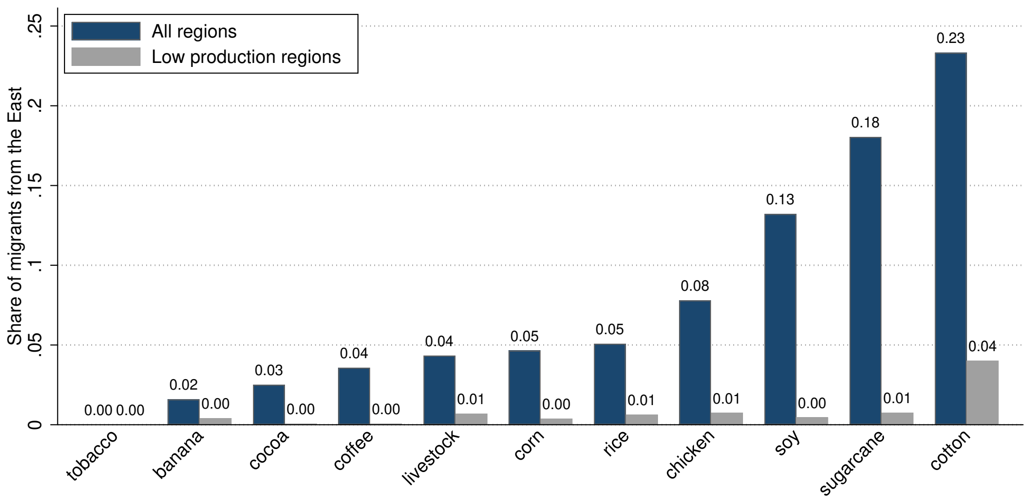
Notes: The man on the right of the figure is Getulio Vargas, who was the Brazilian president first from 1930 to 1945 and second from 1951 to 1954. The quote in the bottom translates to "The true sense of Brazilianness is the March to the West".

Figure 8: Share of Migrants across Meso-Regions in 2010



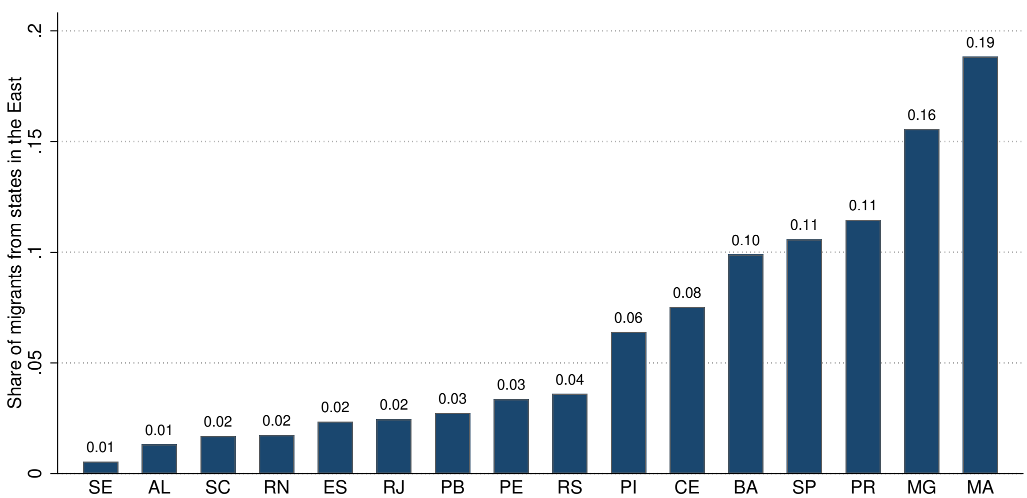
Notes: This figure shows the proportion of migrants as defined by the share of the population that was not born in their state of residence in 2010. The West, which is a recently occupied region, shows a substantially larger share of migrants than the East.

Figure 9: Share of Migrants from the East in the West by Crop (2010)



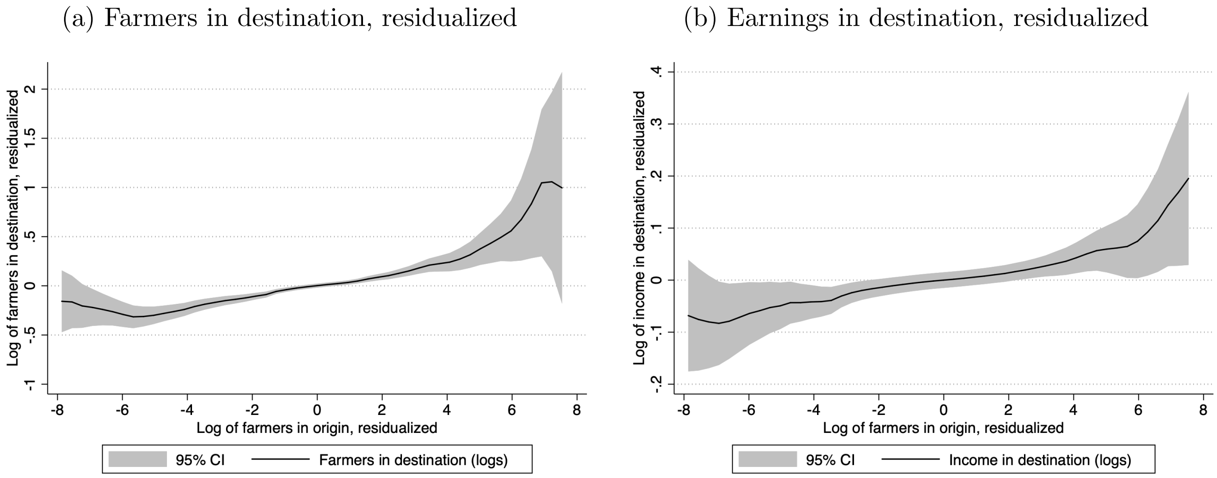
Notes: This figure shows the proportion of migrants from the East producing each crop in the West in 2010. In addition, it shows the proportion of migrants producing each crop in the West coming from regions within the East that have low production of these same crops. We define low producing regions as those that are below the bottom quartile in the distribution of workers employment within the East.

Figure 10: Share of Migrants from the East in the West by State of Origin (2010)



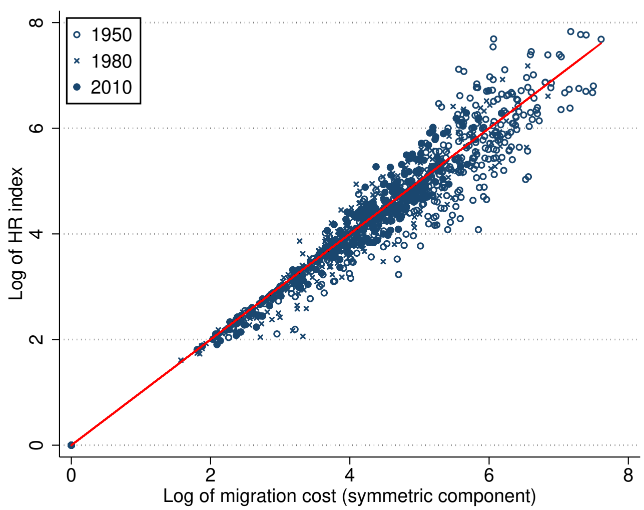
Notes: This figure shows the proportion of migrants from the East in the West according to their state of origin in the East.

Figure 11: Local Polynomial Regressions of the Influence of the Region of Origin on Crop Choice and Earnings of Farmers in their Destination Region



Notes: To construct this figure, we first first absorb destination-crop-year fixed effects from dependent and independent variables. We then run a local polynomial smooth.

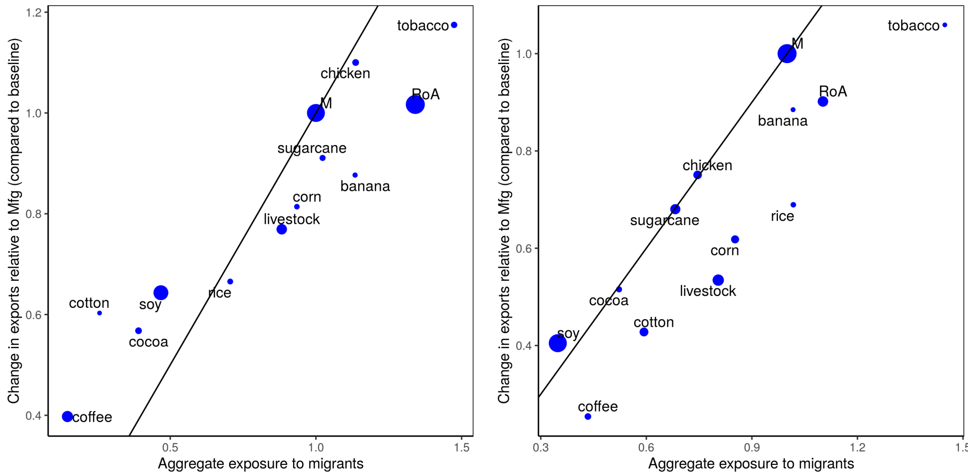
Figure 12: Comparing Estimated Migration Costs with Head-Ries Index



Notes: Panel a shows the migration costs between states as implied by the model. It indicates a sharp decline in migration costs between 1950 and 1980, but a similar distribution between 1950 and 1980. Panel b shows the correlation between our estimates of migration costs and the average travel distance between states.

Figure 13: Exposure to Migrants and Changes in Export in the Aggregate in the West

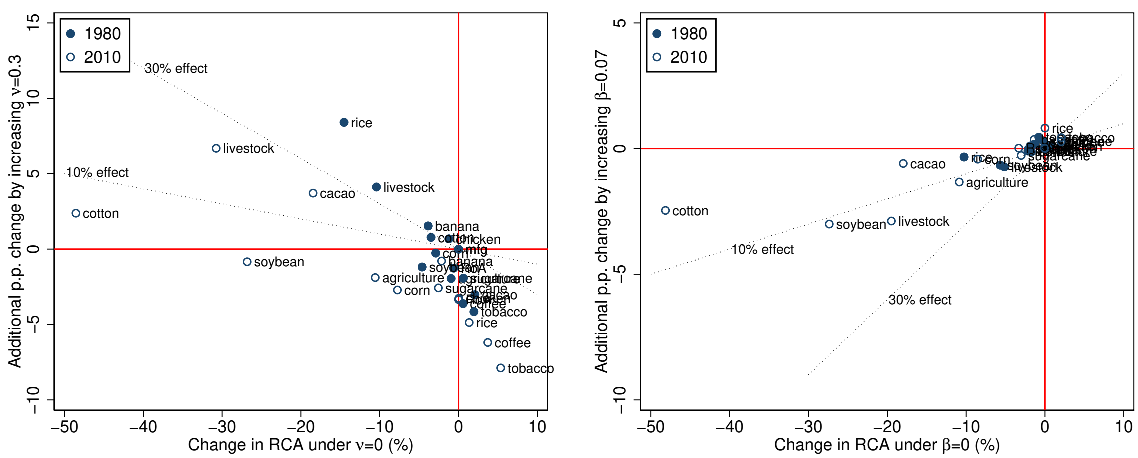
(a) 1980 (b) 2010



Notes: This figure shows the relationship between aggregate changes in relative exports in the West and the aggregate exposure of each crop to migrants from the East.

Figure 14: Examining the Effects of Knowledge and Land

(a) Homogeneous Land Intensity within Agriculture (b) Adding Knowledge



Notes: This figure shows the changes in Brazil's theoretically consistent RCA when we add different margins to the model. The reference dotted lines show the percent effect of 10% and 30%.

Table 8: The Relationship between Farmers' Choices and Earnings and their Region of Origin with Controls for previous Migration

	OLS	OLS	OLS	PPML	PPML
	(1)	(2)	(3)	(4)	(5)
a. Farmers in destination (logs)					
Farmers in origin	0.070***	0.058***	0.079***	0.105***	0.109***
	(0.006)	(0.012)	(0.020)	(0.012)	(0.019)
R^2	0.189	0.731	0.756	-	-
Obs	8265	8265	6142	8265	6142
b. Earnings (logs)					
Farmers in origin	0.020***	0.023**	0.069***	0.030***	0.069***
	(0.004)	(0.009)	(0.016)	(0.011)	(0.016)
R^2	0.345	0.692	-	-	-
Obs	7520	7520	5597	7520	5597
c. Farmers in destination (logs) - Controls for previous migration					
Farmers in origin	0.062***	0.058***	0.080***	0.105***	0.109***
	(0.007)	(0.012)	(0.020)	(0.012)	(0.019)
R^2	0.198	0.732	0.758	-	-
Obs	8265	8265	6142	8265	6142
d. Earnings (logs) - Controls for previous migration					
Farmers in origin	0.021***	0.023**	0.069***	0.030***	0.069***
	(0.004)	(0.009)	(0.016)	(0.011)	(0.016)
R^2	0.345	0.692	-	-	-
Obs	7520	7520	5597	7520	5597
e. Farmers in destination (logs) - Migrants from state of birth					
Farmers in origin	0.082***	0.091***	0.089***	0.134***	0.127***
	(0.008)	(0.017)	(0.028)	(0.017)	(0.026)
R^2	0.241	0.740	0.750	-	-
Obs	5385	5385	3999	5385	3999
f. Earnings (logs) - Migrants from state of birth					
Farmers in origin	0.026***	0.025**	0.063***	0.038***	0.063***
	(0.006)	(0.012)	(0.023)	(0.012)	(0.023)
R^2	0.361	0.696	-	-	-
Obs	4975	4975	3708	4975	3708
Dest-Act-Year FE	Y	Y	Y	Y	Y
Dest-Orig-Year FE		Y	Y	Y	Y
Above Q1			Y		Y
Include zeros					

Notes: * / ** / *** denotes significance at the 10 / 5 / 1 percent level. Standard errors clustered at the destination-crop-year level in parenthesis. This table replicates our main table in Panels a and b. In panels c and d we include the share of return migrants and the share of farmers who come from their region of origin, which control for third migration of workers between regions. In panels e and f we run regressions only with workers who come from their state of birth, which accounts for 40% of our sample.

Table 9: The Relationship between Farmers' Choices and Earning and their Region of Origin (OLS - Robustness)

Geographic Unit	Meso	Meso	Meso	Meso	Meso	Micro	Municipality	Meso	Meso	Meso	Meso
Lag (years)	30	20	10	20	10	30	30	30	30	30	30
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
a. Farmers in destination (logs)											
Farmers in origin	0.070***	0.079***	0.089***	0.096***	0.121***	0.050***	0.037***	0.070***	0.078***	0.079***	0.080***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.004)	(0.002)	(0.006)	(0.008)	(0.008)	(0.006)
R^2	0.189	0.192	0.197	0.206	0.204	0.254	0.616	0.189	0.190	0.190	0.183
Obs	8265	9337	9289	15638	26266	16864	24536	8265	8265	8265	10901
b: Earnings (logs)											
Farmers in origin	0.020***	0.016***	0.020***	0.014***	0.019***	0.019***	0.018***	0.018***	0.014**	0.015***	0.020***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)	(0.003)
R^2	0.345	0.370	0.373	0.333	0.343	0.383	0.533	0.468	0.346	0.468	0.344
Obs	7520	8454	8413	14661	25020	15292	22190	7520	7520	7520	9901
c. Farmers in destination (logs) - Above Q1											
Farmers in origin	0.071***	0.100***	0.155***	0.113***	0.175***	0.072***	0.054***	0.071***	0.070***	0.070***	0.083***
	(0.011)	(0.010)	(0.012)	(0.010)	(0.009)	(0.007)	(0.005)	(0.011)	(0.012)	(0.012)	(0.010)
R^2	0.215	0.218	0.230	0.228	0.224	0.270	0.607	0.215	0.215	0.215	0.207
Obs	6142	6952	6895	11646	19522	12599	18311	6142	6142	6142	8096
b: Earnings (logs) - Above Q1											
Farmers in origin	0.020**	0.022**	0.031***	0.021***	0.029***	0.029***	0.035***	0.022***	0.015	0.020**	0.018**
	(0.009)	(0.009)	(0.010)	(0.008)	(0.005)	(0.006)	(0.007)	(0.008)	(0.010)	(0.009)	(0.007)
R^2	0.365	0.391	0.392	0.352	0.365	0.403	0.557	0.486	0.366	0.486	0.359
Obs	5597	6319	6268	10956	18651	11497	16622	5597	5597	5597	7353
Dest-Act-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dest-Orig-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SES								Y		Y	
Control function									Y	Y	
Years: 2000-2010	Y	Y	Y			Y	Y	Y	Y	Y	Y
Years: 1990-2010				Y							
Years: 1980-2010					Y						
Age: 30-60	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
Age: 20-											Y

Notes: * / ** / $^{+}$ denotes significance at the 10 / 5 / 1 percent level. Standard errors clustered at the destination-crop-year level in parenthesis. This table shows results using several alternative specifications relative to our main results in the paper using OLS. Our sample exclude non-migrants.

Table 10: The Relationship between Farmers' Choices and Earning and their Region of Origin (PPML - Robustness)

Geographic Unit	Meso	Meso	Meso	Meso	Meso	Micro	Municipality	Meso	Meso	Meso	Meso
Lag (years)	30	20	10	20	10	30	30	30	30	30	30
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
a. Farmers in destination (logs)											
Farmers in origin	0.140***	0.149***	0.168***	0.213***	0.283***	0.091***	0.058***	0.140***	0.142***	0.143***	0.148***
	(0.013)	(0.013)	(0.014)	(0.023)	(0.022)	(0.008)	(0.004)	(0.013)	(0.015)	(0.015)	(0.013)
Obs	8265	9337	9289	15638	26266	16864	24536	8265	8265	8265	10901
b: Earnings (logs)											
Farmers in origin	0.049***	0.031**	0.051**	0.017	0.024**	0.034***	0.034***	0.045**	0.040**	0.042**	0.041***
	(0.017)	(0.012)	(0.022)	(0.011)	(0.009)	(0.009)	(0.010)	(0.017)	(0.018)	(0.017)	(0.014)
Obs	7520	8454	8413	14661	25020	15292	22190	7520	7520	7520	9901
c. Farmers in destination (logs) - Above Q1											
Farmers in origin	0.147***	0.183***	0.266***	0.260***	0.404***	0.122***	0.090***	0.147***	0.129***	0.129***	0.154***
	(0.021)	(0.024)	(0.022)	(0.037)	(0.034)	(0.014)	(0.009)	(0.021)	(0.024)	(0.024)	(0.020)
Obs	6142	6952	6895	11646	19522	12599	18311	6142	6142	6142	8096
b: Earnings (logs) - Above Q1											
Farmers in origin	0.041**	0.023	0.046*	0.018	0.022	0.035**	0.070***	0.035**	0.042*	0.039*	0.019
	(0.020)	(0.024)	(0.024)	(0.021)	(0.017)	(0.014)	(0.018)	(0.017)	(0.023)	(0.020)	(0.018)
Obs	5597	6319	6268	10956	18651	11497	16622	5597	5597	5597	7353
Dest-Act-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dest-Orig-Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SES								Y		Y	
Control function									Y	Y	
Years: 2000-2010	Y	Y	Y			Y	Y	Y	Y	Y	Y
Years: 1990-2010				Y							
Years: 1980-2010					Y						
Age: 30-60	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
Age: 20-											Y

Notes: * / ** / $^{+}$ denotes significance at the 10 / 5 / 1 percent level. Standard errors clustered at the destination-crop-year level in parenthesis. This table shows results using several alternative specifications relative to our main results in the paper using PPML. Our sample exclude non-migrants.

Table 11: The Relationship between Farmers' Earning and their Region of Origin (Individual Level Regressions)

	OLS	OLS	OLS	PPML	PPML
	(1)	(2)	(3)	(4)	(5)
a. Earnings (logs)					
Farmers in origin	0.009**	0.018***	0.055***	0.041***	0.084***
	(0.003)	(0.005)	(0.011)	(0.009)	(0.016)
R^2	0.261	0.389	0.394	-	-
Obs	20107	20107	14848	20107	14848
b. Earnings (logs) - SES controls					
Farmers in origin	0.007**	0.016***	0.044***	0.036***	0.067***
	(0.003)	(0.005)	(0.011)	(0.009)	(0.016)
R^2	0.372	0.464	0.465	-	-
Obs	20107	20107	14848	20107	14848
c. Earnings (logs) - Controls for previous migration					
Farmers in origin	0.007**	0.016***	0.044***	0.036***	0.067***
	(0.003)	(0.005)	(0.011)	(0.009)	(0.016)
R^2	0.372	0.464	0.465	-	-
Obs	20107	20107	14848	20107	14848
d. Earnings (logs) - Migrants from state of birth					
Farmers in origin	0.012**	0.015*	0.052***	0.044***	0.086***
	(0.005)	(0.008)	(0.015)	(0.012)	(0.020)
R^2	0.283	0.413	0.413	-	-
Obs	12801	12801	10046	12801	10046
Dest-Act-Year FE	Y	Y	Y	Y	Y
Dest-Orig-Year FE		Y	Y	Y	Y
Above Q1			Y		Y

Notes: See table 2 for a description of columns. This table shows results using individual level data. Our sample excludes non-migrants.

Table 12: Summary Statistics by Activity (in %)

	services	mfg	agriculture	rest of agri	banana	chicken	cocoa	coffee	corn	cotton	livestock	rice	soy	sugarcane	tobacco
				Percentage within Agriculture											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
a. Value Added															
-1950	53.5	21.6	24.8	26.3	1.2	5.5	1.3	14.9	7.0	6.5	25.8	6.5	0.2	3.9	1.0
-1980	56.5	33.8	9.7	22.0	1.4	7.7	2.3	6.7	8.7	2.7	25.3	6.7	8.6	6.7	1.1
-2010	74.0	21.0	5.0	20.4	2.2	7.8	0.6	6.1	6.6	1.7	23.9	3.6	15.0	9.7	2.4
b. Workers															
-1950	37.7	4.2	58.1	59.3	0.7	0.1	0.8	11.2	10.2	3.7	3.0	5.6	0.2	4.2	0.9
-1980	68.2	14.7	17.1	36.7	0.7	0.6	1.9	8.0	14.0	6.1	10.2	11.3	3.9	4.7	1.8
-2010	87.5	7.6	4.9	53.6	1.3	4.1	1.2	8.2	7.5	0.1	13.1	2.9	1.4	3.4	3.1
c. Exports from Brazil to the ROW															
-1950	0.0	20.1	79.9	14.6	0.5	0.0	8.9	58.6	1.3	6.2	2.1	0.4	0.0	5.6	1.7
-1980	0.0	55.6	44.4	25.7	0.2	2.5	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	17.8	0.1	10.6	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	0.0	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	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	87.6	0.0	0.2	2.2	0.3	1.6	1.2	2.1	3.3	0.8	0.0	0.6

Notes: This table shows summary statistics of the aggregate data for 1950, 1980 and 2010.

Table 13: Summary Statistics - Aggregate

	1950	1960	1970	1980	1990	2000	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a. Migration							
- East to East	0.874	0.845	0.825	0.759	0.735	0.719	0.687
- East to West	0.085	0.105	0.137	0.191	0.205	0.205	0.218
- East to West + West to East	0.107	0.129	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	0.019	0.026	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	99.416	94.234	68.613	43.941	50.180	51.971	31.768

Notes: This table shows summary statistics of the main data used in the calibration of the model for 1950, 1980 and 2010. Data from the national account comes from the annual statistical yearbooks from Brazil.