Trade, Jobs, and Worker Welfare*

Erhan Artuc†

Paulo Bastos‡

Eunhee Lee§

November 2019

Abstract

We study welfare effects of trade on workers across different regions and sectors by introducing a new dynamic discrete choice model of labor mobility with endogenous number of choices. In our general equilibrium model, trade shocks impact worker welfare not only through wages, but also via the number of job opportunities available to workers in different labor markets. First, we exploit differential exposure of sectors and regions to destination-specific demand shocks to estimate the impacts of export shocks on wages, employment and labor mobility, using detailed employer-employee panel data for Brazil. Second, we employ the same empirical strategy to estimate structural parameters and the different components of the change in model-implied worker welfare. Third, we use our model and the estimated structural parameters to perform counterfactual policy simulations. The structural IV estimates reveal that the job opportunities channel that we introduce accounts for a sizable share of the losses in worker lifetime welfare following a negative shock to exports.

Keywords: Trade shocks, jobs, wages, worker mobility, adjustment costs, worker welfare.

JEL Classification: F1

*The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. Research for this paper has been supported in part by the World Bank's Umbrella Facility for Trade. We are grateful to Costas Arkolakis, Tibor Besedes, Lorenzo Caliendo, Rafael Dix-Carneiro, Max Dvorkin, Pablo Fajgelbaum, Aaditya Mattoo, Guy Michaels, Ricardo Reyes-Heroles, Ana Maria Santacreu, Sharon Traiberman, and participants at several seminars and conferences for helpful comments and discussions. Nicolas Santos provided excellent research assistance. We remain responsible for any errors.

†Development Research Group, World Bank, eartuc@worldbank.org.

‡Lisbon School of Economics and Management, Universidade de Lisboa, pbastos@iseg.ulisboa.pt.

§Department of Economics, University of Maryland, eunhee11@umd.edu.

1 Introduction

How do trade shocks impact workers? Answering this question requires an understanding of how trade shocks affect their wages and job options they can choose from. In this paper, we develop and estimate a new dynamic discrete choice model with endogenous number of choices to quantify the impacts of trade shocks on labor mobility along various dimensions and worker's lifetime utility. Dynamic models of trade-induced labor mobility have explored wage differentials and idiosyncratic utility as drivers of mobility across sectors, regions and occupations. We emphasize an additional important motive of mobility: *the number of job opportunities* provided by different sectors and regions.

This new channel matters for workers for two main reasons. First, if a worker can choose her job out of more opportunities, it is more likely that the best one delivers higher welfare. Second, even when she is hit by a negative labor demand shock in the future, it is more likely that she will be able to find another job without having to move to a different region or sector. Thus, a region-sector pair (henceforth referred to as labor market) receiving a positive trade shock will attract more workers not just because it provides a higher wage, but also because of the larger number of job opportunities that are created there. In addition, a labor market with a positive trade shock will see a larger internal churning, *i.e.*, more job switching within the labor market, which has been largely overlooked in the literature. Our model clearly shows how trade shocks affect workers differently through labor mobility between and within labor markets. The model delivers a structural equation of trade-induced change in worker's lifetime welfare, which can be conveniently estimated through an instrumental-variable strategy exploiting differential exposure of sectors and regions to destination-specific import demand shocks.

Our framework is motivated by reduced-form evidence on the effects of export shocks on labor markets. The empirical analysis draws on rich employer-employee panel data combined with customs records on export transactions from Brazil in the period 2003-2015. To account for the endogeneity of exports, we construct an instrument at the labor market level, exploiting variation over time in sectoral import demand directed to the region—defined as trade-weighted sectoral imports of the initial set of destinations, sourced from all countries other than Brazil. The empirical analysis draws on data on more than 500 regions and 3 broad sectors. The IV estimates reveal a positive causal effect of exports on residual wages, employment, worker inflows, and job turnover rates in the corresponding labor market. This evidence supports the relevance of the various drivers of trade-induced labor mobility we explore.

Building on this reduced-form evidence, we develop a new dynamic general equilibrium model of labor mobility with an endogenous number of choices. Different labor markets offer different wages and different numbers of job opportunities to workers. A worker chooses the job which gives her the highest utility, where the number of jobs in each labor market is endogenously determined. This is a distinctive feature of our framework compared to previous dynamic models which assume that workers choose a labor market and that the number of different jobs is exogenously fixed, and the same across labor markets. In our model, a worker's job choice determines which labor market she belongs to. Both the wage and the set of job opportunities provided by each labor market are factored into her optimal job choice.

In a labor market with relatively more job opportunities, workers can choose optimally out of more potential jobs, to each of which workers attach idiosyncratic preference. We assume that this idiosyncratic preference for jobs follows a type I extreme value distribution. This is the first channel through which a labor market with more job opportunities provides workers of a greater utility, because the maximum utility will be higher with more options. The second channel stems from frictions to worker mobility. We assume that a job switch requiring a change of labor market implies incurring a higher switching cost compared to a job switch within a labor market. Therefore, a growing labor market with more job opportunities reduces the risk of having to pay a higher switching cost in the future. The prospect of job switch generates an option value in worker's welfare. Our model further decomposes this option value into the option value associated alternative job opportunities within the current labor market and the option value from having alternative jobs in all other labor markets.

Our model delivers a structural equation of changes in worker welfare which is a function of only the estimated probability of moving between labor markets and the labor supply elasticity. The welfare result does not depend on the moving cost structure, observed changes in future wages, or moving probabilities across jobs within a labor market. The effects of a trade shock are fully embedded in the gross flows between labor markets. This is a powerful result which greatly simplifies the analysis of the welfare impacts of trade shocks.

For the welfare analysis, we structurally estimate the model using the worker-firm data from Brazil. In the first stage of the estimation, we pin down the common value attached to each labor market and the moving cost between labor markets for each worker group using a gravity-like equation. The implied probability of moving between labor markets is then calculated with the estimated value of each labor market and the estimated moving cost. We find the the regional moving cost coefficient in the gravity equation is approximately equal to -1, which is consistent with the migration literature. We find that the moving cost between sectors is equivalent to one time loss of approximately 65% of annual wage, which is also consistent with the estimates of Dix-Carneiro (2014). In the second stage of the estimation, we pin down the labor supply elasticity of our model. We first derive an estimable equation describing the relationship between a change in the transformed value of the labor market and a change in wages, with the labor supply elasticity governing the responsiveness of the former with respect to the latter. We exploit variation in residual wages induced by the instrument we used earlier: the trade-weighted change in import demand directed to the labor market. Armed with the estimate of the labor supply elasticity, we estimate the effect of trade shocks in Brazil on workers' welfare, employment and wages using the same instrument. We find that the lifetime welfare of a median formal sector worker increases by 68% of the annual wage, following the rise in exports observed during the sample period; wages increase by 32%, while employment increases by 23%.

Related literature. This paper bridges dynamic models of labor mobility and reduced-form differential exposure methods of quantifying the impacts of trade shocks on worker welfare. Artuç, Chaudhuri, and McLaren (2010, ACM, henceforth), Dix-Carneiro (2014) and Traiberman (2019) study the dynamic transmission of international trade shocks on labor markets via labor mobility by modeling worker's idiosyncratic preference for a labor market with an extreme value distribution.[1] Using these models, they structurally estimate labor market frictions on the basis of differential labor market outcomes across sectors or occupations, such as wages or labor flows. We follow the convention of this literature when modeling worker preferences, but we introduce a new channel which affects worker welfare: the number of jobs within each labor market. Workers choose their job, and which labor market they belong to is a consequence of their choice. In our model, the number of jobs in each labor market is endogeneous. Hence, trade shocks can affect labor mobility and thus welfare not just through wages but also through the number of jobs.

By endogenizing the number of jobs and bringing it to the welfare analysis of trade shocks, we combine the main strength of dynamic models of labor mobility with that of the reduced-form liter-

[1] McLaren (2017) offers a review of this literature. A related strand of work develops trade models featuring search generated labor market frictions frictions, including Davidson, Martin, and Matusz (1999), Coşar, Guner, and Tybout (2016), Helpman, Itskhoki, Muendler, and Redding (2017) and Ritter (2015). In standard search models explored in the trade literature, the number of jobs matters as it affects employment probabilities. In our model, however, workers are matched to multiple jobs. Therefore, our framework can account for the welfare effects generated by changes in both employment probabilities and number of job options. The model remains tractable and allows for the estimation of deep parameters using simple and transparent reduced-form econometric methods.

ature on local labor market effects of trade, including influential contributions by Topalova (2010), Kovak (2013), Autor, Dorn, and Hanson (2013), Autor, Dorn, and Hanson (2015), McLaren and Hakobyan (2016), Dix-Carneiro and Kovak (2015), Dix-Carneiro and Kovak (2017) and Dix-Carneiro and Kovak (2019). This literature builds on the existence of frictions to spatial labor mobility to establish a strong reduced-form relationship between trade shocks and employment changes in local labor markets. However, the reduced-form approach is unable to estimate the implications of these effects for worker lifetime welfare.[2] Our model answers this welfare question by bringing this employment channel into a structural dynamic model of labor mobility and providing a welfare equation. We estimate this welfare equation using an instrumental-variable strategy analogous to that used in the reduced-form literature.

In a recent important contribution, Caliendo, Dvorkin, and Parro (forthcoming, CDP, henceforth) embed the dynamic discrete choice worker problem in Artuç, Chaudhuri, and McLaren (2010) into a multi-country, multi-region, and multi-sector general equilibrium model with trade and migration costs. In important departures from Caliendo, Dvorkin, and Parro (forthcoming), we endogenize the number of job opportunities available to workers in each labor market, and structurally estimate the welfare effects and other important primitives of the model using an instrumental variables strategy. These structural estimates allow us to perform interesting policy simulations. Since we explicitly estimate the moving cost for each time period, we can examine how specific policy interventions influencing different frictions to labor mobility impact workers' welfare gains from trade. By separately changing the distance-related component and the sector-related component of the moving cost, our model can shed light on which dimension policies should target with a prioritize in order to mitigate losses in workers' imposed by trade shocks.[3]

Interestingly, the welfare equation we derive relates closely to several others in the trade literature. Arkolakis, Costinot, and Rodriguez-Clare (2012) provide a sufficient statistic of welfare

[2] A related body of literature uses reduced-form methods to examine the effects of trade shocks on labor market outcomes at the industry, firm or worker level, including contributions by Revenga (1992), Verhoogen (2008), Brambilla, Lederman, and Porto (2012), Amiti and Davis (2012), Bertrand (2004), Hummels, Jøorgensen, and Xiang (2014), Autor, Dorn, Hanson, and Song (2014), Acemoglu, Autor, Dorn, Hanson, and Price (2016), Pierce and Schott (2016) and Frías, Kaplan, Verhoogen, and Alfaro-Serrano (2018). Harrison, McLaren, and McMillan (2011) provide an overview of the literature on trade and inequality.

[3] Galle, Rodriguez-Clare, and Yi (2017) develop a static multi-sector gravity model with heterogeneous workers to quantify the aggregate and group-level welfare effects of trade, and estimate a key structural parameter using the China shock as in Autor, Dorn, and Hanson (2013). Also in static setting, Adao, Arkolakis, and Esposito (2019) exploit the same source of variation in a model implied optimal IV to estimate the labor allocation elasticity. In a key departure from their work, we develop and structurally estimate a dynamic trade model of labor mobility with an endogenous number of choices. This framework allows us to emphasize the importance of the changing number of job opportunities for the dynamics of labor reallocation across regions and sectors following a trade shock, and quantify the implications of this channel for worker lifetime welfare.

gains from trade, which is consistent with various classes of trade models. Since workers are assumed to be homogeneous in the trade models covered by that sufficient statistic result, each worker has the same welfare gains from trade. In our model, worker welfare depends on their actual mobility, and thus is allowed to differ across workers. However, we still maintain the same spirit of Arkolakis, Costinot, and Rodriguez-Clare (2012) by deriving a new formulation of the change in the welfare experienced by workers following a trade shock. The structural welfare equation we derive makes it possible to compare the welfare implications of the previously discussed existing models, isolating the relative importance of each transmission channel of trade shocks.

Roadmap. The paper is organized as follows. In Section 2 we document various reduced-form evidence about the effect of export shocks on labor market outcomes using the RAIS database from Brazil. Motivated by the reduced-form evidence, in Section 3, we introduce a new framework of dynamic labor mobility with international trade and the endogenous number of job opportunities. In Section 4, we discuss how we estimate the key structural parameters of the model and evaluate trade-induced welfare changes by combining the local labor market approach with our structural model. In Section 5, we validate our mechanism with macro-level data. We then discuss how we simulate our model in Section 6, before summarizing our main conclusions in Section 7.

2 Data and reduced-form evidence

2.1 Data sources

The empirical analysis in this paper combines and examines several sources of panel data from Brazil spanning the period 2003-2015. In this section, we provide a brief description of each data source, while giving further details in Appendix A.1.

The main source of data is *Relacao Anual de Informacoes Sociais* (RAIS), a labor census gathering longitudinal data on the universe of formal workers and firms in Brazil, covering the period 2003-2015. RAIS is a high-quality administrative census of formal employees and employers collected every year by the Brazilian Ministry of Labor. These records are used by the government to administer several government benefits programs. Workers are required to be in RAIS in order to receive payments of these programs and firms face fines for failure to report.

RAIS covers virtually all formal workers and provides yearly information on demographics (age, gender, and schooling), job characteristics (detailed 6-digit occupation, wage, hours worked), as well as hiring and termination dates. For each job, the RAIS annual record reports average yearly earnings, as well as the monthly wage in December. We use the information on the December wage, so as to ensure that all labor market outcomes are measured at the same time and avoid potential mismeasurement for workers that did not work full year. RAIS also includes information on a number of establishment-level characteristics, notably number of employees, geographical location (municipality) and industry code (according to the 5-digit level of the Brazilian National Classification of Economic Activities). Unique identifiers (tax identification numbers) for workers and establishments make it possible to follow them over time. The establishment identifier contains 12 digits and the first 8 digits make it possible to uniquely identify the firm. Therefore, it is possible to identify and track multi-establishment firms.

While the RAIS data cover also segments of the public sector, we restrict the analysis to the private sector. We will use the detailed classification of occupations as a measure of the number of different jobs available in a labor market. The Brazilian Classification of Occupations changed in 2002 (CBO-2) and has been reported consistently since 2003. Although the RAIS data are available for earlier years, we restrict the analysis to the post-2003 period in order to ensure that this important variable is defined in a consistent way throughout the period of analysis. There are 2637 occupation codes at the 6-digit level during this period. We use information on the establishment's location (municipality) and industry, and worker-level data on gender, age, education and December wage. We focus on workers aged 16 to 64 years old. As in Dix-Carneiro and Kovak (2017), we use the "microregion" concept of the Brazilian Statistical Agency (IBGE) to define regional boundaries. This definition groups together economically integrated contiguous municipalities with similar geographic and productive attributes. We consider a set of 558 consistently defined microregions, grouping the 5571 municipalities in the data. To ensure a consistent definition of microregions over time, when necessary we merge microregions whose boundaries changed over the period of analysis.

We merge the RAIS with customs records on export transactions by microregion, industry and destination in each year. These customs records are administrative data collected by *Secretaria do Comercio Externo* (SECEX) of the Ministry of Development, Industry and Foreign Trade. These data are originally defined at the level of the municipality, detailed product category and destination market, and are available since 1997. For consistency with the RAIS data, we restrict the analysis to the post-2003 period and aggregate the customs records up to microregion-sector level. To construct an instrument for exports, we further use yearly data on the industry-level imports of each Brazilian destination (sourced from all countries except Brazil). There are 189 destinations in total reported in the customs data, to which we link information on sectoral import demand from the UN COMTRADE.

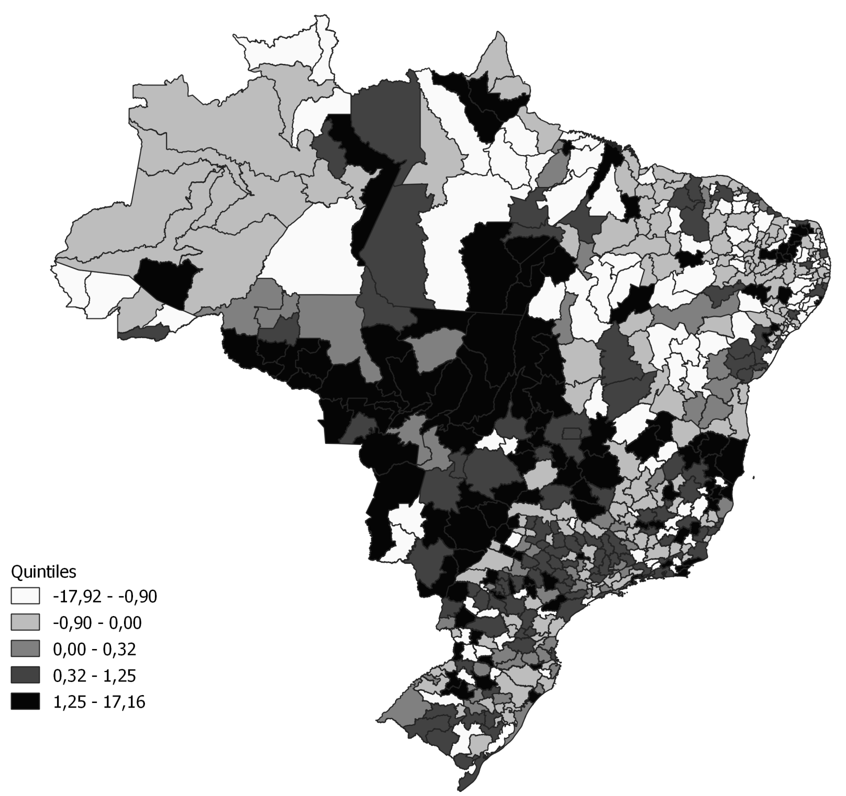
2.2 Econometric model

We now describe the econometric strategy for examining the effects of export shocks on labor markets. We adopt the following baseline specification:

$$\Delta y_t^k = \tilde{\beta} \Delta Z_t^k + \lambda_t + \epsilon_t^k, \tag{1}$$

where y_t^k denotes the log of the outcome variable of interest in the sector-region pair k in year t, Z_t^k denotes the log of export revenue originated in the same labor market, λ_t denotes a year fixed effect, and ϵ_t^k is the error term. The Δ operator denotes the change in a variable between year t and year $t-1$.

Figure 1: Change in export revenue, 2004-2014



Notes: Figure depicts the change in log of (1+exports) in Brazilian microregions during 2004-2014.

Figure 1 depicts the change in export revenue observed in different microregions over the period 2004-2014. It reveals that there exists substantial heterogeneity in the direction and the magnitude of the change in exports across space, which is convenient for identification. Notice also that because some labor markets were initially more export-oriented than others, they differ in the extent to which they are exposed to a given percentage change in export revenue. This heterogeneity is illustrated in Figure A1 in the Appendix, which depicts the distribution of export revenue per worker across microregions in 2003. Initial exports per worker also vary across sectors within each microregion. The same percentage change in exports would therefore be expected to have a stronger impact on labor market outcomes in labor markets where exports per worker were higher to begin with. To account for this, each observation of ΔZ_t^k is weighted by the export revenue per worker observed in the corresponding labor market in 2003.

An important concern is that changes in exports are potentially endogenous to changes in labor market outcomes. For example, lower wages or growing job turnover might cause an increase in export activity. Changes in both variables could also reflect the role of omitted variables, such as underlying changes in infrastructure or technology. To address potential endogeneity in the relationship between exports and labor market outcomes, we adopt an instrumental variables approach. An important challenge in constructing the instrument is to identify a source of variation at the microregion-sector level. Our strategy relies on variation over time in sectoral import demand directed to the region. This strategy builds on the fact that changes in external demand in a particular destination do not matter equally for all labor markets; it matters more for labor markets that initially shipped a larger share of their exports to that destination. Our instrument is therefore defined as the trade-weighted sectoral imports of the initial set of destinations of the microregion-sector (sourced from all countries other than Brazil), where the weights are the export shares of each destination within each microregion-sector cell in 2003. Formally, the change of import demand directed to labor market k in year t can be written as:

$$\Delta \bar{Z}_t^k = \Delta \sum_d \gamma_{d,k,2003} IM_{dI_k t}, \tag{2}$$

where $IM_{dI_k t}$ denotes each destination d's total imports (excluding Brazil) in sector I_k in year t and $\gamma_{d,k,2003}$ the share of exports of labor market k to destination d in 2003. The Δ operator denotes the change in a variable between year t and year $t-1$. We take logs on the summation before taking first differences. Since different labor markets tend to serve different destinations, they vary in the degree to which they are exposed to changes in sectoral import demand from different countries. This heterogeneity is illustrated in Figure A2 in the Appendix, which depicts the main export destination of each microregion in 2003.[4] For the reasons discussed above, we weight each observation of the instrument by the corresponding export revenue per worker observed in 2003. The intuition behind this instrumental variables approach follows closely that adopted in the local labor markets literature, including Topalova (2010), Kovak (2013), Autor, Dorn, and Hanson (2013), Dix-Carneiro and Kovak (2017) and Dix-Carneiro and Kovak (2019). It is also closely related to earlier works using trade-weighted relative prices or import demand as the source of variation in imports or exports at the industry level, such as Revenga (1992) and Bertrand (2004); as well as with more recent work exploiting similar sources of variation at the firm-level, including Brambilla, Lederman, and Porto (2012), Bastos, Silva, and Verhoogen (2018) and di Giovanni, Levchenko, and Mejean (2018).

2.3 Summary statistics and results

Table 1 presents summary statistics on the variables used in the empirical analysis. Each variable is in log change and summarized by sectors across microregions and over time. Employment growth was relatively higher in services, followed by manufacturing, and agriculture and mining. These differential job dynamics across sectors is also reflected in the number of entrants and exiters, as well as in the number of job switchers within the labor market—defined as the number of workers who switched either occupation or establishment within the labor market. Changes in residual wages tended to be relatively more similar across sectors, on average. As detailed in the appendix, this variable is first computed at the individual-level, purging wages from the effects of age, gender and education, and then taking the average at the labor market level. Exports originated from firms whose main activity is within tradable sectors shows significant growth, in line with the growth in external demand.

Table 2 reports the first stage estimates relating changes in log export revenue originated from each labor market to changes in external demand directed to the labor market, as defined in equation (2). The econometric results reveal that our instrument provides a suitable source of variation for examining the impact of plausibly export shocks on labor market outcomes. The coefficient of interest is 0.793 indicating that a 10% increase in external demand directed to the

[4] Notice that the change in exports, initial exports per worker and the relative importance of each destination also vary across sectors within each microregion. This heterogeneity is exploited in the analysis, but is not reported in the figures to avoid clutter.

Table 1: Summary statistics, 2004-2011

Variable	Agriculture and mining	Manufacturing	Services	All
	(1)	(2)	(3)	(4)
Δ Employment	0.033	0.053	0.073	0.053
	(0.278)	(0.236)	(0.138)	(0.226)
Δ Residual wage	0.032	0.025	0.025	0.028
	(0.106)	(0.097)	(0.065)	(0.091)
Δ # exiters	0.023	0.053	0.059	0.045
	(0.712)	(0.667)	(0.475)	(0.625)
Δ # entrants	0.038	0.042	0.071	0.051
	(0.719)	(0.745)	(0.493)	(0.660)
Δ # job switchers	0.013	0.025	0.091	0.046
	(0.659)	(0.749)	(0.358)	(0.599)
Δ Export revenues	0.043	0.04	-0.008	0.039
	(10.506)	(0.944)	(20.404)	(10.284)
$\Delta\bar{Z}$	0.096	0.055	0.052	0.072
	(0.240)	(0.150)	(0.293)	(0.202)

Notes: Table reports summary statistics on the unrestricted estimation sample. Means are reported in plain text, and standard deviations are in parentheses.

labor market leads to a 7.7% increase in exports. This relationship is precisely estimated, with a Kleibergen-Paap rk Wald F-stat of 29.87, which is indicative of a strong instrument.

We proceed by examining the causal effects of export shocks on a range of labor market outcomes. Table 3 presents the instrumental-variables estimates of (1), using the strategy discussed above. Panel A measures effects on total employment and residual wages in the labor market. The estimates reveal that a 10% increase in exports leads to a 2.3% increase in employment and 3.1% increase in average residual wages. The estimates in Panel B reveal that export shocks have important implications for gross worker flows across and within labor markets. A positive export shock leads to a reduction in the number of workers leaving the labor market, an increase in the number of workers entering the labor market, and a significant increase in the number of workers switching occupation and/or establishment within the labor market. The last result implies that a positive export shock increases internal churning within the labor market.

3 Model

In this section, we first introduce a dynamic labor mobility model with the endogenous number of job opportunities. Each labor market is defined by a pair of region and sector. Each labor

Table 2: First stage estimates

Dependent variable:	Δ Exports
$\Delta\bar{Z}$	0.806
	(0.100)
Kleibergen-Paap rk Wald F-stat	27.92
Sectors	3
Microregions	558
Labor markets	857
Observations	4008
Year effects	Y

Notes: Table reports first stage estimates. For each dependent and independent variable, we take the log before computing the first differences. Changes in log exports and sectoral export demand are weighted by exports per worker in the labor market in 2003. Standard errors clustered by microregion and year are presented in parentheses.

market offers a different number of job opportunities as well as a different wage. Conditional on wage, the number of job opportunities, and their idiosyncratic preference, workers optimally choose a job which belongs to a certain labor market. Next, we introduce international trade to the framework to show how wage and the number of job opportunities in each labor market are endogenously determined. Changes in labor market condition due to trade shocks endogenously impact the wage level and the number of job opportunities of each labor market. As a result, labor mobility between labor markets is endogenously determined.

3.1 Labor mobility model with endogenous number of jobs

Consider an economy with a continuum L workers. Each worker is in a discrete state $k \in \{1, 2, .., K\}$ which is a region-sector labor market index.[5] The number of workers in labor market k at time t is denoted as L_t^k with $\sum_k L_t^k = L_t$. We denote R_k as the region of labor market k, S_k as the sector of labor market k. The total number of regions in this economy is R, and the total number of sectors is S, both of which we assume to be fixed over time. In most papers in the literature on dynamic discrete choice model of labor mobility, workers are assumed to choose

[5] We denote a labor market with a single index k instead of a pair of region index and sector index. This notation is particularly convenient when we estimate the model with the data, because not all region-sector pairs are populated in the data. The maximum number of labor market we can have is a product of the number of regions and the number of sectors in the data, but the actual K is not necessarily equal to this maximum number.

Table 3: IV estimates on the impact of export shocks on labor markets

A. Dependent variable:	Δ Employment	Δ Residual wage		
	(1)	(2)		
Δ Exports	0.230	0.318		
	(0.039)	(0.034)		
Sectors	3	3		
Microregions	558	558		
Labor markets	857	857		
Observations	4008	4008		
Year effects	Y	Y		
B. Dependent variable:	Δ # leaving (f)	Δ # leaving (r)	Δ # entering	Δ # switching jobs
	(1)	(2)	(3)	(4)
Δ Exports	-0.960	-0.148	0.403	0.460
	(0.134)	(0.056)	(0.080)	(0.101)
Sectors	3	3	3	3
Microregions	558	558	558	558
Labor markets	857	857	857	857
Observations	4008	4008	4008	4008
Year effects	Y	Y	Y	Y

Notes: Table reports IV results of equation (1) in text, using the baseline estimation sample. For all dependent and independent variables, we take the log before computing the first differences. Changes in log exports are instrumented by changes in sectoral import demand. All explanatory variables are weighted by exports per worker in the labor market in 2003. Standard errors clustered by microregion and year are in parentheses.

a labor market directly. On the other hand, we model worker's problem as a choice of a *job*. There are many job opportunities workers can compare in each labor market, and workers choose one job in every period. Which labor market a worker belongs to is an outcome of her optimal choice of a job. We denote the labor market which job j belongs to by $k = K_j$. In addition to that this assumption is much more realistic, it introduces an important dimension through which labor markets are affected by aggregate shocks. The status of an economy impacts the number of job opportunities from which workers can choose. Depending on the nature of the shock and the differential exposure of each labor market to the shock, the impact of trade shocks on the number of jobs will be different between labor markets.

Instead of choosing a job from job opportunities available in each labor market, workers have an option to move to the residual labor market. Empirically, unemployment, home employment, and working in the informal labor market all belong to this choice. Each worker compares the option to move to the residual labor market with all the other job opportunities available in each labor market. We denote the choice to be in the residual labor market by $j = I$. The residual labor market is assumed to offer only one job opportunity, and thus we do not need to distinguish a job from a labor market for the residual labor market. Therefore, we can express $K_I = I$ without loss of generality.

All jobs within labor market k are identical apart from the iid utility shock associated with them. We denote the real wage in labor market K_j at time t by $w_t^{K_j}$. If the current job j is I, then the real wage is w_t^I. Specifically, an agent who is indexed with h and who is attached to job j within labor market K_j receives instantaneous utility u_t^h at time t defined as

$$u_t^h = w_t^{K_j} + \varepsilon_t^{h,j}, \tag{3}$$

where $\varepsilon_t^{h,j}$ is distributed Gumbel with mean κ and the scale parameter $\nu > 0$.[6]

Workers start period t attached to a job j, and receive $w_t^{K_j}$ in the beginning of the period. After workers receive their wages, they sample some jobs from each labor market and learns about the iid shocks associated with each of those jobs and the residual labor market. Conditional on this idiosyncratic draw, each worker chooses the best available job and receives the iid shock at the end of the current period. Then, the period t ends.

In each period, a worker can sample N_t^k jobs from labor market k. The number of jobs that can be sampled increases with the number of total jobs in the labor market. This relationship is determined at the general equilibrium which we will discuss after characterizing international trade. We assume that workers choose a job within a region-sector jointly with full information about the idiosyncratic component of their utility from a given job across all labor markets. In addition, they have full information about the idiosyncratic preference of the current period for the residual labor market. They do not know the exact future values of the idiosyncratic component, but they form rational expectations. Therefore, each worker compares the expected utility from the current job and $\sum_k N_t^k + 1$ potential jobs which include the option to move to the residual labor market with the expected utility from the current job at every period.[7]

[6]We can assume a log wage in the utility function as well, and this alternative specification does not change our main results. We show the results based on a log utility in robustness check.

[7]We effectively assume that $N_t^I = 1$ in every t.

Switching a job and switching a labor market are both subject to frictions. When a worker moves from a job j to a different job j' at time t, she pays the moving cost δ_t. This cost incurs even when the job switching is within the same labor market. If the job switching involves a switch to a different labor market, i.e., $K_j \neq K_{j'}$, then she pays an additional moving cost $C_t(K_j, K_{j'}) \geq 0$.[8] Since the option to be in the residual labor market is compared with job opportunities in the formal labor market, we assume that switching between the residual labor market and any other job in the formal labor market is subject to both switching costs. In other words, if a worker moves from a job j in a labor market K_j to the residual labor market, the total moving cost is $\delta_t + C_t(K_j, I)$, and it is $\delta_t + C_t(I, K_j)$ for the opposite direction.

Based on the moving cost structure, we derive the labor-market-specific present discounted utility of the agent h with job j in the beginning of t as

$$U_t^{j,h} = w_t^{K_j} + \max_{j'} \left\{ \beta E_t V_{t+1}^{K_{j'}} - \left(C_t(K_j, K_{j'}) + \delta_t \right) \mathbf{1}(j \neq j') + \varepsilon_t^{j',h} \right\}. \tag{4}$$

By taking an expectation over the idiosyncratic component, we define the expected present discounted value in the beginning of time t as

$$\begin{aligned} V_t^{K_j} &= E_\varepsilon U_t^{j,h}, \\ &= w_t^{K_j} + E_\varepsilon \max_{j'} \left\{ \beta E_t V_{t+1}^{K_{j'}} - \left(C(K_j, K_{j'}) + \delta_t \right) \mathbf{1}(j \neq j') + \varepsilon_t^{j',h} \right\}. \end{aligned}$$

3.2 Equilibrium labor mobility and option values

Using the assumption of Gumbel distribution for the idiosyncratic shock, the probability that a worker moves from a labor market k to a labor market l is derived as

$$\begin{aligned} m_t^{kl} &= \frac{\mathbf{1}_{l=k} \exp\left(\beta E_t V_{t+1}^l / \nu\right) + N_t^l \exp\left(\frac{E_t \beta V_{t+1}^l - \delta_t - C_t(k,l)}{\nu}\right)}{\exp\left(\frac{\beta}{\nu} E_t V_{t+1}^k\right) + N_t^k \exp\left(\frac{E_t \beta V_{t+1}^k - \delta_t}{\nu}\right) + \sum_{l' \neq k} N_t^{l'} \exp\left(\frac{\beta E_t V_{t+1}^{l'} - \delta_t - C_t(k,l')}{\nu}\right) + \exp\left(\frac{\beta E_t V_{t+1}^I - \delta_t - C_t(k,I)}{\nu}\right)} \\ &= \frac{\mathbf{1}_{l=k} \lambda_{0,t}^k + \mathbf{1}_{l \neq k \wedge l \neq I} \lambda_{1,t}^l \exp\left(-\frac{C_t(k,l)}{\nu}\right) + \mathbf{1}_{l=I} \lambda_{I,t}^k}{\lambda_{0,t}^k + \lambda_{1,t}^k + \lambda_{2,t}^k + \lambda_{I,t}^k}, \end{aligned} \tag{5}$$

[8] Effectively, we assume $C_t(k,k) = 0$.

where we define

$$
\begin{aligned}
\lambda_{0,t}^{k} &\equiv \exp\left(\frac{\beta}{\nu} E_t V_{t+1}^{k}\right) \\
\lambda_{1,t}^{k} &\equiv N_t^{k} \exp\left(\frac{E_t \beta V_{t+1}^{k} - \delta_t}{\nu}\right) \\
\lambda_{2,t}^{k} &\equiv \sum_{l' \neq k} N_t^{l'} \exp\left(\frac{\beta E_t V_{t+1}^{l'} - \delta_t - C_t(k, l')}{\nu}\right) \\
\lambda_{I,t}^{k} &\equiv \exp\left(\frac{\beta E_t V_{t+1}^{I} - \delta_t - C_t(k, I)}{\nu}\right)
\end{aligned}
$$

for notational simplicity. From this probability of switching a labor market, the role of the number of jobs across labor markets becomes clear. Workers are more likely to move to a labor market in which they can sample more jobs, i.e., a larger N_t^l, conditional on the expected net value of a labor market. Since workers will choose the job which gives them the highest present discounted value, the expectation of the maximum of the idiosyncratic component should increase in the number of jobs that workers can sample, which makes it more likely for workers to move into a labor market with more jobs. Similarly, if formal labor markets offer more job opportunities, then it becomes relatively less likely for workers to move to the residual labor market.

For notational convenience, we define additional probabilities of switching. First, we denote the probability of moving from labor market k to labor market l conditional on changing jobs but staying in a formal labor market by $\widetilde{m}_t^{K_j K_{j'}}$ for $j \neq j'$.[9] In addition, we denote the probability of staying in the same job, thus in the same labor market k, by $\mu_{0,t}^k$, the probability of changing jobs but staying in the same labor market k by $\mu_{1,t}^k$, and the probability of changing labor markets from k to any other labor market $l \neq k$ (thus also changing jobs) by $\mu_{2,t}^k$. Finally, we denote the probability of moving from labor market k to the residual labor market by $\mu_{I,t}^k$, which is effectively equal to m_t^{kI}.[10] The job opportunity channel that we introduce into the model will allow for internal churning between jobs, which is measured by $\mu_{1,t}^k$. Since we assume that the

[9] More formally, this conditional probability for $l \neq I$ is

$$
\widetilde{m}_t^{kl} = \frac{\mathbf{1}_{l \neq k \wedge l \neq I} \lambda_{1,t}^{l} \exp\left(-\frac{C_t(k,l)}{\nu}\right)}{\lambda_{1,t}^{k} + \lambda_{2,t}^{k}},
$$

where $\lambda_{1,t}^l$, $\lambda_{1,t}^k$, and $\lambda_{2,t}^k$ are as defined above.

[10] By construction, $\mu_{0,t}^k + \mu_{1,t}^k + \mu_{2,t}^k + \mu_{I,t}^k = 1$ should hold for any t and j. Using the definition of $\lambda_{0,t}^k$, $\lambda_{1,t}^k$, and $\lambda_{2,t}^k$, each probability can be expressed as $\mu_{0,t}^k = \frac{\lambda_{0,t}^k}{\lambda_{0,t}^k + \lambda_{1,t}^k + \lambda_{2,t}^k + \lambda_{I,t}^k}$, $\mu_{1,t}^k = \frac{\lambda_{1,t}^k}{\lambda_{0,t}^k + \lambda_{1,t}^k + \lambda_{2,t}^k + \lambda_{I,t}^k}$, $\mu_{2,t}^k = \frac{\lambda_{2,t}^k}{\lambda_{0,t}^k + \lambda_{1,t}^k + \lambda_{2,t}^k + \lambda_{I,t}^k}$, and $\mu_{I,t}^k = \frac{\lambda_{I,t}^k}{\lambda_{0,t}^k + \lambda_{1,t}^k + \lambda_{2,t}^k + \lambda_{I,t}^k}$.

residual labor market I is degenerate with a single job opportunity I, $\mu^I_{0,t} = \mu^I_{I,t}$. With the new notations, we can re-write the labor-market-specific value as

$$V^k_t = w^k_t + \beta E_t V^k_{t+1} - \nu \log(\mu^k_{0,t}), \tag{6}$$

where $-\nu \log(\mu^k_{0,t})$ is an option value of moving. This option value can be decomposed into internal and external option values. The external option value is defined as the option value from alternative jobs in a different labor market. In order to net out the effect from switching to a different job in the same labor market, we need to divide $\mu^k_{0,t}$ by $\mu^k_{0,t} + \mu^k_{2,t} + \mu^k_{I,t}$ in the option value term, which gives the external option value term as

$$-\nu \log(\mu^k_{0,t}) + \nu \log(\mu^k_{0,t} + \mu^k_{2,t} + \mu^k_{I,t}). \tag{7}$$

The internal option value is the value from alternative job opportunities within the labor market, which is given as the difference between the total option value and the external option value, i.e.,

$$-\nu \log(\mu^k_{0,t} + \mu^k_{2,t} + \mu^k_{I,t}). \tag{8}$$

3.3 Relative welfare and the number of jobs

The model delivers simple formulae for changes in the relative welfare between workers in two different labor markets and for changes in the number of jobs in each labor market in response to the change in an underlying policy variable. First, we assume that a change in a policy variable x does not change the moving cost.

If we further assume that workers receive their wage in the beginning of time t before any change in policy parameter x is realized, then the change in the relative welfare of workers in labor market k compared to workers in labor market l is equal to

$$\Delta_x \left(V^k_t - V^l_t\right) = \Delta_x \left[\nu \left(\log \widetilde{m}^{lk}_t - \log \widetilde{m}^{kk}_t\right) - \nu \left(\log\left(1 - \mu^k_{0,t} - \mu^k_{I,t}\right) - \log\left(1 - \mu^l_{0,t} - \mu^l_{I,t}\right)\right)\right], \tag{9}$$

where Δ_x denotes the change induced by a change of x.

In the next section, we estimate changes in $\widetilde{m}^{lk}_t$ due to trade shocks using variations explored in the local labor market approach. This result also suggests that the change in the relative welfare due to a policy change depends on the relative wage and the relative number of job opportunities between labor markets, but their impact on the welfare change is entirely captured by the change in switching probabilities.

Next, our model shows that the change in the number of job opportunities in each labor market driven by a change in x can be written as

$$\Delta_x \log N_t^k = \Delta_x \left(\nu \log \mu_{1,t}^k - \nu \log \mu_{0,t}^k \right). \tag{10}$$

Equations (9) and (10) suggest that moving probabilities and the labor supply elasticity ν are the sufficient statistics for changes in welfare and changes in the number of job opportunities.

3.4 International trade with love for variety of jobs

The dynamic structural labor mobility model that we introduce can be used to quantify the effect of various labor demand shocks on worker welfare. In this paper we focus particularly on the effect of international trade. Different wages and different numbers of job opportunities across labor markets are two key driving forces which generate labor mobility in our model. In order to characterize how labor mobility is affected by trade shocks, we introduce international trade to our model, where trade shocks endogenously affect both wage and the number of job opportunities at the labor market level.

We characterize international trade based on the Eaton and Kortum (2002) model. This assumption will make international trade driven only from the standard Ricardian force. We assume that there are N countries ($n = 1, \ldots, N$), but only country 1 is populated with more than one regions. For other countries $n \neq 1$, there is only one region which is the country itself. In other words, $R > 1$ for country 1 and $R = 1$ for all other countries. In order to characterize the country-level trade while we still have multiple regions in the country that we are interested in, we assume that there is a national aggregator for each sector in that particular country of interest. The aggregator of each sector sources each variety within that sector from the lowest cost region in the country at no trade cost. This way the national price level of each product is determined regardless of the consumption location. Then this aggregator trades each product with partner countries. If no region in country 1 is the lowest cost supplier of a certain product for consumers and producers of country 1, then country 1 imports that product, and no region in this country produces that product at the equilibrium. In order to simplify our analysis, we assume no production in the residual labor market of country 1 and further assume that there is only formal sector in other countries $n' \neq 1$.

Labor, fixed factor, and composite intermediate inputs are used for production of each product variety ω. We assume that there is a continuum of products in $[0, 1]$, and that each product variety is traded between countries. Products can be used either as intermediate inputs or for final consumption by consumers. We assume a Cobb-Douglas production function, where the production technology for a variety ω in labor market k of country 1 at time t is:

$$Q_{1,t}^k(\omega) = z_1^k(\omega)(\widetilde{l}_{1,t}^k)^{\gamma_l}(M_{1,t}^k)^{\gamma_m}(B_1^k)^{1-\gamma_l-\gamma_m},$$

where $M_{1,t}^k$ is a composite intermediate input and B_1^k is a fixed factor used in the labor market k of country 1. We assume that this fixed factor is specific to a labor market and does not change over time. To simplify the analysis, we do not impose an explicit input-output structure between sectors. Instead, each sector and each region considers the aggregate price index $P_{1,t}$ as the price of composite intermediate inputs in country 1 at time t. As a result, the share of a particular sector in the composite intermediate inputs is the same across all sectors that demand intermediates. This assumption is to simplify the production and trade structure and focus on the worker side.

In the production function, $\tilde{l}_{1,t}^k$ is in terms of the number of efficiency units provided by all workers in labor market k of country 1 at time t. This variable is different from $L_{1,t}^k$ which is the actual number of workers in labor market k of country 1 at time t. Here we introduce a notion of task which produces labor efficiency units. Producers within each labor market k will decide how to allocate the total labor force into tasks. Empirically, this allocation can be across differentiated occupations, different establishments, or both. We assume that workers are equally productive regardless of the task they are assigned, as long as they are in the same labor market. Therefore, producers will simply choose the total mass of a continuum of tasks to operate as well as the total efficiency units of labor they use. For each labor market k, $\tilde{l}_{1,t}^k$ is a CES aggregate of all efficiency units provided by each task in labor market k. We define a continuum of tasks operated in labor market k as Ω_t^k and denote the mass of Ω_t^k as O_t^k. Then, the total labor aggregate is

$$\widetilde{l}_{1,t}^k = \left[\int_{\tau \in \Omega_t^k} (l_{1,t}^k(\tau))^{\frac{\tilde{\sigma}-1}{\tilde{\sigma}}} d\tau\right]^{\frac{\tilde{\sigma}}{\tilde{\sigma}-1}},$$

where $l_{1,t}^{k}(\tau)$ is the total efficiency units provided by task τ in labor market k of country 1; and $\tilde{\sigma} > 1$ is the elasticity of substitution between tasks. In addition, we assume that producers have to pay an additional cost $\tilde{c}(O_t^k)^{\tilde{\alpha}}$ in order to operate a continuum of tasks with a mass O_t^k, where $\tilde{\alpha}$ is the curvature in the cost of operating tasks. Workers are equally productive at each task thus are paid the same wage regardless of assignment into a task or choice of a job. This assumption is the key difference from the standard search model. Under this assumption, $\widetilde{l}_{1,t}^{k}$ can be rewritten as $\widetilde{l}_{1,t}^{k} = L_{1,t}^{k}(O_t^k)^{\frac{1}{\tilde{\sigma}-1}}$, where $L_{1,t}^{k}$ is the total number of workers in labor market k at time t.[11]

If $\tilde{\sigma} > 1$ as we assumed, then there are more tasks in a cell if the optimal demand for total labor force $\widetilde{l}_{1,t}^{k}$ is larger, given the labor supply $L_{1,t}^{k}$.[12] This is the love for variety of tasks channel, which is analogous to the familiar love for variety of products in the Armington trade model. The degree of love for variety of tasks depends on the elasticity of substitution between tasks, $\tilde{\sigma}$. Therefore, if there is an exogenous shock which expands the labor market k, then the total labor demand in k will increase and thus the number of tasks operated in that labor market also will increase. Empirically, producers may post new type of tasks to operate a more differentiated task structure or open new establishments to meet the increased demand.

In order to link this number of tasks to the number of job opportunities that workers perceive when they solve their problem of job choice, we assume that

$$N_t^l \quad = \quad \rho\left(O_t^l\right),$$

where N_t^l is a positive integer.[13] We assume that $\rho(.)$ is a monotonically increasing mapping in O_t^l. In other words, the number of job opportunities perceived by workers in a certain labor market is assumed to be increasing in the number of tasks operated in the labor market.

We assume that the wage in the residual labor market is a fraction of the average wage of all formal labor markets, i.e., $w_{1,t}^{I} = \frac{\eta}{K}\sum_{k=1}^{K} w_{1,t}^{k}$, where $0 < \eta < 1$. In addition to the wage-induced labor mobility, the effect of a labor demand shock on labor mobility through the number of job opportunities is two-fold. First, a positive labor demand shock to a certain labor market increases the number of tasks operated by producers through the love for variety of tasks channel and thus increases the number of job opportunities that workers can compare. As a result, a labor market with a positive labor demand shock attracts more workers by providing

[11] The assumption of the same productivity across tasks makes it optimal for producers to evenly allocate labor forces into different tasks, i.e., $l_{1,t}^{k}(\tau) = L_{1,t}^{k}/O_t^k$ for all $\tau \in \Omega_t^k$.

[12] The actual value of $\tilde{\sigma}$ is to be estimated in the next section.

[13] We will define the sampling function $\rho(.)$ formally in the appendix. We can think of it as a step function mapping positive real numbers to positive integers.

more opportunities. Second, a positive labor demand shock in a non-residual labor market will decrease the probability of moving to the residual labor market. This second channel is similar to the standard search model, where the effect of a macro shock on welfare operates only through unemployment margin. In our model, a labor demand shock affects labor mobility and thus workers' welfare not only through the transition between formal and residual labor markets but also through higher utility from being able to compare more opportunities to choose the best one.

Other countries $n \neq 1$ have a simpler production function using only aggregate labor and composite intermediate inputs for each labor market k:

$$Q^k_{n,t}(\omega) = z^k_n(\omega)(\bar{L}^k_{n,t})^{\bar{\gamma}_l}(M^k_{n,t})^{1-\bar{\gamma}_l}.$$

For other countries $n \neq 1$, they all have the same number of sectors $\bar{S}$ such that $\bar{S}$ is the number of unique S_k over all k in country 1, so the sector-level trade flows between country 1 and $n \neq 1$ are well-defined. In addition, for all k in country $n \neq 1$, $R_k = 1$, as there is only one region in each country. Therefore, the superscript k effectively denotes a sector for country $n \neq 1$. At the general equilibrium, $\sum_k^{\bar{S}} \bar{L}^k_{n,t} = \bar{L}_n$ should hold for every (n, t), where $\bar{L}_n$ is the total labor endowment of country n which is assumed to be time-invariant. $M^k_{n,t}$ is the composite intermediate inputs demanded by sector k of country n at time t. $\bar{\gamma}_l$ is the value-added share which is assumed to be the same across all countries $n \neq 1$.

Factor-neutral productivity for each product variety is randomly drawn from a Fréchet distribution. We assume that the factor-neutral productivity does not vary over time. For each country n, $z^k_n(\omega)$ is randomly drawn from

$$F^k_n(z) = \exp(-T^k_n z^{-\theta}).$$

As in the EK model, we assume that the productivity draws are independent across countries and labor markets. Since the labor market is defined in a different way between country 1 and country $n \neq 1$, we impose a further structure on the scale parameter. We assume that $T^k_1 \equiv T^{S_k}_1 T^{R_k}_1$, where $T^{S_k}_1$ determines country 1's overall productivity level in sector S_k and $T^{R_k}_1$ is for region-specific productivity.

We assume iceberg trade costs between countries, $d^s_{nn',t}$, for products of sector s shipped from country n to n' at time t. If $n = 1$ or $n' = 1$, then this is trade cost between the national aggregator in country 1 and its partner country. We assume that there is no trade cost between regions within country 1.

3.5 Equilibrium trade flows and price indices

In our model, producers decide not only how much of each factor to employ but also how to allocate total labor force into different tasks. For the latter, producers have to decide the optimal number of tasks. There is a clear trade-off of having more tasks, i.e., more specialized labor technology. Having more tasks is beneficial because it increases the total efficiency units of labor conditional on $L^k_{1,t}$. However, to have more diversified production technology, producers have to pay a higher cost for such as training or building a new establishment. We assume that producers have to pay marginal cost of $\tilde{c} > 0$ per each additional task τ and that this cost is the same across all labor markets.

The Cobb-Douglas production function gives us the following unit cost function for all firms in labor market k of country 1 at time t,

$$c^k_{1,t} = \Upsilon_1(\widetilde{w}^k_{1,t}\tilde{c}^{\frac{1}{\tilde{\sigma}-1}})^{\gamma_l}(P_{1,t})^{\gamma_m}(b^k_{1,t})^{1-\gamma_l-\gamma_m}, \tag{11}$$

where $\widetilde{w}^k_{1,t}$ is the nominal wage of workers in labor market k of country 1 at time t, $P_{1,t}$ is the aggregate price index in country 1 at time t, and $b^k_{1,t}$ is the price of the fixed factor for labor market k of country 1 at time t.[14] Υ_1 is Cobb-Douglas constant which is a function of γ_l and γ_m. Similarly, the unit cost function for all producers in country $n' \neq 1$ at time t is

$$c_{n',t} = \Upsilon_{n'}(\widetilde{w}_{n',t})^{\bar{\gamma}_l}(P_{n',t})^{1-\bar{\gamma}_l}, \tag{12}$$

with the assumption of perfect labor mobility between sectors for all countries $n' \neq 1$.

At the equilibrium under perfect competition as in EK, bilateral trade share in sector s between regions r and r' of country 1 at time t is determined by[15]

$$\lambda^s_{(1,r),(1,r'),t} = \frac{T^{(r,s)}_{1,t}(c^{(r,s)}_{1,t})^{-\theta}}{\sum_{r''}T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t})^{-\theta} + \sum_{n'\neq 1}T^s_{n',t}(c_{n',t}d^s_{n'1,t})^{-\theta}} = \frac{X^s_{(1,r),(1,r'),t}}{X^s_{(1,r'),t}}. \tag{13}$$

From the assumption of no trade cost between regions in country 1, $\lambda^s_{(1,r),(1,r'),t}$ is equalized across all r' of country 1. Similarly, the equilibrium trade flow of sector s from country $n \neq 1$ to region

[14] Note that the wage is per worker, not per efficiency unit.

[15] In order to make a clear distinction between region and sector in trade flow equations, each labor market k is denoted as (r, s).

r of country 1 at time t is determined by

$$\lambda^s_{n,(1,r),t} = \frac{T^s_{n,t}(c_{n,t}d^s_{n1,t})^{-\theta}}{\sum_{r''} T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t})^{-\theta} + \sum_{n'\neq 1} T^s_{n',t}(c_{n',t}d^s_{n'1,t})^{-\theta}} = \frac{X^s_{n,(1,r),t}}{X^s_{(1,r),t}}. \tag{14}$$

Since all regions take the sector-level price index as given for the use of intermediate inputs and consumers have the identical preference regardless where they live or in which sector they work, incoming trade flows should be the same across all regions. The actual demand level between regions depends only on their real income which is allowed to be different across regions in our setting. As a result, we have $\lambda^s_{n,1,t} = \bar{R}\lambda^s_{n,(1,r),t}$ for sector-level trade shares from country $n \neq 1$ to country 1 in aggregate.

The reverse trade flow from region r of country 1 to a country $n \neq 1$ is

$$\lambda^s_{(1,r),n,t} = \frac{T^{(r,s)}_{1,t}(c^{(r,s)}_{1,t}d^s_{1n,t})^{-\theta}}{\sum_{r'} T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t}d^s_{1n,t})^{-\theta} + \sum_{n'\neq 1} T^s_{n',t}(c_{n',t}d^s_{n'n,t})^{-\theta}} = \frac{X^s_{(1,r),n,t}}{X^s_{n,t}}. \tag{15}$$

Finally, trade flows between countries $n \neq 1$ and $n'' \neq 1$ are derived as

$$\lambda^s_{n,n'',t} = \frac{T^s_{n,t}(c_{n,t}d^s_{nn'',t})^{-\theta}}{\sum_{r'} T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t}d^s_{1n'',t})^{-\theta} + \sum_{n'\neq 1} T^s_{n',t}(c_{n',t}d^s_{n'n'',t})^{-\theta}} = \frac{X^s_{n,n'',t}}{X^s_{n'',t}}. \tag{16}$$

The exact price index for sector s in country 1 at time t is

$$P^s_{1,t} = \bar{\Gamma}\left[\sum_{r''} T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t})^{-\theta} + \sum_{n'\neq 1} T^s_{n',t}(c_{n',t}d^s_{n'1,t})^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{17}$$

and the exact price index for sector s in country $n \neq 1$ at time t is

$$P^i_{n,t} = \bar{\Gamma}\left[\sum_{r'} T^{(r'',s)}_{1,t}(c^{(r'',s)}_{1,t}d^s_{1n,t})^{-\theta} + \sum_{n'\neq 1} T^s_{n',t}(c_{n',t}d^s_{n'n,t})^{-\theta}\right]^{-\frac{1}{\theta}}, \tag{18}$$

where $\bar{\Gamma}$ is the same constant for both cases.[16] We assume that all consumers have an identical nested CES preference with a common elasticity of substitution σ at the variety level and a Cobb-Douglas aggregation across sectors with expenditure share ϕ^s. Then, the aggregate price index is given by $P_{n,t} = \prod_s (\frac{P^s_{n,t}}{\phi^s})^{\phi^s}$.

[16] More precisely, $\bar{\Gamma} \equiv \left[\Gamma(\frac{\theta+1-\sigma}{\theta})\right]^{1/(1-\sigma)}$. We assume $\sigma < \theta + 1$ so that the price index is well-defined.

3.6 Market clearing

To simplify the general equilibrium, we assume that country-level trade deficit $D_{n,t}$ is exogenously fixed as the share of world GDP. We further assume that the country 1's trade deficit is distributed across consumers in different labor markets of country 1 proportionally to the income share. The trade deficit distributed to labor market k is denoted by $D^k_{1,t}$. The total expenditure on sectors s products by all agents in region r of country 1 is

$$X^s_{(1,r),t} = \phi^s \gamma_m \sum_{s'} \left(\sum_{r'} \lambda^{s'}_{(1,r),(1,r'),t} X^{s'}_{(1,r'),t} + \sum_{n' \neq 1} \lambda^{s'}_{(1,r),n',t} X^{s'}_{n',t} \right) + \phi^s \left(\sum_{k \in \{k | R_k = r\}} \left(\widetilde{w}^k_{1,t} L^k_{1,t} + D^k_{1,t} \right) \right), \tag{19}$$

where $\{k \mid R_k = r\}$ is the set of labor markets within region r. Note that we can denote $X^i_{1,t} \equiv \sum_r X^i_{(1,r),t}$. Similarly, the total expenditure on sector s products by all agents in country $n \neq 1$ at time t is

$$X^s_{n,t} = \phi^s (1 - \bar{\gamma}_l) \sum_{s'} \left(\sum_{r'} \lambda^{s'}_{n,(1,r'),t} X^{s'}_{(1,r'),t} + \sum_{n' \neq 1} \lambda^{s'}_{n,n',t} X^{s'}_{n',t} \right) + \phi^s \left(\widetilde{w}_{n,t} \bar{L}_{n,t} + D_{n,t} \right). \tag{20}$$

Finally, the set of market clearing conditions is given by

$$\widetilde{w}^k_{1,t} L^k_{1,t} = \gamma_l \left(\sum_{r'} \lambda^{S_k}_{(1,R_k),(1,r'),t} X^{S_k}_{(1,r'),t} + \sum_{n' \neq 1} \lambda^{S_k}_{(1,R_k),n',t} X^{S_k}_{n',t} \right), \tag{21}$$

$$b^k_{1,t} B^k_1 = (1 - \gamma_l - \gamma_m) \left(\sum_{r'} \lambda^{S_k}_{(1,R_k),(1,r'),t} X^{S_k}_{(1,r'),t} + \sum_{n' \neq 1} \lambda^{S_k}_{(1,R_k),n',t} X^{S_k}_{n',t} \right) \tag{22}$$

for each labor market k of country 1, and

$$\widetilde{w}_{n,t} \bar{L}_{n,t} = \bar{\gamma}_l \sum_{s'} \left(\sum_{r'} \lambda^{s'}_{n,(1,r'),t} X^{s'}_{(1,r'),t} + \sum_{n' \neq 1} \lambda^{s'}_{n,n',t} X^{s'}_{n',t} \right), \tag{23}$$

for country $n' \neq 1$. The equilibrium labor supply for each labor market of country 1 $L^k_{1,t}$ is pinned down by the labor model. We further normalize $\widetilde{c} = 1$ to write O^k_t as a function of γ_l, $\tilde{\sigma}$, and the equilibrium output from the producer's first order condition. We also normalize the average wage to be equal to 1 at every period.

To complete the model, we now characterize the residual labor market further. Workers in the residual labor market receive a fixed wage η in every period as long as they stay in the residual sector. A worker who was in the residual labor market but will be in a formal labor market in the next period can draw iid shocks for available jobs and choose a job like the incumbent formal workers.

A temporary equilibrium is a vector of wages

$$\widetilde{\mathbf{w}}_t = (\widetilde{w}^1_{1,t}, \ldots, \widetilde{w}^K_{1,t}, \widetilde{w}_{2,t}, \ldots, \widetilde{w}_{N,t}),$$

which satisfies (11)-(23), conditional on the labor supply

$$\mathbf{L}_t = (L^1_{1,t}, \ldots, L^K_{1,t}, \bar{L}_2, \ldots, \bar{L}_N),$$

and other fundamental parameters. A sequential competitive equilibrium is the sequence of $\mathbf{L}_t$, $\widetilde{\mathbf{w}}_t$, and $\mathbf{m}_t = \left\{m^{kl}_{1,t}\right\}^{\infty,\infty}_{k=1,l=1}$ which solve the labor mobility model at each time period t, conditional on the initial labor allocation $\mathbf{L}_0$ and fundamental parameters of the model.

4 Estimation

We estimate our structural model with the sample described in Section 2. In addition to the sample selection rules discussed previously, we restrict the sample to the labor markets where at least 100 workers move in and out respectively in every time period. Since the identification of moving probabilities are based on the worker mobility, values of labor markets with little labor mobility cannot be identified. Unlike Artuç, Chaudhuri, and McLaren (2010) or Caliendo, Dvorkin, and Parro (forthcoming), we allow corridors with zero mobility. For example, if there are less than 100 workers move out of (or move into) labor market A, we drop labor market A. However, if there are zero workers moving from labor market B to labor market C, but there are more than 100 workers total moving into other cells from labor market B, then we keep labor market B and also keep the B-C corridor in the estimation. The idea is similar to dropping small countries in the gravity estimation of trade flows, where it is common to keep corridors with zero flows. After this restriction, we end up with 857 labor markets.[17]

The main objects to be estimated are the moving probabilities and the structural parameters such as moving costs, the labor supply elasticity ν, and the elasticity of substitution between

[17]All main results that we present in this paper are robust to the mobility cutoff. This number is close to the maximum number of labor markets we can consider due to computational limitations.

tasks, $\tilde{\sigma}$. We combine the local labor market approach to estimate our structural parameters. With these estimates in hand, we then turn to our variables of interest: changes in welfare, wage, option values, and the number of job opportunities due to trade shocks.

4.1 Estimation of the moving probabilities

As we showed in Section 3, our welfare results depend on the following three factors: the probability of switching a labor market conditional on changing a job and staying in a formal labor market ($\widetilde{m}_t^{kl}$), the probability of staying in the current job ($\mu_{0,t}^k$), the probability of moving to the residual labor market ($\mu_{I,t}^k$), and the labor supply elasticity (ν). In practice, it is not possible to get $\log \widetilde{m}_t^{kl}$ directly from data without any estimation. The bin-estimator, i.e. imputing the probability of moving by dividing the number of switchers by the total number of workers, only works with a large sample size and a small number of choices. We have 857 labor markets, and each labor market offers more than one job opportunities, which makes the bin-estimator unfeasible. This problem is more serious for the probability of switching than for the probability of staying. Instead of using the simple bin-estimator, we can estimate $\log \widetilde{m}_t^{kl}$ by imposing a structure on the moving costs.

The probability of moving from a labor market k to a labor market l conditional on changing jobs is equal to

$$\log \widetilde{m}_t^{kl} = \widetilde{V}_t^l - \widetilde{C}_t(k,l) + \widetilde{\Gamma}_t^k - \log \widetilde{L}_t^k, \tag{24}$$

where

$$\widetilde{V}_t^l = E_t \frac{\beta}{\nu} V_{t+1}^l - \log \mu_{0,t}^l + \log \mu_{1,t}^l \tag{25}$$

$$\widetilde{C}_t(k,l) = \frac{C_t(k,l)}{\nu} \tag{26}$$

$$\widetilde{\Gamma}_t^k = -\log \sum_{l'} \exp\left(\widetilde{V}_t^{l'} - \widetilde{C}_t(k,l')\right) + \log \widetilde{L}_t^k, \tag{27}$$

and $\widetilde{L}_t^k$ is the number of workers who change jobs within labor market k or move out of the labor market k but stay in a formal labor market. After estimating $\widetilde{C}_t(k,l)$ and $\widetilde{V}_t^l$, it is straightforward to calculate the implied probabilities using (24). If we define the number of workers observed in the sample moving from a labor market k to l conditional on changing jobs by $\widetilde{y}_t^{kl}$, then we have $\widetilde{L}_t^k = \sum_l \widetilde{y}_t^{kl}$. This result means that the likelihood function is equal to

$$\mathcal{L} = \prod_k \prod_l \left(\widetilde{m}_t^{kl}\right)^{\widetilde{y}_t^{kl}}, \tag{28}$$

or, alternatively,

$$\log \mathcal{L} = \sum_k \sum_l \widetilde{y}_t^{kl} \left[\widetilde{\Gamma}_t^k + \widetilde{V}_t^l - \widetilde{C}_t(k,l) - \log(\widetilde{L}_t^k)\right], \tag{29}$$

using (24). As discussed above, if $L_t^k \to \infty$, then $\widetilde{L}_t^k \to L_t^k(\mu_{1,t}^k + \mu_{2,t}^k)$, and $\widetilde{y}^{kl}/\widetilde{L}_t^k \to \widetilde{m}_t^{kl}$. Therefore, as the sample size goes to infinity, the maximum likelihood (ML) estimator becomes equivalent to the bin-estimator. We use the Poisson pseudo-maximum-likelihood (PPML) method to estimate $\widetilde{\Gamma}_t^k$, $\widetilde{V}_t^l$, and $\widetilde{C}_t(k,l)$ for each period, because we can write $\widetilde{y}_t^{kl}$ as

$$\widetilde{y}_t^{kl} = \exp\left(\widetilde{\Gamma}_t^k + \widetilde{V}_t^l - \widetilde{C}_t(k,l)\right) + \epsilon_t^{kl}, \tag{30}$$

where ϵ is a sampling error; $\widetilde{\Gamma}_t^k$ is the origin fixed effect; $\widetilde{V}_t^l$ is the destination fixed effect; and $\widetilde{C}_t(k,l)$ is the moving cost between labor markets. The expected number of workers who move from labor market k to l is equal to $E_t\widetilde{y}_t^{kl} = \widetilde{m}_t^{kl}\left(\mu_{1,t}^k + \mu_{2,t}^k\right) L_t^{s,k}$.

Guimaraes, Figueirdo, and Woodward (2003) prove that the maximum likelihood estimation of gravity equation based on Fréchet distribution is identical to PPML. The same intuition applies to our model for Gumbel distribution as well. In the appendix, we prove that the PPML and MLE are equivalent herein.

In the estimation, we consider a simple moving cost structure as follows:

$$\widetilde{C}_t(j,k) = \widetilde{c}_{1,t}D^{jk} + \widetilde{c}_{2,t}\mathbf{1}_{S_j \neq S_k} + \widetilde{c}_{3,t}\mathbf{1}_{S_j \neq S_k \wedge R_j \neq R_k}, \tag{31}$$

where D^{jk} is the log of distance between labor markets j and k, and $\mathbf{1}_{S_j \neq S_k}$ is an indicator function that is equal to one if labor markets j and k are associated with different sectors, and $\mathbf{1}_{S_j \neq S_k \wedge R_j \neq R_k}$ is an indicator function that is equal to one if k and j are associated with different

Table 4: Estimated moving cost parameters

Year	$\widetilde{c}_{1,t}$ (log distance)	s.e.	$\widetilde{c}_{2,t}$ (sector)	s.e.	$\widetilde{c}_{3,t}$ (both)	s.e.
2003	1.0775	(0.0008)	1.9983	(0.0048)	-0.2839	(0.0097)
2004	1.0602	(0.0008)	1.9275	(0.0049)	-0.2786	(0.0098)
2005	1.0473	(0.0008)	1.6034	(0.0047)	0.0507	(0.0099)
2006	1.0447	(0.0007)	1.8967	(0.0048)	-0.2808	(0.0093)
2007	1.0667	(0.0007)	2.0184	(0.0048)	-0.4659	(0.0091)
2008	1.0478	(0.0007)	1.9298	(0.0045)	-0.3153	(0.0086)
2009	1.0448	(0.0006)	1.7939	(0.0041)	-0.2755	(0.0079)
2010	1.0355	(0.0006)	1.8190	(0.0042)	-0.2546	(0.0080)
2011	1.0250	(0.0007)	1.8127	(0.0050)	-0.2146	(0.0094)
2012	1.0221	(0.0007)	1.8331	(0.0053)	-0.2194	(0.0098)
2013	1.0290	(0.0006)	1.8491	(0.0044)	-0.2546	(0.0083)
2014	1.0399	(0.0006)	1.9684	(0.0044)	-0.2752	(0.0083)
Average	1.0451		1.8709		-0.2556	

sectors and regions. We impose $D^{jj} = 0$ for every j and $D^{jk} = 0$ if $R_j = R_k$. All coefficients in equation (31) are divided by ν to be consistent with the definition of $\widetilde{C}_t(j,k)$.[18]

4.1.1 Results

Table 4 reports the estimated moving cost parameters $\widetilde{c}_{1,t}$, $\widetilde{c}_{2,t}$, and $\widetilde{c}_{3,t}$ as well as their standard errors from the PPML estimation. The moving cost between two labor markets increases in log distance between the two as expected. This number is also close to the number that has been found in the migration literature. The estimated coefficients are not identical over years, but they are very stable over time.

4.2 Estimation of the structural parameters

The two key structural parameters we estimate are the labor mobility elasticity ν and the elasticity of substitution between tasks $\tilde{\sigma}$ which governs the degree of love for variety of tasks. We estimate both parameters by combining the local labor market approach into our structural model. We derive a regression equation similar to Autor, Dorn, and Hanson (2013), where the reduced-form coefficient has a direct link with our structural parameters.[19]

[18] Alternatively, we can simply have $\mathbf{1}_{R_j \neq R_k}$ instead of D^{jk}, but we use the information on physical distance to back out the region-level mobility friction.

[19] Our estimation strategy is in a similar spirit of Galle, Rodriguez-Clare, and Yi (2017) and Adao, Arkolakis, and Esposito (2019) who estimate the labor allocation elasticity in a static setting. For example, Galle, Rodriguez-Clare,

From the equilibrium probability of moving between labor markets m_t^{kl} as derived in (5), the shape parameter of the distribution of worker's idiosyncratic shock on jobs, ν, is essentially the labor mobility elasticity in our model. Our key results about changes in welfare, option values, and the number of jobs all depend on this key parameter, since the responsiveness of outcomes of interest crucially depends on how elastic labor mobility is.

Once we estimate $\widetilde{V}_t^k$ as the destination fixed effect in the PPML equation, we can calculate the present discounted value of the expected wage as

$$E_t \frac{\beta}{\nu} w_{t+1}^k \quad = \quad \widetilde{V}_t^k + \left(\log \mu_{0,t}^k - \log \mu_{1,t}^k\right) - \beta E_t \left[\widetilde{V}_{t+1}^k + \log \mu_{1,t+1}^k\right].$$

We use the right hand side of the equation above as the dependent variable to estimate ν based on the local labor market approach by estimating

$$\Delta y_t^k = \alpha + \tilde{\beta} \Delta Z_t^k + \lambda_t + \epsilon_t^k, \tag{32}$$

where we set $y_t^k = \widetilde{V}_t^k + \left(\log \mu_{0,t}^k - \log \mu_{1,t}^k\right) - \beta \left[\widetilde{V}_{t+1}^k + \log \mu_{1,t+1}^k\right]$ and $Z_t^k = w_{t+1}^k$, in the equation above. Conditional on $\beta = 0.95$, we back out the elasticity ν from $\tilde{\beta}$.[20]

Since wages can be endogenous we need to use an instrument for wages. In the first stage regression, we regress wages on import instrument $\Delta \bar{Z}_t^k$ discussed in Section 2 for the formal sector. Then, we use the predicted wage as an explanatory variable in the regression equation above.

Similarly, our model derives the following regression equation to estimate the structural parameter $\tilde{\sigma}$, the elasticity of substitution between tasks in the production function:

$$\log N_t^k \quad = \quad (\tilde{\sigma} - 1) \left[\log Y_t^k - \log L_t^k - \log w_t^k - \log \gamma_L\right] + \epsilon_t^k, \tag{33}$$

allowing for the error term to capture the discrepancy between O_t^k and N_t^k in the model. We can also derive $\log N_t^k = \log \mu_{1,t}^k - \log \mu_{0,t}^k$ from the model. We then use the same instrument to predict the entire terms in the bracket in the first stage. Then, in the second stage, we use the reduced form regression equation where $y_t^k = \log \mu_{1,t}^k - \log \mu_{0,t}^k$ and $Z_t^k = \left[\log Y_t^k - \log L_t^k - \log w_t^k - \log \gamma_L\right]$.

and Yi (2017) also estimate their key structural parameter using the China shock as in Autor, Dorn, and Hanson (2013). In our model, the elasticity is constructed for labor mobility both across regions and sectors, and we estimate one additional parameter which is crucial for the number of job opportunities mechanism we introduce. The labor market adjustment in our model is dynamic and subject to mobility friction, while the two aforementioned papers interpret the elasticity in a completely static setting without explicit mobility frictions.

[20] Our main results are robust to the choice of the discount factor.

Table 5: Estimation results for β/ν and $(\tilde{\sigma}-1)/\nu$

	(a) β/ν	(b) $(\tilde{\sigma}-1)/\nu$
A. First stage		
$\Delta\bar{Z}$	0.412	0.452
	(0.028)	(0.217)
F-stat	138.480	5.286
B. Second stage		
	1.962	1.37
	(0.757)	(0.62)

We then back out $\tilde{\sigma}$ from $\tilde{\beta}$ conditional on the estimated ν from the first step.

In summary, our model delivers a simple estimable equation for each of the two key structural parameters ν and $\tilde{\sigma}$. We can use the same Bartik-type instrument that we used in the reduced-form analysis, which provides a clean identification of the key structural parameters of the model.

4.2.1 Results

Panel (a) of Table 5 reports the estimates of β/ν as well as the first stage result for a change in wage with the same instrument we used for the reduced-form analysis. If we assume $\beta = 0.95$, then the implied ν is 0.484.[21] Conditional on the implied $\nu = 0.484$, the estimate for $(\tilde{\sigma}-1)/\nu$ reported in Panel (b) of Table 5 implies that the estimate of the task elasticity $\tilde{\sigma}$ is 1.663. This result confirms the love for variety of tasks channel featured in our model, which operates if $\tilde{\sigma} > 1$.

4.3 Welfare results from the trade shock

We revisit the simple regression equation from Section 2 again to estimate the impact of trade on labor market outcomes. We use predicted exports from Section 2 as the explanatory variable Z_t^k in the equation. We replace the dependent variable y_t^k with the welfare expression in equation (9), the number of jobs formula in equation (10), and the option value formulas in equations (8) and (7). We use estimated moving probabilities to construct the welfare, jobs and option value measures, without using any structural parameters, except for the labor supply elasticity ν. With a larger sample or smaller number of choices, it would be possible to plug the data directly into

[21] Our estimates are similar to what other papers in the literature have found: e.g., Artuç and McLaren (2015) find $\nu = 0.56$ with $\beta = 0.9$ and $\nu = 1.613$ with $\beta = 0.97$.

the equations with a simple bin-estimator. The bin-estimators are feasible for μ's in our model. For $\widetilde{m}$'s, on the other hand, we estimate them using the PPML as discussed before, since the bin-estimators are not feasible due to the reasons discussed above. After the estimation, we use the structural elasticity parameter ν to express the estimated numbers in annual average wages.

We regress changes in welfare-related variables on changes in exports with the same instrument as before. Table 6 reports the estimation results for each labor market outcome of interest. Since all dependent variables are divided by ν as shown in each of (8), (7), (9), and (10), we use the estimated ν to back out the implied elasticities of each outcome variable with respect to export revenues of each labor market. For the baseline implied elasticities, we assume $\nu = 0.501$ which is the obtained estimate with $\beta = 0.95$.

Table 6: Export-induced changes in welfare-related variables

	Coefficients	s.e.	Implied elasticities with $\nu = 0.484$
Welfare	0.700	(0.150)	0.339
Job opportunities	0.622	(0.131)	0.301
Internal option value	0.146	(0.022)	0.071
External option values	-0.147	(0.048)	-0.071

The result shows that a positive export shock increases worker's welfare in the corresponding labor market and the number of jobs provided in the labor market. One of the interesting results is that a positive shock increases internal option values but decreases external option values. In existing models such as ACM and CDP, a positive export shock should decrease the option value, as other labor markets become relatively less valuable after a positive export shock in your own labor market. This is captured by the effect on the external option value of our model. On the other hand, in our model with the endogenous number of job opportunities, the internal option value moves towards the opposite direction. As the number of job opportunities increases with the number of job opportunities, the internal option value increases. Due to this additional effect that our model is able to capture through the number of job opportunities, a positive export shock generates extra positive effects on the total option values compared to the existing models.

5 Model fit and robustness

In this section we show that the model's predictions are consistent with the empirical patterns observed in the data. We provide additional evidence for the importance of the channels discussed herein and the model's fit in replicating these channels. The model predicts that labor mobility should be highly correlated not only with trade shocks, but with all aggregate shocks. We will show how labor mobility changes over the business cycle consistent with the predictions of the model, which is different from the other structural labor mobility models used in international trade literature.

We first define the stylized "positive labor-market-neutral productivity shock" as a productivity shock that uniformly increases number of jobs N_t^k by $z_1 > 0$ and values V_t^k by $z_2 > 0$ in all labor markets, where $\frac{\partial}{\partial x} N_t^k = z_1$ and $\frac{\partial}{\partial x} V_t^k = z_2$, $\forall k \in \{1, 2, .., K\}$ for $s = s_1$. Using this balanced shock simplifies the algebra since we do not need to worry about changes in worker flows across labor markets.

Proposition 1. *A positive labor-market-neutral productivity shock increases average number of workers changing jobs within a labor market relative to workers staying with the same job.*

Proof. Follows directly from the equation about the number of jobs:

$$\nu \log \mu_{1,t}^k - \nu \log \mu_{0,t}^k = \log N_t^k - \delta_t,$$

thus

$$\frac{\partial}{\partial x} \frac{1}{K} \sum_k \left(\nu \log \mu_{1,t}^k - \nu \log \mu_{0,t}^k \right) > 0.$$

□

Proposition 1 implies that economy-wide positive shocks are expected to increase average log churning within cells.

Proposition 2. *A positive labor-market-neutral productivity shock increases across-labor-market churning (i.e. inter-labor-market mobility) conditional on formality.*

Proof. Follows from the lemma above.

$$
\begin{aligned}
\frac{\partial}{\partial x}\frac{1}{K}\sum_k \log(\frac{\mu_0+\mu_1}{1-\mu_I}) &= \frac{\partial}{\partial x}\frac{1}{K}\sum_k \left(-\log\left(1+\sum_{l\neq k}\exp\left(V_t^l - V_t^k - \bar{C}_t(k,l)\right)\right)\right) \\
&= \frac{1}{K}\sum_k\sum_{l\neq k} m_t^{kl}\left(\frac{z_1}{N_t^k+\exp(\delta)} - \frac{z_1}{N_t^l}\right), \\
&= \frac{1}{K}\sum_k\left(\sum_{l\neq k} m_t^{kl}\right)\left(\frac{z_1}{N_t^k+\exp(\delta)} - \frac{z_1}{N_t^k}\right), \\
&= \frac{1}{K}\sum_k \mu_{2,t}^k\left(\frac{z_1}{N_t^k+\exp(\delta)} - \frac{z_1}{N_t^k}\right) < 0,
\end{aligned}
$$

thus

$$
\frac{\partial}{\partial x}\frac{1}{K}\sum_k \log(\frac{\mu_2}{1-\mu_I}) > 0.
$$

□

We could express our main model as a standard discrete choice model (with fixed number of choices), with utility shifters and moving costs that are functions of the number of jobs. We call this alternative isomorphic model as the auxiliary model, and discuss it in the appendix. We denote the moving cost as in the auxiliary model as $\bar{C}_t(k,l)$. If we ignore the fact that number of choices can change over time, empirically it would look like the moving costs are changing. In other words, the moving costs estimated using ACM would fluctuate over time as the number of jobs are changing in labor markets.

Proposition 3. *A positive labor-market-neutral productivity reduces the average implied moving cost difference between the auxiliary and main models, $\bar{C}_t(k,l) - C_t(k,l)$.*

$$
\frac{\partial}{\partial x}\frac{1}{K^2-K}\sum_k\sum_l\left[\bar{C}_t(k,l) - C_t(k,l)\right] < 0. \tag{34}
$$

Proof.

$$\begin{aligned}
\frac{\partial}{\partial x}\frac{1}{K^2-K}\sum_k\sum_l\left[\bar{C}_t(k,l)-C_t(k,l)\right] &= \frac{1}{K^2-K}\sum_k\sum_{l\neq k}\left(\frac{z_1}{N_t^k+\exp(\delta)}-\frac{z_1}{N_t^l}\right), \\
&= \frac{1}{K^2-K}K\sum_k\left(\frac{z_1}{N_t^k+\exp(\delta)}-\frac{z_1}{N_t^k}\right), \\
&= \frac{1}{K-1}\sum_k\left(\frac{z_1}{N_t^k+\exp(\delta)}-\frac{z_1}{N_t^k}\right)<0.
\end{aligned}$$

□

Based on the propositions above, we expect to see positive correlation between aggregate productivity and churning within and across labor markets. Figure 2 shows the GDP growth rate in Brazil between 2003 and 2014, plotted together with the average number of job opportunities N_t^k implied by the model. The correlation coefficient between the two series is equal to 0.68, showing a strong positive correlation as discussed in Proposition 1. Figure 3 shows the GDP growth rate in Brazil between 2003 and 2014, plotted together with the average number of switchers across labor markets. The correlation coefficient between the two series is also positive as shown in Proposition 2 and equal to 0.8. Finally, following Proposition 3, Figure 4 shows the inverse of moving costs differences is equal to 0.62. The figures give us some confidence about the model's ability to capture channels related to the labor mobility. In addition to the evidence from Brazil, we show that $\bar{C}$ is indeed negatively correlated with aggregate shocks, as implied by the model, using data from the U.S. in the appendix Figure A5.

Figure 2: Change in GDP growth and estimated average N_t

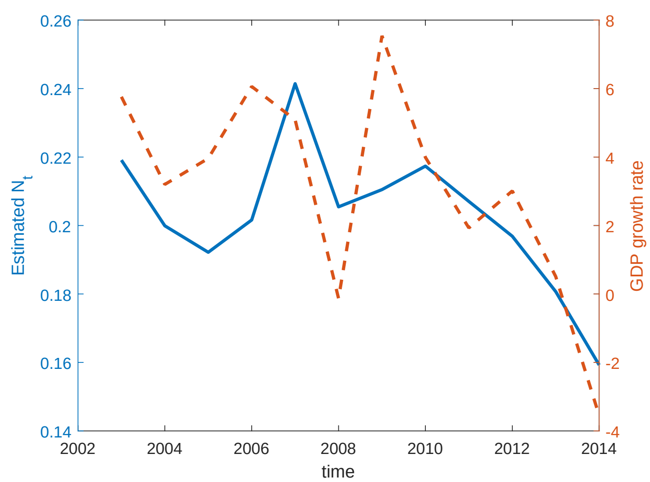


Figure 3: Change in GDP growth and inter-cell mobility

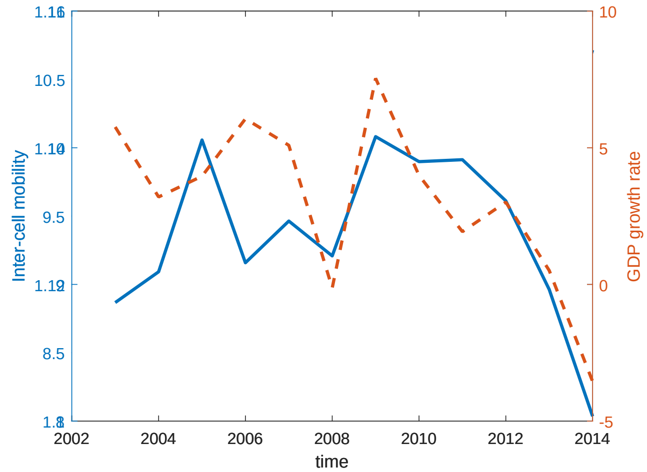
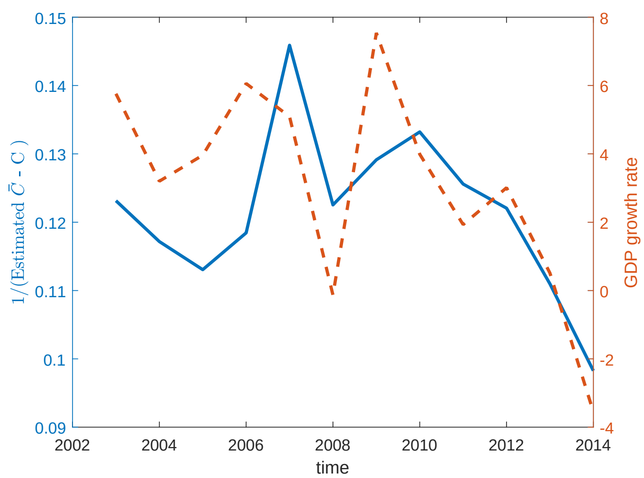


Figure 4: Change in GDP growth and $C - \bar{C}$



6 Simulation

In this section we run counterfactual simulations using our model. First, we assume that Brazil is a small open economy then increase the price index of a particular sector. After the small open economy simulation, we solve the full general equilibrium model using the dynamic hat algebra proposed by Caliendo, Dvorkin, and Parro (forthcoming) with an exogenous change in bilateral trade costs. We also run counterfactual simulations with an exogenous change in moving cost along different margins for implications on policies to alleviate adjustment friction.

[IN PROGRESS]

7 Conclusion

We have introduced a new framework to quantify the impacts of trade shocks on labor mobility and worker welfare that combines the advantages of the structural and reduced-form methodologies. Our framework features various drivers of labor mobility across sectors and regions, and identifies how trade shocks impact those determinants endogenously. Models of trade-induced labor mobility have explored wage differentials and idiosyncratic utility as drivers of mobility. We have introduced an additional important motive of mobility: the number of job opportunities provided by different sectors and regions. If a worker can choose her job out of more opportunities, it is more likely that the best one delivers higher welfare. Even when she is hit by a negative labor demand shock in the future, it is more likely that she will be able to find another job without having to move to a different region or sector, which would imply a higher switching cost. Therefore, a labor market experiencing a positive trade shock will attract more workers not just because it provides a higher wage, but also because of the larger number of job opportunities that are created there. This mechanism of dynamic labor adjustment in response to trade shock impacts worker's lifetime welfare.

We have first provided empirical evidence on the causal effects of export shocks on labor markets. The analysis draws on rich employer-employee panel data combined with customs records on export transactions from Brazil. Using changes in external demand directed to the labor market as a source of variation in exports, we documented a positive causal effect of export shocks on employment, residual wages, and job turnover rates in the corresponding labor market. Motivated by this reduced-form evidence, we developed and structurally estimated a dynamic general equilibrium model of labor mobility. Different labor markets offer different wages and different numbers of job opportunities to workers. A worker chooses the job which gives her the highest utility, where the number of jobs in each labor market is endogenously determined. In a labor market with relatively more job opportunities, workers can choose optimally out of more potential jobs, to each of which workers attach idiosyncratic preference. A job switch requiring a change of labor market implies incurring a higher switching cost compared to a job switch within a labor market. Therefore, a growing labor market with more job opportunities reduces the risk of having to pay a switching cost in the future. The prospect of job switch generates an option value in worker's welfare. Our model further decomposes this option value into the option value associated alternative job opportunities within the current labor market and the option value from having alternative jobs in all other labor markets.

Our model delivers a structural equation of changes in worker welfare which is a function of only the estimated probability of moving between labor markets and the labor supply elasticity. The welfare result does not depend on the moving cost structure, observed changes in future wages, or moving probabilities across jobs within a labor market. The effects of a trade shock are fully embedded in the gross flows between labor markets. This is a powerful result which greatly simplifies the analysis of the welfare impacts of trade shocks.

We have structurally estimated the model using the worker-firm data from Brazil. In the first stage of the estimation, we pin down the common value attached to each labor market and the moving cost between labor markets for each worker group using a gravity-like equation. The implied probability of moving between labor markets is then calculated with the estimated value of each labor market and the estimated moving cost. In the second stage of the estimation, we pin down the labor supply elasticity of our model. We first derive an estimable equation describing the relationship between a change in the transformed value of the labor market and a change in wages, with the labor supply elasticity governing the responsiveness of the former with respect to the latter. We instrument the change in residual wages with our Bartik-type instrument, exploiting variation in external demand directed to the region. Armed with the estimate of the labor supply elasticity, we estimated the causal effect of trade shocks in Brazil on workers' welfare using the same instrument. Our structural IV estimates reveal that the lifetime welfare of a median formal sector worker increased by 68% of the annual wage, following the rise in exports observed during the sample period; while wages and employment rose by 32% and 23%, respectively.

References

Acemoglu, D., D. Autor, D. Dorn, G. Hanson, and B. Price (2016): "Import Competition and the Great U.S. Employment Sag of the 2000s," *Journal of Labor Economics*, 34(S1), S141–S198. 2

Adao, R., C. Arkolakis, and F. Esposito (2019): "Spatial linkages, global shocks, and local labor markets: Theory and evidence," *NBER Working paper*, 25544. 3, 19

Amiti, M., and D. R. Davis (2012): "Exports, Export Destinations and Skills," *Review of Economic Studies*, 79(1), 1–36. 2

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

Artuç, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100(3), 1008–1045. 1, 4

Artuç, E., and J. McLaren (2015): "Trade Policy and Wage Inequality: A Structural Analysis with Occupational and Sectoral Mobility," *Journal of International Economics*, 97(2), 278–294. 21

Autor, D., D. Dorn, and G. Hanson (2015): "Untangling Trade and Technology: Evidence from Local Labor Markets," *Economic Journal*, 125(584), 621–646. 1

Autor, D., D. Dorn, G. Hanson, and J. Song (2014): "Trade Adjustment: Worker-Level Evidence," *Quarterly Journal of Economics*, 129(4), 1799–1860. 2

Autor, D., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 6(103), 2121–2168. 1, 3, 2.2, 4.2, 19

Bastos, P., J. Silva, and E. Verhoogen (2018): "Export Destinations and Input Prices," *American Economic Review*, 108(2), 353–392. 2.2

Bertrand, M. (2004): "From the Invisible Handshake to the Invisible Hand? How Import Competition Changes the Employment Relationship," *Journal of Labor Economics*, 2(22), 723–765. 2, 2.2

Brambilla, I., D. Lederman, and G. Porto (2012): "Exports, Export Destinations and Skills," *American Economic Review*, 102(7), 3406–3488. 2, 2.2

Caliendo, L., M. Dvorkin, and F. Parro (forthcoming): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*. 1, 4, 6

Coşar, K., N. Guner, and J. Tybout (2016): "Firm Dynamics, Job Turnover and Wage Distributions in an Open Economy," *American Economic Review*, 106(3), 625–663. 1

Davidson, C., L. Martin, and S. Matusz (1999): "Trade and search generated unemployment," *Journal of International Economics*, 48(2), 271 – 299. 1

di Giovanni, J., A. A. Levchenko, and I. Mejean (2018): "The Micro Origins of International Business Cycle Comovement," *American Economic Review*, 108(1), 82–108. 2.2

Dix-Carneiro, R. (2014): "Trade Liberalization and Labor Market Dynamics," *Econometrica*, 82(3), 825–885. 1

Dix-Carneiro, R., and B. Kovak (2015): "Trade Liberalization and the Skill Premium: A Local Labor Markets Approach," *American Economic Review Papers & Proceedings*, 105(5), 551–557. 1

——— (2017): "Trade Liberalization and Regional Dynamics," *American Economic Review*, 107(10), 2908–2946. 1, 2.1, 2.2, A.1

——— (2019): "Margins of Labor Adjustment to Trade," *Journal of International Economics*. 1, 2.2

Eaton, J., and S. Kortum (2002): "Technology, Geography and Trade," *Econometrica*, 70(5), 1741–79. 3.4

Frías, J. A., D. S. Kaplan, E. Verhoogen, and D. Alfaro-Serrano (2018): "Exports and Wage Premia: Evidence from Mexican Employer-Employee Data," *Unpublished Manuscript*. 2

Galle, S., A. Rodriguez-Clare, and M. Yi (2017): "Slicing the Pie, Quantifying the Aggregate and Distributional Effects of Trade," NBER Working Papaer 23737. 3, 19

Guimaraes, P., O. Figueirdo, and D. Woodward (2003): "A Tractable Approach to the Firm Location Decision Problem," *Review of Economics and Statistics*, 85(1), 201–204. 4.1

Harrison, A., J. McLaren, and M. S. McMillan (2011): "Recent Perspectives on Trade and Inequality," *Annual Review of Economics*, 3(1), 261–289. 2

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2017): "Trade and Inequality: From Theory to Estimation," *Review of Economic Studies*, 84, 1–36. 1

Hummels, D., R. Jøorgensen, and J. M. . C. Xiang (2014): "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," *American Economic Review*, 104(6), 1597–1629. 2

Kovak, B. (2013): "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?," *American Economic Review*, 103(5), 1960–1976. 1, 2.2

McLaren, J. (2017): "Globalization and Labor Market Dynamics," *Annual Review of Economics*, 9, 177–200. 1

McLaren, J., and S. Hakobyan (2016): "Looking for Local Labor Market Effects of NAFTA," *Review of Economics and Statistics*, 98(4), 728–741. 1

Pierce, J., and P. K. Schott (2016): "The Surprisingly Swift Decline of US Manufacturing Employment," *American Economic Review*, 106(7), 1632–1662. 2

Revenga, A. L. (1992): "Exporting Jobs? The Impact of Import Competition on Employment and Wages in U.S. Manufacturing," *Quarterly Journal of Economics*, 1(107), 255–284. 2, 2.2

Ritter, M. (2015): "Trade and inequality in a directed search model with firm and worker heterogeneity," *Canadian Journal of Economics/Revue canadienne d'économique*, 48(5), 1902–1916. 1

Topalova, P. (2010): "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India," *American Economic Journal: Applied Economics*, 2(4), 1–41. 1, 2.2

Traiberman, S. (2019): "Occupations and Import Competition," *Unpublished Manuscript*. 1

Verhoogen, E. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 123(2), 489–530. 2

A.1 Data sources and description

Here we provide further details about the data sets used in the empirical analysis.

Employer-employee panel data: *Relacao Anual de Informacoes Sociais* (RAIS) is a labor census gathering longitudinal data on the universe of formal workers and firms in Brazil. We use data for the period 2003-2015. RAIS is a high-quality administrative census of formal employees and employers collected every year by the Brazilian Ministry of Labor. These records are used by the government to administer several government benefits programs. Workers are required to be in RAIS in order to receive payments of these programs and firms face fines for failure to report, until they do report. These requirements ensure that RAIS is an accurate and complete census of the formal sector in Brazil (Dix-Carneiro and Kovak, 2017).

RAIS covers virtually all formal workers and provides yearly information on demographics (age, gender, and schooling), job characteristics (detailed 6-digit occupation, wage, hours worked), as well as hiring and termination dates. For each job, the RAIS annual record reports average yearly earnings, as well as the monthly wage in December. We use the information on the December wage, so as to ensure that all labor market outcomes are measured at the same time and avoid potential mismeasurement for workers that did not work full year. RAIS further includes information on a number of establishment-level characteristics, notably number of employees, geographical location (municipality) and industry code, defined according to the 5-digit level of the Brazilian National Classification of Economic Activities (CNAE). Unique identifiers for workers and firms are consistently defined across years and therefore make it possible to follow them over time. The worker unique identifier is the number associated with her registration in *Programa de Inserção Social* (PIS). The establishment unique identifier (*Cadastro Nacional de Pessoa Jurídica*) (CNPJ) is an identification number issued to Brazilian companies by the Secreteriat of the Federal Revenue of Brazil. It consists of a 12 digit number, of which the first 8 digits uniquely identify the firm and the remaining four digits identify the establishment. Therefore, it is possible to identify and track multi-establishment firms. While the RAIS data cover also segments of the public sector, we restrict the analysis to the private sector. The industry classification contain 572 industries at the 5-digit level, of which 42 are in agriculture and natural resources, 286 are in manufacturing and the remainder are in services.

The information on the level of education of the worker is reported in 9 categories: illiterate (corresponding to 0 years of education); primary school dropout (indicating from 1 to 3 years of education), primary school graduate (4 years education), middle school dropout (5 to 7 years of education), middle school graduate (8 years of education), high school dropout (9 to 10 years of education), high school graduate (11 years of education), college dropout (12 to 14 years of education), college graduate (15 years of education), Masters (18 years of education) and PhD (22.5 years of education). To compute average years of education of school dropouts, we consider the mid-point of the interval.

We use the detailed classification of occupations as a measure of the number of different jobs available in a labor market. The Brazilian Classification of Occupations changed in 2002 (CBO-2) and has been reported consistently since 2003. Although the RAIS data are available for earlier years, we restrict the analysis to the post-2003 period in order to ensure that this important variable is defined in a consistent way throughout the period of analysis. CBO-2 aims to portray the reality of professions of the Brazilian labor market. It was established with legal basis in Administrative Rule no. 397 of October 10, 2002. There are 2637 occupation codes at the 6-digit level during this period. The description of each 6-digit code is available at http://portalfat.mte.gov.br/wp-content/uploads/2016/04/CBO2002_Liv3.pdf.

We use information on the establishment's location (municipality) and industry, and worker-level data on gender, age, education and December wage. We focus on workers aged 16 to 64 years old. As in Dix-Carneiro and Kovak (2017), we use the "microregion" concept of the Brazilian Statistical Agency (IBGE) to define regional boundaries. This definition groups together economically integrated contiguous municipalities with similar geographic and productive attributes. The documentation supporting this definition is available at `https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?id=22269&view=detalhes`. We consider a set of 558 consistently defined microregions, grouping the 5571 municipalities in the data. To ensure a consistent definition of microregions over time, when necessary we merge microregions whose boundaries changed over the period of analysis.

Customs records: We also use customs data on export transactions by microregion, industry and destination in each year. These customs records are administrative data collected by Secretaria do Comercio Externo (SECEX) of the Ministry of Development, Industry and Foreign Trade. These data are available since 1997 and contain information on FOB export values and quantities, and are originally defined at the level of the municipality, detailed product category and destination market. The customs records were originally collected by SECEX at the firm-product-destination level. To aggregate up to the municipal level, SECEX attributed each firm-level export transaction to the municipality where the headquarters of the exporting firm are located. The product classification is *Nomenclatura Comum do MERCOSUL* (NCM), at the 8-digit level. For consistency with RAIS, we restricted the analysis to the post-2003 period, and aggregated up to microregion-sector level. To aggregate exports from the NCM 8-digit level to the 5-digit level of the CNAE, we used a concordance made available to us by SECEX.

Industry-level imports of Brazil's destinations: To construct an instrument for exports, we further use yearly data on the industry-level imports of each of Brazil's export destinations. To capture changes in sectoral import demand that are plausibly exogenous to microregions in Brazil, we consider the imports of these countries sourced from all countries other than Brazil (i.e. we exclude imports sourced from Brazil from total imports of each country in a given industry-year). There is a total of 189 destinations reported in the customs data, to which we link information on sectoral import demand from the UN COMTRADE data set.

A.2 Variable definitions and summary statistics

This section describes in detail the variables used in the econometric analysis:

Δ Employment: log change in the number of employees in microregion-sector k between years $t-1$ and t;

Δ Residual wage: log change in average residual wage in the microregion-sector k between years $t-1$ and t;

Δ # unique occupations: log change in the number of unique occupations (defined at the 6-digit level of CBO-2) in the microregion-sector k between years $t-1$ and t;

Δ # of leavers: log change in the number of workers leaving microregion-sector k (i.e., incumbent workers changing either microregion or sector) between years $t-1$ and t;

- Δ # of entrants: log change in the number of workers entering microregion-sector pair k between year $t-1$ and t from other microregion and/or sector (thus excluding new entrants to the formal labor force);

- Δ # of job switchers: log change in the number of workers that switch jobs (i.e., switch either occupation or establishment), while staying in the same microregion-sector between years $t-1$ and t;

- Δ Exports: log change in the value of the exports originated in the microregion-sector between years $t-1$ and t;

- Δ $\bar{Z}$: log change in the value of import demand directed to the microregion-sector between years $t-1$ and t, as defined in equation (2) in text.

A.3 Estimation of switching probabilities, values and moving costs

Here we show that PPML orthogonality conditions are equal to the MLE first order conditions for our model for the estimation of values and moving costs (subject to a normalization ν). We omit the type superscript s. We denote the number of agents moving from j to k with y^{jk}, the expected value (i.e. destination fixed) effect with $\widetilde{V}^k$, the origin fixed effect with $\widetilde{\Gamma}^j$ and the moving cost with $\widetilde{C}(j,k)$.

Consider the following moving cost structure:

$$\widetilde{C}(j,k) = \widetilde{c_1} D^{jk} + \widetilde{c_2} \mathbf{1}_{S_j \neq S_k}, \tag{A1}$$

where $\widetilde{c_j}$ is the distance coefficient (divided by ν) and D^{jk} is the log of distance between l and k, $\widetilde{c_2}$ is the sector switching cost (divided by ν), $\mathbf{1}_{S_l \neq S_k}$ is an indicator function that is equal to one if l and k are associated with different sectors. Note that we impose $D^{ll} = 0$ for every j. We omit the time sub-scripts and the last component of moving cost to simplify exposition.

A.3.1 Maximum Likelihood Estimation (First Order Conditions)

The likelihood function is

$$\mathcal{L} = \prod_j \prod_k \left(m^{jk}\right)^{y^{jk}}, \tag{A2}$$

or alternatively using logarithm

$$\log \mathcal{L} = \sum_j \sum_k y^{jk} \log(m^{jk}). \tag{A3}$$

Note that the moving probability can be expressed as

$$\begin{aligned} m^{jk} &= \frac{\exp\left(\widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right)}{\sum_l \exp\left(\widetilde{V}^l - \widetilde{c}_1 D^{jl} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_l}\right)}, \\ &= \exp\left(\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k} - \log \widetilde{L}_t^j\right), \end{aligned}$$

where $\widetilde{\Gamma}^j = -\log\left[\sum_k \exp\left(\widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right)\right] + \log(\widetilde{L}_t^j)$.

The log-likelihood function, then, can be written as

$$\log \mathcal{L} = \sum_j \sum_k y_t^{jk} \left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k} - \log(\widetilde{L}^j)\right],$$

subject to

$$\widetilde{\Gamma}^j = -\log\left[\sum_k \exp\left(\widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right)\right] + \log(\widetilde{L}^j). \tag{A4}$$

The goal is to find $\widetilde{V}^j$, $\widetilde{c}_2$ and $\widetilde{c}_1$ coefficients that maximize the log likelihood function.

Note that

$$\partial \widetilde{\Gamma}^j / \partial \widetilde{V}^k = -m^{jk},$$

$$\partial\widetilde{\Gamma}^j/\partial\widetilde{c_1} = -\sum_{j\neq k} D^{jk} m^{jk},$$

$$\partial\widetilde{\Gamma}^j/\partial\widetilde{c_2} = -\sum_{j\neq k} \mathbf{1}_{S_j\neq S_k} m^{jk},$$

and

$$m^{jk} = \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c_1}D^{jk} - \widetilde{c_2}\mathbf{1}_{I_j\neq I_k} - \log(\widetilde{L}^j)\right].$$

Values:

We take the derivative of the log likelihood function with respect to $\widetilde{V}^k$ to find the first order condition

$$\frac{d\log\mathcal{L}}{d\widetilde{V}^k} = \frac{\partial\log\mathcal{L}}{\partial\widetilde{V}^k} + \sum_j \frac{\partial\log\mathcal{L}}{\partial\widetilde{\Gamma}^j}\frac{\partial\widetilde{\Gamma}^j}{\partial\widetilde{V}^k} = 0 \tag{A5}$$

We rearrange the terms:

$$\begin{aligned}
0 &= \frac{\partial\log\mathcal{L}}{\partial\widetilde{V}^k} - \sum_j \frac{\partial\log\mathcal{L}}{\partial\widetilde{\Gamma}^j} m^{jk} \\
&= \sum_j y^{jk} - \sum_j \left(\sum_k y^{jk}\right) m^{jk} \\
&= \sum_j y^{jk} - \sum_j \widetilde{L}^j m^{jk} \\
&= \sum_j y^{jk} - \sum_j \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c_1}D^{jk} - \widetilde{c_2}\mathbf{1}_{S_j\neq S_k}\right] \\
&= \sum_j \left(y^{jk} - \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c_1}D^{jk} - \widetilde{c_2}\mathbf{1}_{S_j\neq S_k}\right]\right)
\end{aligned}$$

thus the first order condition associated with values is

$$\sum_j \left(y^{jk} - \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c_1}D^{jk} - \widetilde{c_2}\mathbf{1}_{I_j\neq I_k}\right]\right) = 0. \tag{A6}$$

Distance coefficient:

Then we take the derivative of the log likelihood function with respect to distance coefficient

$\widetilde{c}_1$:

$$\frac{d\log\mathcal{L}}{d\widetilde{c}_1}=\frac{\partial\log\mathcal{L}}{\partial\widetilde{c}_1}+\sum_i\frac{\partial\log\mathcal{L}}{\partial\widetilde{\Gamma}^i}\frac{\partial\widetilde{\Gamma}^i}{\partial\widetilde{c}_1}=0 \tag{A7}$$

$$\begin{aligned}
0 &= -\sum_j\sum_{k\neq j}D^{jk}y^{jk}+\sum_j\sum_l y^{jl}\sum_{k\neq j}D^{jk}m^{jk}\\
&= \sum_j\sum_{k\neq j}D^{jk}y^{jk}\sum_j\widetilde{L}^j\sum_{k\neq j}D^{jk}m^{jk}\\
&= \sum_j\sum_{k\neq j}D^{jk}y^{jk}-\sum_j\sum_{k\neq j}D^{jk}\exp\left[\widetilde{\Gamma}^j+\widetilde{V}^k-\widetilde{c}_1D^{jk}-\widetilde{c}_2\mathbf{1}_{S_j\neq S_k}\right]
\end{aligned}$$

thus the first order condition associated with $\widetilde{c}_1$ is

$$\sum_j\sum_{k\neq j}D^{jk}\left(y^{jk}-\exp\left[\widetilde{\Gamma}^j+\widetilde{V}^k-\widetilde{c}_1D^{jk}-\widetilde{c}_2\mathbf{1}_{S_j\neq S_k}\right]\right)=0. \tag{A8}$$

Sector switching coefficient:

Then we take the derivative of the log likelihood function with respect to distance coefficient $\widetilde{c}_1$:

$$\frac{d\log\mathcal{L}}{d\widetilde{c}_2}=\frac{\partial\log\mathcal{L}}{\partial\widetilde{c}_2}+\sum_t\sum_i\frac{\partial\log\mathcal{L}}{\partial\widetilde{\Gamma}^i_t}\frac{\partial\widetilde{\Gamma}^i_t}{\partial\widetilde{c}_2}=0 \tag{A9}$$

$$\begin{aligned}
0 &= -\sum_j\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}y^{jk}+\sum_j\sum_l y^{jl}\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}m^{jk}_t\\
&= \sum_j\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}y^{jk}\sum_j\widetilde{L}^j\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}m^{jk}\\
&= \sum_j\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}y^{jk}-\sum_j\sum_{k\neq j}\mathbf{1}_{I_j\neq I_k}\exp\left[\widetilde{\Gamma}^j+\widetilde{V}^k-\widetilde{c}_1D^{jk}-\widetilde{c}_2\mathbf{1}_{I_j\neq I_k}\right]
\end{aligned}$$

thus the first order condition associated with c is

$$\sum_j\sum_{k\neq j}\mathbf{1}_{S_j\neq S_k}\left(y^{jk}-\exp\left[\widetilde{\Gamma}^j+\widetilde{V}^k-\widetilde{c}_1D^{jk}-\widetilde{c}_2\mathbf{1}_{S_j\neq S_k}\right]\right)=0. \tag{A10}$$

A.3.2 PPML (Orthogonality Conditions)

Now we turn to the PPML regression equation. We will show that the orthogonality conditions implied by the PPML regression equation are identical to the ML first order conditions. PPML can be preferable to ML for two reasons: (i) Since is straightforward to take analytical derivatives of the orthogonality conditions, PPML is very low cost computationally. (ii) There are many software packages to estimate PPML. We will prove that PPML and ML estimators are identical for our model.

The PPML equation (without type superscript) is

$$y_t^{jk} = \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right] + \epsilon_t^{jk}, \tag{A11}$$

The regression equation can be written in matrix form as

$$y = \exp\left[XB\right] + \epsilon, \tag{A12}$$

where y is a vector with elements y_t^{jk}, X is a matrix of destination and origin dummies and switching cost variables, B is the vector of coefficients.

The orthogonality condition of PPML regression is

$$0 = X'\left(y - \exp\left[XB\right]\right). \tag{A13}$$

This matrix operation implies a vector of equations.

We can group the rows (i.e. equations) in of the orthogonality condition matrix above into four categories:

I. Equations associated with the origin coefficients

$$\sum_k \left[y^{jk} - \exp\left(\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right)\right] = 0, \ \forall j, \tag{A14}$$

II. Equations associated with the destination coefficients

$$\sum_j \left(y^{jk} - \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right]\right) = 0, \ \forall k, \tag{A15}$$

III. Equation associated with the distance coefficient

$$\sum_j \sum_{k \neq j} D^{jk} \left(y^{jk} - \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right]\right) = 0, \tag{A16}$$

IV. Equation associated with the sector switching cost coefficient

$$\sum_j \sum_{k \neq j} \mathbf{1}_{S_j \neq S_k} \left(y^{jk} - \exp\left[\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c}_1 D^{jk} - \widetilde{c}_2 \mathbf{1}_{S_j \neq S_k}\right]\right) = 0. \tag{A17}$$

Note that equation (A6) is same as (A15); equation (A8) is same as (A16); and equation (A10) is same as (A17).

To conclude the proof, we have to show the restriction (A4) of the ML estimation is same as the equation (A14) of the PPML regression.

Consider equation (A14) from above

$$y^{jj} - \exp(\widetilde{\Gamma}^j + \widetilde{V}^j) + \sum_{k \neq j} \left[y^{jk} - \exp\left(\widetilde{\Gamma}^j + \widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k} \right) \right] = 0$$

we can arrange the terms as

$$\begin{aligned}
0 &= \sum_k \left[y^{jk} - \exp\left(\widetilde{\Gamma}^j\right) \exp\left(\widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k} \right) \right] \\
\sum_k y_t^{jk} &= \exp\left(\widetilde{\Gamma}_t^j\right) \sum_k \exp\left(\widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k} \right) \\
\widetilde{L}^j &= \exp\left(\widetilde{\Gamma}^j\right) \sum_k \exp\left(\widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k} \right) \\
\exp\left(\widetilde{\Gamma}^j\right) &= \frac{L^j}{\sum_k \exp\left(\widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k} \right)}
\end{aligned}$$

thus we get

$$\widetilde{\Gamma}^j = \log(\widetilde{L}^j) - \log\left(\sum_k \exp(\widetilde{V}^k - \widetilde{c_1} D^{jk} - \widetilde{c_2} \mathbf{1}_{S_j \neq S_k}) \right), \tag{A18}$$

which is equal to the restriction (A4) in the ML estimator. Therefore, solving the first order conditions of the ML estimator is equivalent to solving the orthogonality conditions in PPML.

A.4 Auxiliary model

Consider the following model with constant number of choices.

The economy has L agents, and each agent is attached to a region and/or sector labor market k, where $k \in \{1, 2, .., K\}$. The number of agents in labor market k is denoted as L_t^k. An agent, who is indexed with h and attached to labor market k, will receive instantaneous utility $\bar{u}_t^h$ at time t defined as

$$\bar{u}_t^h = w_t^k + \eta_t^k + \varepsilon_t^{h,k}, \tag{A19}$$

where $\varepsilon_t^{h,k}$ is distributed Gumbel with mean 0 and scale ν. The labor-market-specific utility shifter η is defined as

$$\eta_t^k = \log\Big(N_t^k + \exp(\delta_t)\Big) - \delta_t, \tag{A20}$$

where $N_t^{s,k}$ and δ_t^s are as defined in the main model section.

The workers pay moving cost

$$\bar{C}_t(k,l) = C_t(k,l) + \log\left(\frac{N_t^k + \exp(\delta_t)}{N_t^l}\right), \tag{A21}$$

where $\bar{C}_t(k,l)$ is the implied moving cost with $\bar{C}_t(k,k) = 0$, and $C_t(k,l)$ is the structural moving cost parameter herein. The auxiliary model above is isomorphic to the model described in Section 3, and as $N_t^k \to 0$ the model becomes equivalent to ACM.

A.5 Sampling rate and asymptotics

In this section, we show how it is possible to use a real number for the number of sampled jobs instead of an integer for simulation purposes. When the number of sampled jobs goes to infinity, welfare and other important variables in the model can still be finite.

Consider $N_t^k = \rho(O_t^k)$, a step function $\rho : \mathbb{R}^+ \to \mathbb{N}^+$ where

$$\vartheta O_t^k < \rho(O_t^k) \leq \vartheta O_t^k + 1. \tag{A22}$$

Imagine that ε is distributed Gumble with mean $\kappa = 0$ and scale parameter ν. The moving cost $\delta = \tilde{\delta} + \nu\log(\vartheta)$. Welfare

$$W \quad = \quad \nu\log\left[\sum N_t^k \exp(\frac{V_t^k - \delta}{\nu})\right], \tag{A23}$$

then

$$
\begin{aligned}
W &= \nu \log \left[\sum N_t^k \exp(\frac{V_t^k - \tilde{\delta} - \nu \log(\vartheta)}{\nu}) \right], && \text{(A24)} \\
&= \nu \log \left[\sum \frac{N_t^k}{\vartheta} \exp(\frac{V_t^k - \tilde{\delta}}{\nu}) \right], && \text{(A25)} \\
&= \nu \log \left[\sum \tilde{N}_t^k \exp(\frac{V_t^k - \tilde{\delta}}{\nu}) \right], && \text{(A26)} \\
& && \text{(A27)}
\end{aligned}
$$

where

$$
\tilde{N}_t^k = \rho(O_t^k) \frac{1}{\vartheta}. \tag{A28}
$$

Note that

$$
\lim_{\vartheta \to \infty} \tilde{N}_t^k = O_t^k. \tag{A29}
$$

Proof.

$$
\begin{aligned}
\lim_{\vartheta \to \infty} \frac{\vartheta O_t^k + 1}{\vartheta} &= \lim_{\vartheta \to \infty} \left(O_t^k + \frac{1}{\vartheta} \right), && \text{(A30)} \\
&= O_t^k, && \text{(A31)}
\end{aligned}
$$

and

$$
\lim_{\vartheta \to \infty} \frac{\vartheta O_t^k}{\vartheta} = O_t^k. \tag{A32}
$$

Note that (squeezing functions with the same limit)

$$
\frac{\vartheta O_t^k}{\vartheta} < \frac{\rho(O_t^k)}{\vartheta} \leq \frac{\vartheta O_t^k + 1}{\vartheta}. \tag{A33}
$$

thus

$$\lim_{\vartheta \to \infty} \frac{\rho(O_t^k)}{\vartheta} = O_t^k. \tag{A34}$$

A.6 Appendix Tables

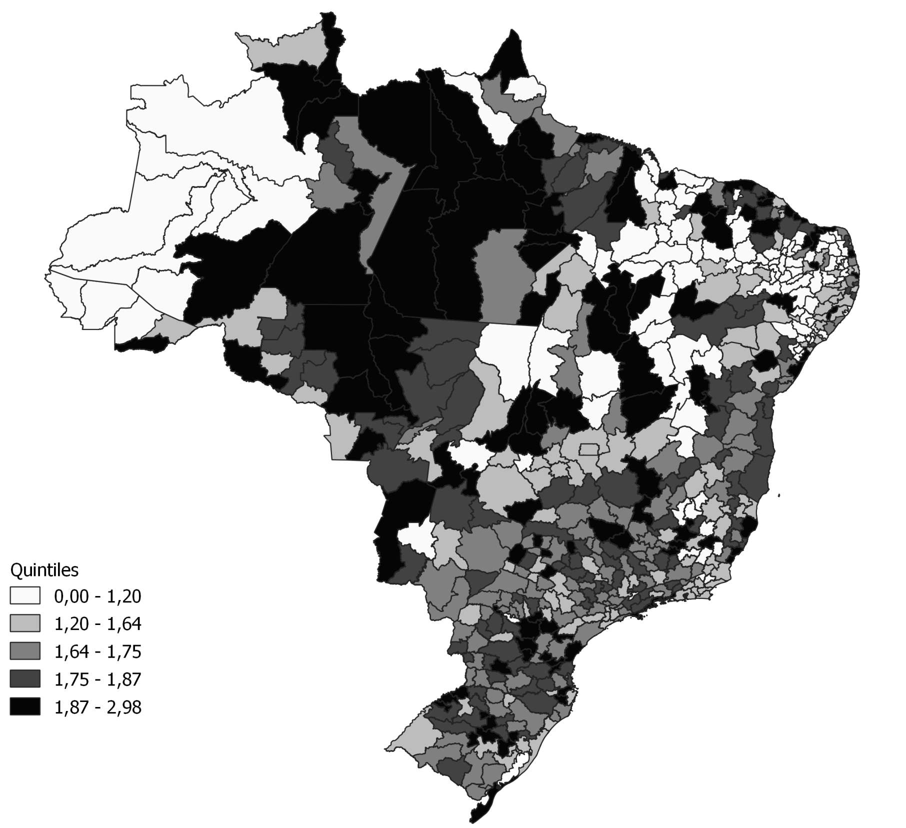
Table A1: Gross job flow rates by sector and year

	Agriculture and mining		Manufacturing		Services		All sectors	
year	inflows	outflows	inflows	outflows	inflows	outflows	inflows	outflows
2004	0.084	0.086	0.066	0.063	0.054	0.076	0.068	0.075
2005	0.091	0.091	0.067	0.094	0.059	0.082	0.072	0.089
2006	0.112	0.091	0.087	0.071	0.060	0.076	0.087	0.079
2007	0.087	0.087	0.069	0.072	0.060	0.079	0.072	0.079
2008	0.083	0.089	0.067	0.074	0.062	0.075	0.071	0.079
2009	0.090	0.102	0.068	0.090	0.067	0.079	0.075	0.090
2010	0.089	0.101	0.075	0.080	0.068	0.082	0.077	0.088
2011	0.094	0.098	0.074	0.082	0.072	0.086	0.080	0.089
2012	0.107	0.094	0.074	0.078	0.072	0.081	0.084	0.084
2013	0.094	0.087	0.076	0.073	0.070	0.073	0.080	0.077
2014	0.092	0.078	0.070	0.065	0.064	0.063	0.075	0.069
Average	0.093	0.091	0.072	0.076	0.064	0.077	0.076	0.082

Notes: Table reports average gross job inflow and outflow rates for each sector and year in Brazilian microregions in 2004-2014

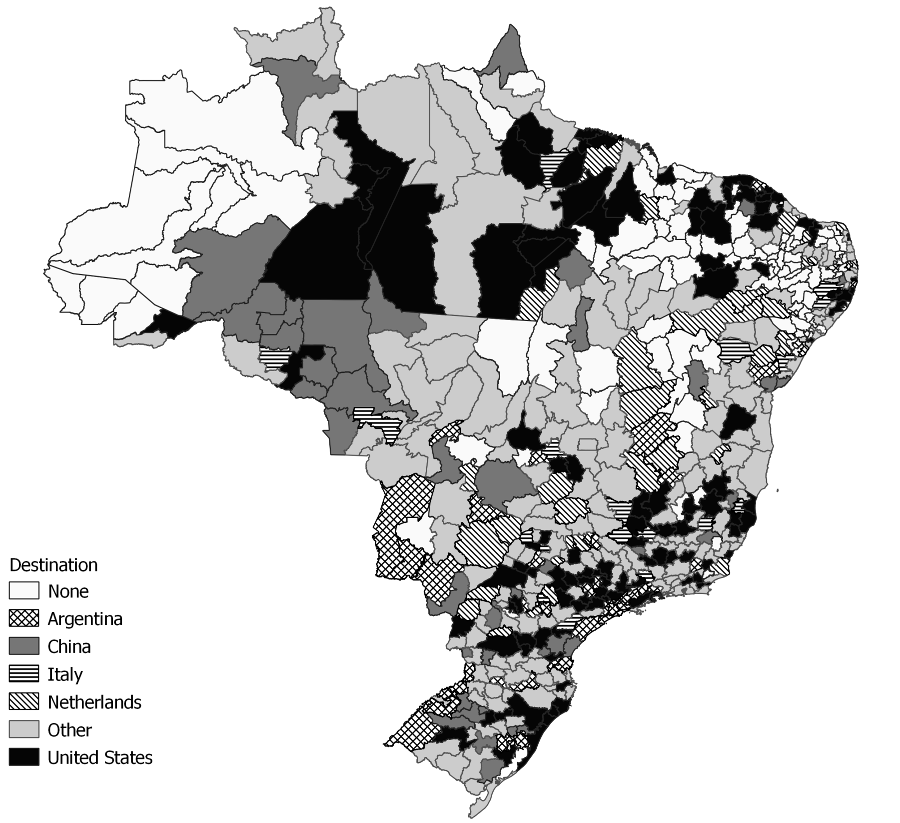
A.7 Appendix Figures

Figure A1: Export revenue per worker, 2003



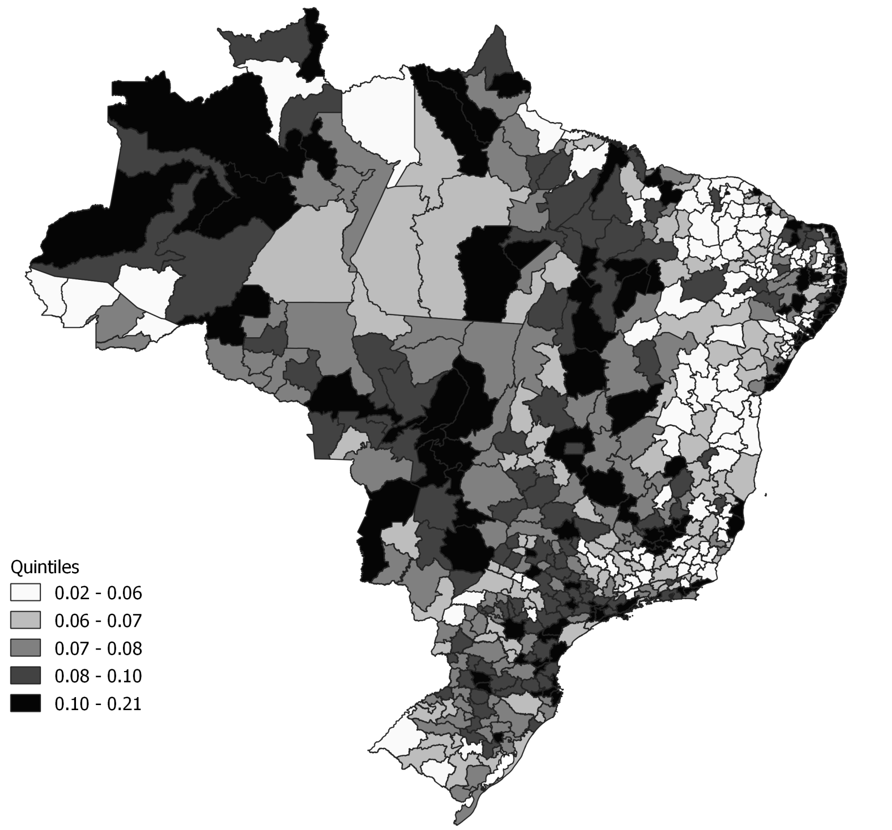
Notes: Figure depicts the log of (1+exports) per worker in Brazilian microregions in 2003.

Figure A2: Top export destinations, 2003



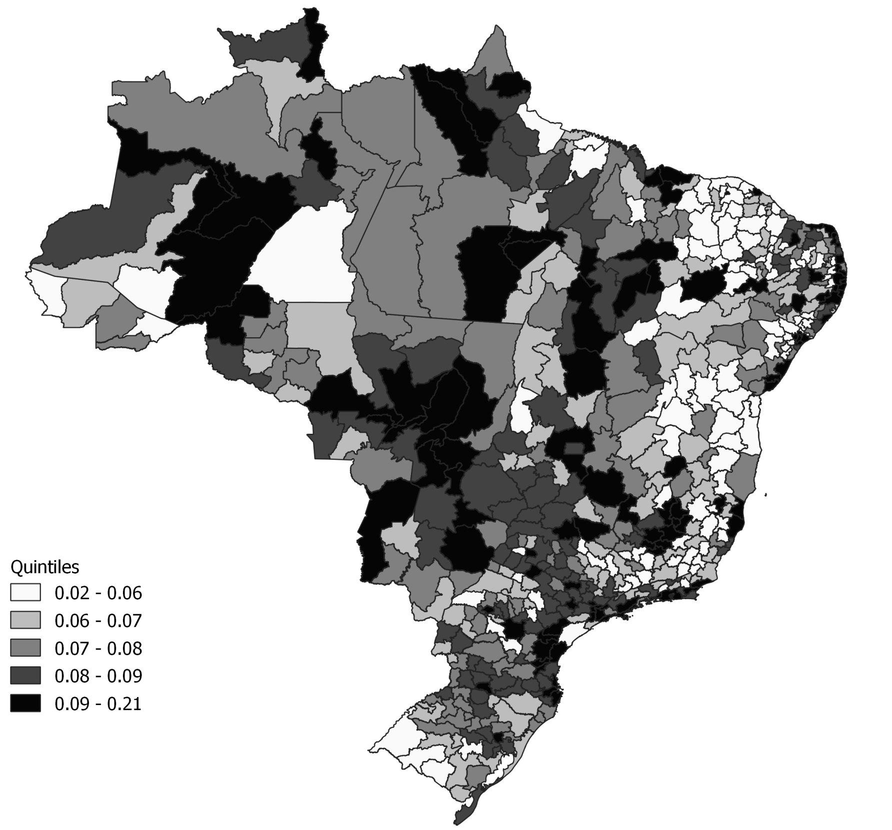
Notes: Figure depicts the top export destination of each Brazilian microregion in 2003.

Figure A3: Gross outflow rates by microregion, 2004-2014



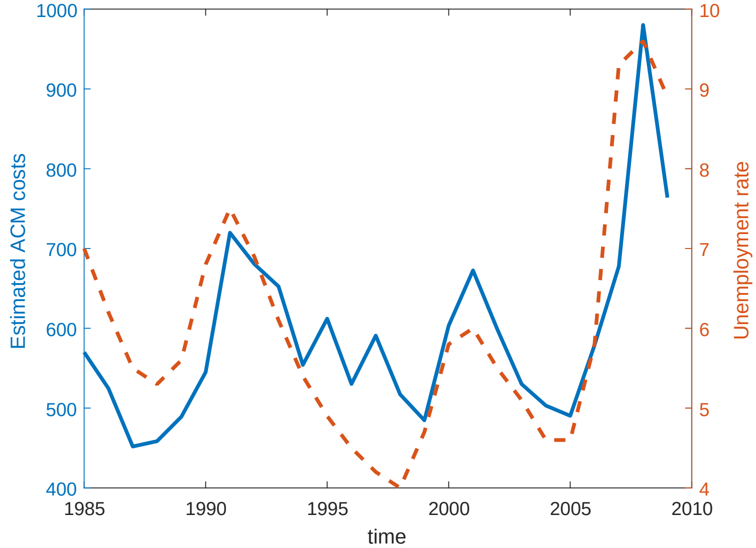
Notes: Figure depicts the average gross job outflow rates observed in each Brazilian microregion in 2004-2014.

Figure A4: Gross inflow rates by microregion, 2004-2014



Notes: Figure depicts the average gross job inflow rates observed in each Brazilian microregion in 2004-2014.

Figure A5: Additional evidence on model fit from the USA: change in unemployment rate and estimated time varying ACM moving costs



Notes: Figure depicts the correlation between the ACM moving cost and unemployment rate in USA between 1985 and 2009. The correlation coefficeint between the two series is equal to 0.58.