Spatial linkages, Global Shocks, and Local Labor Markets: Theory and Evidence*

Rodrigo Adao
Chicago Booth

Costas Arkolakis
Yale

Federico Esposito
Tufts

February 2018

Abstract

We analyze and measure the impact of economic shocks on local labor markets in the presence of rich cross-market links, arising from trade, agglomeration spillovers and labor mobility. We model these links within a generalized perfectly competitive framework and characterize the key aggregate elasticities that fully determine both the differential and aggregate impact of trade and productivity shocks. In this environment, we show that the presence of spatial links across markets generates indirect spillover effects and thus reduced form empirical analysis may not identify the correct elasticities. We instead propose an empirical methodology that relies on a Model-Implied Instrumental Variable, which we show to be optimal, in the sense that it yields the most efficient estimator of the elasticities. These insights are general since our framework is observationally equivalent to a large class of existing trade and geography models with cross-market links. We apply our estimator using disaggregated regional trade data to measure the impact of trade shocks on US local labor markets. Our counterfactual exercises reveal that the importance of general equilibrium indirect effects are directly related to the magnitude of the migration elasticities.

*We thank Pol Antras, Martin Beraja, Lorenzo Caliendo, Arnaud Costinot, Javier Cravino, Dave Donaldson, Max Dvorkin, Farid Farrokhi, Gordon Hanson, Samuel Kortum, Andrew McCallum, Eduardo Morales, Steve Redding, Esteban Rossi-Hansberg, Erik Hurst, Daniel Trefler, Jonathan Vogel, as well as seminar participants at Wisconsin-Madison, Princeton, Tufts, World Bank, WTO, HKUST, SED Edinburgh, SAET Faro, NBER Summer Institute, UPenn, University of Maryland, Fed Board of Governors, Bank of Canada, University of Zurich, University of Toronto, University of Cambridge, University of Zurich, University of Amsterdam, EIIT Conference, Brown, Yale, BFI Chicago, Midwest Trade Conference for helpful suggestions and comments. All errors are our own.

1 Introduction

How do foreign and domestic shocks, such as international trade, technological change, and macroeconomic policy affect the performance of local labor markets? A growing literature exploits differential local exposure to shocks from various sources to measure differential effects on employment and wages. However, when local markets are interconnected, inference about total effects of shocks from these estimates is not straightforward.[1] The challenge is the existence of cross-market links itself: it implies that any local shock percolates through the entire network, indirectly affecting other markets in equilibrium. In this paper, we argue that the theoretical assessment and the empirical measurement of these cross-market links are crucial to evaluate the effect of economic shocks on labor market outcomes.

Our starting point is a Generalized Spatial Competitive Economy featuring countries constituted by multiple local labor markets, i.e. sector-region pairs. In each market, production of a composite good is subject to agglomeration and congestion forces – i.e., productivity is a function of employment in different markets. Countries have a representative household with preferences for consumption and labor supply across markets. The utility maximization problem determines, for each market, the bilateral trade flows as a function of good prices, and the labor supply as a function of real wages. Thus, our general equilibrium economy features three aggregate mappings that control cross-market links in terms of (i) productivity, (ii) labor supply, and (iii) bilateral trade flows.

In this environment, we theoretically assess the impact on local labor markets of changes in local shifters, ''shocks'' henceforth, of labor supply, productivity, or trade costs. Following shocks to any particular market, we characterize how the shape of the aggregate mappings determines not only the direct response of employment and wages in that market, but also the indirect spillover effects on all other markets arising from cross-market linkages. We establish that knowledge of the three aggregate mappings of cross-market links is sufficient to compute the effect of trade and productivity shocks on local labor markets. This insight holds in a large class of geography, trade, and spatial sorting models existing in the literature. In fact, we establish that, by properly specifying the shape of the aggregate mappings, our model generates labor market outcomes that are observationally equivalent to those implied by existing frameworks.[2] This implies that the different microeconomic structure of

[1] There is a growing literature studying the local effects of international trade shocks – e.g., see Topalova (2010), Kovak (2013), Autor et al. (2013) and, for a review, Muendler (2017). Several papers have also documented local effects of technological shocks – see Autor and Dorn (2013), Bustos et al. (2016), and Acemoglu et al. (2017). Finally, an emerging literature evaluates the local effects of macroeconomic shocks – see Mian and Sufi (2014), Nakamura and Steinsson (2014) and Beraja et al. (2016). Muendler (2017) discusses the issues associated with extrapolating aggregate effects of a shock from the estimated regional effects.

[2] These frameworks include: (i) Neoclassical trade theories with external economies of scale (as in Ethier

these frameworks matters only insofar it affects the elasticity structure of the key aggregate mappings.

These indirect spillover effects are a general equilibrium source of local exposure to shocks affecting the world economy. Crucially, we show that their omission in any empirical analysis induces the mismeasurement of both differential and aggregate effects of economic shocks on local labor markets. Only under the restrictive assumption of a Small Open Economy, in which local shocks do not spillover to other markets in general equilibrium, simple difference-in-difference empirical specifications are able to recover *both* the differential and aggregate impact of trade cost shocks.[3] Therefore, the consequences of model mis-specifications depend on the magnitude and dispersion of indirect effects across markets.

Our theoretical results indicate that it is essential to empirically measure the cross-market links embedded in the elasticity structure of the central aggregate mappings in our model. To estimate this elasticity structure, we characterize a class of consistent estimators based on observed trade shocks that are orthogonal to unobservable local shocks in productivity and labor supply. We exploit the fact that, conditional on the shape of the aggregate mappings, these unobservable local shocks can be recovered from changes in trade and labor outcomes across local markets. Thus, the aggregate mappings can be estimated using moment conditions that combine the recovered local shocks and *arbitrary* functions of observed trade shocks.[4]

We then use our general equilibrium model to characterize the ''optimal'' estimator within this class of consistent estimators. That is, we compute the function of observed trade shocks that minimizes the estimator's asymptotic variance. We show that such a function is the impact of *observed* trade shocks on the *endogenous* variables predicted by our general equilibrium model – the Model-implied Optimal IV (MOIV). Intuitively, through the lens of the model, these predicted changes correspond to the total effect of the exogenous trade shock on the endogenous local outcomes and, therefore, contain all sources of variation available to estimate the unknown structural parameters. We further show that the optimal estimator can be easily implemented through a Two-Stage GMM estimator.

(1982a,b)); (ii) New trade theory models with homogeneous firms (as in Krugman (1980)); (iii) New Economic Geography models (as in Krugman (1991) and Allen and Arkolakis (2014)); (iv) Roy spatial assignment models (as in Galle, Rodriguez-Clare, and Yi (2015) and Adão (2015)).

[3]This is the main intuition behind the multiple-sector model that motivates the empirical specifications in Autor, Dorn, and Hanson (2013) and Kovak (2013).

[4]The central orthogonality assumption of our methodology is typically maintained in empirical papers measuring the impact of trade shocks on local labor markets (see Topalova (2010), Kovak (2013), Autor, Dorn, and Hanson (2013), and Hakobyan and McLaren (2016)). However, as in Berry and Haile (2014) and Adao et al. (2017), the recovery of unobservable local shocks requires the invertibility of the central functions in the model. In our setting, the mappings of labor supply and bilateral trade must be invertible, which holds in a variety of existing frameworks in the literature.

In the second tier of our analysis, we apply our methodology to empirically evaluate the role of cross-market linkages in determining total effect of trade shocks on local labor markets in the United States. We consider a parametrization of the Generalized Spatial Competitive Economy that captures labor supply and productivity linkages embedded in existing quantitative papers. To implement our methodology, we take advantage of the globalization process that brought down tariffs around the world between 1993 and 2012. Such a widespread decline in trade barriers is plausibly exogenous to regional shocks in productivity and labor supply in the US. We use this to estimate cross-market linkages from its induced variation on trade and labor outcomes across markets.

To measure local labor market outcomes, we use US regional data on employment, wages and prices. To measure spatial links across regions, we construct a bilateral matrix of trade flows across US states and 59 foreign countries, and aggregate this data at the manufacturing and non-manufacturing sectors, for the years 1993-2012. We obtain estimates of intensive and extensive margins elasticities of labor supply that are consistent with the estimates reviewed by Chetty (2012). While we find evidence on mobility across sectors within regions, we only find weak evidence on between-state mobility in response to changes in real wages, consistent with previous evidence in Autor et al. (2013), Kovak (2013) and Dix Carneiro and Kovak (2016). We finally find strong *local* agglomeration and congestion forces, close to the local agglomeration forces implied by models with increasing returns to scale and firm entry as in Krugman (1980). To gain confidence in our model, we attest that changes in trade and labor market outcomes observed in the data are highly correlated with those predicted by our general equilibrium model in response to the observed tariff shocks. In other words, the first stage of our model-implied IV methodology is very strong.

Finally, we use our estimates to assess the empirical relevance of the general equilibrium effects arising from trade shocks highlighted in the theoretical model. As a first step, we recover changes in trade costs between US states and foreign countries from the observed changes in trade flows and prices of expenditures between 1993 and 2007.[5] We then show that the direct and indirect exposures present a negative correlation across local labor markets in the US, but indirect effects are economically small compared to direct effects. In fact, on average, direct effects can be up to 3 times higher than the indirect, but the relative importance of indirect effects increases with the magnitude of the migration elasticity across regions. We establish that trade shocks accounted for 7% of the reduction in manufacturing employment in the US in the last decade. The decomposition of these shocks across countries indicates that exposure to the so-called ‘‘China shock’’ is correlated with exposure from shocks

[5]We extend the methodology proposed by Head and Ries (2001) to the case of asymmetric trade costs. This requires using information on aggregate price indices as well, similarly to Eaton et al. (2016).

to many other emerging markets. Our results indicate the limited role of trade policy in reverting recent declines in manufacturing employment – especially those aimed at reducing imports from China.

A strand of recent empirical works has assessed the differential effect on labor market outcomes of differential shock exposure across local labor markets – see Topalova (2010), Kovak (2013) and Autor et al. (2013) and Muendler (2017), for a review. We explicitly show that the existence of cross-market links gives rise to indirect spillover effects in general equilibrium that and using reduced form empiral analysis may result in an incomplete measurement of the differential and aggregate effect of economic shocks on labor markets. We provide an extension to existing empirical strategies: using the model's predicted local exposure to shocks, one can capture the entire impact of the shocks on local markets in general equilibrium. Furthermore, we show that such general equilibrium channels are empirically relevant in the context of the US regional economy.

Our paper is related to a growing body of literature proposing structural models to measure the aggregate effect of trade and productivity shocks. This literature follows two distinct approaches. On one hand, some papers have proposed general equilibrium models that yield a structural interpretation for the response of regional economies to local shocks – see Kovak (2013), Beraja et al. (2016) and Acemoglu et al. (2017). The simplicity of the expressions for the differential effects is obtained by imposing symmetry assumptions regarding the effect of spatial linkages in general equilibrium, which are common across all regions. On the other hand, several papers carefully specify structural spatial models featuring rich patterns of spatial linkages in general equilibrium – see Allen and Arkolakis (2014), Caliendo et al. (2014), Galle et al. (2015), Caliendo et al. (2015a), Lee (2015) and Burstein et al. (2016), and, for a review, Redding and Rossi-Hansberg (2016). These structural frameworks incorporate great heterogeneity in the general equilibrium spatial linkages across markets. However, given the many primitive assumptions dictating the strength of aggregate linkages across markets, it is often hard to evaluate how robust the model's predictions are. In contrast, we estimate these relationships in an empirical setting that is robust to the underlying microeconomic structure.

The estimation procedure that we propose consists of two steps: i) recovering from the data the structural residuals representing the unobserved production efficiencies and labor supply shifters, ii) using these residuals together with plausibly exogenous variation in the data to implement an efficient estimator of the key structural elasticities. In the first step, we provide a distinct approach in that we can invert structural residuals without requiring the use of a market clearing conditions as proposed in Allen and Arkolakis (2014), Monte et al. (2015), Faber and Gaubert (2016). Past work, such as Eaton and Kortum (2002), suggested

methods of recovering production efficiencies without imposing market clearing conditions, by instead using strong functional form restrictions that leads to log-linear trade shares equations.[6] In the second step, we propose an estimator by applying the Optimal IV approach of Chamberlain (1987) in General Equilibrium. We establish conditions so that our MOIV estimator, effectively an aggregation function over a plausibly exogenous instrument, is the variance minimizing estimator among a set of consistent estimators. The general equilibrium exposure as a consistent instrument has been recently proposed in spatial models by Allen et al. (2014), while partial model-implied approaches have been also explored by Monte et al. (2015), Faber and Gaubert (2016). The approach of Chamberlain (1987) has found some limited use in the Industrial Organization literature (e.g. in the computational approaches by Berry et al. (1999) and Reynaert and Verboven (2014)), whereas our theoretical analysis allows for the complete characterization of the MOIV in general equilibrium, as well as a simple implementation through a Two-Stage GMM methodology.

Finally, notice that our results generalize previous findings in gravity trade and new economic geography models (see Proposition 2 in Arkolakis, Costinot, and Rodríguez-Clare (2012) and Allen, Arkolakis, and Takahashi (2014)) where the bilateral trade, labor supply, and agglomeration mappings take a constant elasticity functional form. Recently, Adao, Costinot, and Donaldson (2017), have discussed equivalence outcomes to representative agent economies for neoclassical models of trade with multiple sectors and factors, but without externalities nor regional labor mobility. Thus, our approach provides a new robust way of estimating the various spatial linkages emphasized in the literature, and use them to credibly measure the general equilibrium effect of global shocks on local labor markets.

This paper is structured as follows. Section 2 describes our Generalized Spatial Competitive Economy and its equivalence to existing models in the literature. Section 3 characterize the impact on local labor markets of shocks in labor supply, productivity, or trade costs. Section 4 describes our novel empirical methodology, which we then implement in Section 5 using US state-level data. Using the theoretical model and the estimates, in Section 6 we conduct a number of counterfactual exercises. Section 7 concludes.

2 Generalized Spatial Competitive Economy

We begin by proposing a general environment to analyze international trade and labor markets. The framework is based on the premise that each location potentially produces

[6] An alternative approach proposed by Adao et al. (2017) allows for a recovery of a origin-destination specific term that can be used without explicitly identifying production efficiencies. With a proper instrument this procedure allows to estimate trade elasticities.

one differentiated commodity, as in Armington (1969) and Anderson (1979), but departs from this setup in three fundamental ways. First, we consider a general utility function across goods, instead of CES preferences. Second, we allow for a flexible structure of labor supply choice across regions and sectors. Third, we incorporate agglomeration and congestion externalities into production that depend on employment in different regions and sectors. These three departures imply that our model is flexible enough to capture the forces in three main strands of the geography and trade literature: i) generalizations of the gravity setup of trade; ii) new economic geography models with labor mobility; iii) trade and geography models incorporating agglomeration and congestion forces in production.

2.1 Environment

We assume that the world economy is constituted of countries, c, that are collections of regions, $o \in \mathcal{R}_c$, with multiple sectors, $s \in \mathcal{K}$. In the rest of the paper, we follow the convention of calling a region-sector pair a ''local market'', and denote origin markets as $i \equiv (s, o)$, and destination markets as j. We denote the vector of outcome x for all markets in the world as $\{x_j\}_j$, and the vector for all markets in country c as $\{x_j\}_{j \in \mathcal{J}_c}$ with $\mathcal{J}_c \equiv \{j = (s, o) \,|o \in \mathcal{R}_c, s \in \mathcal{K}\}$.[7]

Preferences. We assume that in each country there exists a representative agent with preferences for consumption and labor supply in different markets. In particular, the representative agent's preferences are given by

$$U_c \left(\{\nu_j C_j\}_{j \in \mathcal{J}_c}, \{L_j\}_{j \in \mathcal{J}_c} \right), \tag{1}$$

where, in each market j, L_j is the labor supply, C_j is a consumption index, and ν_j is a market-specific preference shifter. The utility function $U_c(\cdot)$ is increasing in consumption and quasi-concave in consumption and labor. Notice that we allow labor supply in different markets to be perfect substitutes in the utility of the representative agent.[8]

The consumption index is an aggregator of the quantities consumed of the differentiated

[7] We treat regions and sectors in a similar way because there is a labor market clearing constraint that holds at that level for all models covered in our analysis. Of course, different assumptions on the degree of mobility of labor across sectors and regions imply different labor supply mappings, affecting wage differentials across markets in equilibrium. We will get back to this point in Section 2.3.

[8] Note that representative household's preferences in (1) does not explicitly impose that the sum of labor supply across markets is constant, allowing our model to capture endogenous changes in aggregate labor supply in the country. The assumption of constant aggregate labor supply is common in the geography literature – see, for example, Allen, Arkolakis, and Takahashi (2014); Allen and Arkolakis (2014); Bartelme (2015). As discussed below, such a restriction can be imposed directly in the specification of the utility function in (1).

goods produced by different origin markets,

$$C_j \equiv V_j\left(\{c_{ij}\}_i\right), \tag{2}$$

where, in market j, c_{ij} is the consumption of the good produced in market i. The function $V_j(\cdot)$ is a consumption aggregator that is assumed to be increasing and quasi-concave in all arguments. Importantly, we restrict $V_j(\cdot)$ to be homogeneous of degree one, so that we can separate the problem of allocating spending shares across origin markets from the problem of determining labor supply across markets in the country.

We assume that consumption in each market is financed by labor income and transfers. Specifically, in each market $j \in \mathcal{J}_c$, the representative agent faces the following budget constraint:

$$\sum_i c_{ij} p_{ij} = \rho_j\left(w_j L_j\right), \tag{3}$$

where w_j is the wage rate per unit of labor, ρ_j is the transfer rate per unit of labor income, and p_{ij} is the price of the good produced in market i. We consider a competitive environment where, in deciding consumption and labor supply, the representative agent takes as given the transfer rate ρ_j, the wage rate w_j and the goods prices $\{p_{ij}\}$.

Production. In each market, there exists a representative firm that operates under perfect competition. Production requires only labor and it is subject to external economies of scale. We assume that there are iceberg trade costs of shipping a good to a different market. The production function of market i to deliver goods to market j is

$$Y_{ij} = \frac{\zeta_i}{\tau_{ij}} \Psi_i\left(\left\{L_j\right\}_j\right) L_{ij} \tag{4}$$

where L_{ij} is the labor input and τ_{ij} is the iceberg trade cost.

In equation (4), $\zeta_i \Psi_i(\cdot)$ is the labor productivity of origin i. The productivity term has two components: ζ_i is an exogenous productivity shifter, and $\Psi_i(\cdot)$ is an endogenous component that is allowed to depend on equilibrium employment levels in *all* markets. This endogenous component of productivity governs agglomeration and congestion forces in our model.

2.2 Competitive Equilibrium

We now proceed to characterize the optimal choices of consumption, labor supply and production levels.

Utility Maximization. We start with the problem of the representative agent in country c, which maximizes the utility in (1) subject to the budget constraint in (3) for every $j \in \mathcal{J}_c$. We solve this problem in two stages. In the first-stage, the representative agent chooses the demand for products from various origins conditional on prices and total expenditure. Using the homotheticity of $V_j(\cdot)$, we define the price index in j as

$$P_j = P_j\left(\{p_{lj}\}_i\right) \equiv \min_{\{c_{ij}\}} \sum_i p_{ij} c_{ij} \quad \text{s.t.} \quad V_j\left(\{c_{ij}\}_i\right) = 1, \tag{5}$$

with the associated spending share on goods from origin i

$$x_{ij} \in X_{ij}\left(\{p_{ij}\}_i\right). \tag{6}$$

The price index and spending shares inherit the usual properties of demand implied by utility maximization. Specifically, $P_j\left(\{p_{lj}\}_i\right)$ is homogeneous of degree one, concave, and differentiable almost everywhere on the vector of prices, $\{p_{lj}\}_i$. The assumption that $V_j(\cdot)$ is quasi-concave implies that $X_{ij}\left(\{p_{ij}\}_i\right)$ is a convex set. Notice also that any $x_{ij} \in X_{ij}\left(\{p_{lj}\}_i\right)$ is a solution of the minimization problem in the definition of $P_j\left(\{p_{lj}\}_i\right)$, so the envelope theorem in Milgrom and Segal (2002) implies that, for any selection x_{ij} satisfying (6), we can write

$$x_{ij} = \frac{\partial \log P_j\left(\{p_{lj}\}_l\right)}{\partial \log p_{ij}}.$$

This relation yields integrability: we can use any selection x_{ij} to recover $P_j(\cdot)$ from $X_{ij}(\cdot)$. Thus, if we know the functional form of $X_{ij}(\cdot)$, we have immediate knowledge of the price index. We make use of this property below.

We now turn to the second-stage problem of choosing employment across markets in the country. We use the solution of the first-stage problem to write the budget constraint (3) as $P_j C_j = \rho_j w_j L_j$, which we then substitute into the consumer's problem to obtain

$$\max_{\{L_j\}_{j\in\mathcal{J}_c}} U_c\left(\{\nu_j \rho_j \omega_j L_j\}_{j\in\mathcal{J}_c}, \{L_j\}_{j\in\mathcal{J}_c}\right)$$

where $\omega_j \equiv w_j / P_j$ is the real wage in market j.

The solution of this problem yields the labor supply in market j:

$$L_j \in \Phi_j\left(\{\nu_i \rho_i \omega_i\}_{i\in\mathcal{J}_c}\right), \tag{7}$$

where $\Phi_j(\cdot)$ is a convex set because of the quasi-concavity of $U_c(\cdot)$. Notice that the labor

supply in market j depends on the entire vectors of real wages, productivity shifters and transfers in the country.

Profit Maximization. Assume that there are traders that can arbitrage any price differentials across markets. Thus, iceberg trade costs imply that

$$p_{ij} = \tau_{ij} p_{ii}. \tag{8}$$

Taking prices and wages as given, the profit maximization problem of the representative firm in i yields

$$p_{ii} \leq \frac{w_i}{\zeta_i \Psi_i \left(\{L_j\}_j \right)} \quad \text{with equality if } L_i > 0.$$

By dividing the two sides of this expression by P_i, we obtain the following price ratio of the cost of the goods supplied by domestic producers, p_{ii}, compared to the domestic cost of the goods supplied by producers around the world, P_i:

$$Q_i \equiv \frac{p_{ii}}{P_i} \leq \frac{\omega_i}{\zeta_i \Psi_i \left(\{L_j\}_j \right)} \quad \text{with equality if } L_i > 0. \tag{9}$$

By revealed preferences, Q_i is directly related to consumption choices and import flows in market i. We exploit this property to connect equation (9) to the data on bilateral trade flows in Section 4.

Market Clearing. To close the model, we specify the labor market clearing condition: in each market, total labor payments must be equal to total revenues,

$$w_i L_i = \sum_j x_{ij} \rho_j w_j L_j. \tag{10}$$

We consider the world equilibrium where transfers $\{\rho_j\}_j$ follow a set of rules that are determined outside of our model. However, we require that any rule guarantees trade balance at the world level, for any vector of wages and prices.[9]

The Competitive Equilibrium of our economy consists of vectors $\{p_{ii}, P_i, L_i, \omega_i\}_i$ such that equations (5)–(10) hold, conditional on a set of transfers rules, $\{\rho_i\}_i$, that satisfy the world-level trade balance and a numeraire, $p_{mm} \equiv 1$, for an arbitrary reference market m.

[9] An obvious special case of this specification is an environment without transfers, with $\rho_j = 1$. The transfer rule of constant transfers suggested by Dekle, Eaton, and Kortum (2007), which has become a staple of quantitative trade models, can be attained by specifying the transfer rule $\rho_j = \frac{T_j}{w_j L_j} + 1$, where T_j is a constant in terms of world GDP. We discuss further examples in conjunction with specific models in the Appendix of Section 2.3.

2.3 Discussion and Equivalences

We now discuss how our theoretical environment unifies a number of existing frameworks in spatial economics. We show that the shape of the mappings $\{X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot)\}_i$ encompasses the central forces in a wide rage of spatial and trade models. We start by introducing a formal definition of the models for which the equilibrium outcomes of the Generalized Spatial Competitive Economy are observationally equivalent to.

Definition 1. The Generalized Spatial Competitive Economy is observationally equivalent to Economy E with respect to the shifters $\{\tau_{ij}, \zeta_i, \nu_i\}_{i,j}$ if

1. There exist *unique* mappings $\left\{X_{ij}^E(\cdot), \Phi_i^E(\cdot), \Psi_i^E(\cdot)\right\}_i$ such that the equilibrium of Economy E is characterized by conditions (6)–(10) for *any* levels of $\{\tau_{ij}, \zeta_i, \nu_i\}_{i,j}$;

2. There exist preferences, (1)–(2), and technology, (4), that imply $\left(X_{ij}^E(\cdot), \Phi_i^E(\cdot), \Psi_i^E(\cdot)\right)$.

This definition requires that, independent of the levels of the exogenous shifters, the equilibrium of Economy E must satisfy conditions (6)–(10) for unique mappings $\left(X_{ij}^E(\cdot), \Phi_i^E(\cdot), \Psi_i^E(\cdot)\right)$. This implies that any combination of shocks to the shifters $\{\tau_{ij}, \zeta_i, \nu_i\}_{i,j}$ yield identical counterfactual outcomes in labor markets. We use Definition 1 to establish that our model is observationally equivalent to several existing frameworks under specific parametric restrictions on the shape of $(X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot))$. We discuss the main equivalence results below by means of some parametric examples, and present the detailed proofs in Appendix D.1.

In the Generalized Competitive Spatial Economy, the mapping $X_{ij}(\cdot)$ controls how bilateral trade costs affect bilateral trade flows. To see this, consider the case in which $X_{ij}(\cdot)$ is a differentiable function, and define the cross-price elasticity $\chi_{lij} \equiv \frac{\partial \log X_{ij}}{\partial \log p_{lj}}$ to represent the response of market j's demand for goods from i to changes in the price of goods from l.[10] To explain how this cross-price elasticity matrix governs the dependence of the demand for labor in a market to changes in production costs in all other markets, we consider the following constant-elasticity gravity form.

Assumption 1a. *The bilateral trade demand is CES with an elasticity structure given by:*

$$\chi_{lij} = -\chi\left(\mathbb{I}_{[i=l]} - X_{ij}\left(\{p_{ij}\}_i\right)\right)$$

where $\mathbb{I}_{[i=l]}$ is an indicator function that equals one if $i = l$.

Assumption 1a is, by far, the most common assumption in trade models. There is a variety of ways to microfound this functional form, starting from simple Armington preferences,

[10] The assumption that the utility function of the representative agent in (1)–(2) is *strictly* quasi-concave and twice differentiable implies that $X_{ij}(.)$ and $\Phi_j(.)$ are differentiable functions.

as in Anderson (1979), to a number of other forms of competition and demand structures discussed in Arkolakis, Costinot, and Rodríguez-Clare (2012), Costinot and Rodriguez-Clare (2013), and Arkolakis et al. (2017). In addition, the generalized mapping we consider is observationally equivalent to the one studied by Adao, Costinot, and Donaldson (2017) with one factor and $\Psi_i(\cdot) = \Phi_i(\cdot) = 1$ for all i.

The function $\Psi_i(\cdot)$ determines how productivity endogenously changes in response to changes in employment in the country. Whenever $\Psi_i(\cdot)$ varies with employment, our model is observationally equivalent to models with Marshallian externalities in production and "New Trade Theory" models with firm entry. The connection between our model and existing frameworks relies on the shape of $\Psi_i(\cdot)$, as summarized by the elasticity matrix, $\psi_{ij} \equiv \frac{\partial \log \Psi_i}{\partial \log L_j}$. To see this, we consider the following parametric restriction.

Assumption 1b. *The agglomeration function has an elasticity structure given by:*

$$\psi_{ij} = \psi \mathbb{I}_{[i=j]} - \tilde{\psi} \Psi_j \left(\{L_o\}_o \right).$$

Again, there are different micro foundations leading to this functional form. In Krugman (1980), firm entry and increasing returns to scale in production give rise to agglomeration forces that depend on the size of the employment in each market – in this case, $\psi = 1/(\sigma - 1)$ and $\tilde{\psi} = 0$ for $i \neq j$, where σ is the elasticity of substitution across varieties. In the same spirit, the specification in Allen and Arkolakis (2014) corresponds to the case of $\psi = \alpha$ and $\tilde{\psi} = 0$. Marshallian external economies of scale, as introduced in Ethier (1982b), correspond to similar specifications, with i denoting a sector in a region. This specification also captures congestion forces arising from other factors in production, such as housing or capital. In this case, production costs are increasing in other market's employment, and thus $\tilde{\psi} > 0$. Furthermore, we show that the combination of the mappings $X_{ij}(\cdot)$ and $\Psi_i(.)$ implies that our framework is observationally equivalent to a general one-factor neoclassical economy with a flexible form of market-level external economies of scale.[11]

Finally, restrictions on $\Phi_j(\cdot)$ allow our model to accommodate different assumptions regarding worker mobility across regions and sectors. Our environment encompasses perfect mobility between any two markets whenever any real wage differential leads to employment

[11] Technology levels that are proportional to population are also postulated in Kortum (1997); Eaton and Kortum (2001). Models with spatial diffusion of knowledge, as in Fujita, Krugman, and Venables (1999) and Lucas and Rossi-Hansberg (2003), specify cross-location spillovers of knowledge so that $\psi_{ij} > 0$ even if $i \neq j$. Our environment can also easily accommodate models with multiple sectors that differ in terms of market structure and strength of economies of scale – e.g., Krugman and Venables (1995), Balistreri, Hillberry, and Rutherford (2010), Kucheryavyy, Lyn, and Rodríguez-Clare (2016).

concentration in the market with the highest real wage. Our model also covers worker immobility between any two markets whenever the labor supply does not depend on the real wage of the other market. Between these two extreme cases, our Generalized Spatial Competitive Economy is observationally equivalent to most of the traditional models of trade and geography in the literature. This is the case for ''New Economic Geography'' models as in Krugman (1991) and Helpman (1998). Similarly, this is the case for traditional trade models, such as the neoclassical setup and more recent quantitative multiple-sector gravity setups surveyed by Costinot and Rodriguez-Clare (2013), where there is mobility across sectors but not across regions.

In addition, the flexibility of $\Phi_j(\cdot)$ allows our model to cover spatial quantitative models exhibiting partial worker mobility across markets. To illustrate the dependence on labor supply on real wages, we consider the case in which $\Phi_j(\cdot)$ is a differentiable function with the following elasticity structure, $\phi_{ij} \equiv \frac{\partial \log \Phi_i}{\partial \log \omega_i}$.

Assumption 1c. *The labor supply mapping has an elasticity structure given by:*

$$\phi_{ij} = \phi \mathbb{I}_{[i=j]} - \tilde{\phi} \Phi_j \left(\{\omega_o\}_o \right).$$

Several recent geography models impose this constant-elasticity gravity structure on labor supply across either sectors or regions. Under specific parametric assumptions, this labor supply structure arises in geography models from either externalities in preference for locations or worker heterogeneity in location-specific preferences or efficiency – e.g., Allen and Arkolakis (2014), Redding (2012), Allen, Arkolakis, and Takahashi (2014), Bartelme (2015) and Bryan and Morten (2015). In recent trade models, this gravity-like labor supply function is implied by the sorting decision of heterogeneous workers that draw sector-specific efficiency from an extreme value distribution – e.g., Galle, Rodriguez-Clare, and Yi (2015).[12] More generally, the unrestricted function $\Phi_j(\cdot)$ implies that our models is observationally equivalent to a generalized Roy model without parametric restrictions on the distribution of market-specific preferences or efficiency units across workers.

Furthermore, combined restrictions on $(X_{ij}(\cdot), \Phi_j(\cdot), \Psi_j(\cdot))$ imply that the Generalized Spatial Competitive Economy is observationally equivalent to existing quantitative spatial models reviewed in Redding and Rossi-Hansberg (2016). The formal equivalence requires transfer rules that specify how income of non-labor factors are allocated across markets. For instance, Allen and Arkolakis (2014) impose a local transfer that is proportional to the

[12] We can also extend our model to allow markets to be region-sector-occupation triples. With this more general definition of a market, the model is also observationally equivalent to environments with gravity-like labor supply function across occupations, as in Burstein et al. (2016) and Lee (2015). See Appendix D.2.

income of the residents in a location, and Caliendo et al. (2014) allow for the possibility that rental income is concentrated in a country portfolio and then split equally to the residents of each location.

Finally, our Generalized Spatial Economy can be easily extended to incorporate additional features highlighted by the trade and geography literature. As we describe in Appendix D.2, we extend the model to allow for: i) workers commuting across regions, as in Ahlfeldt et al. (2015), Monte et al. (2015) and Allen et al. (2015); ii) input-output linkages in production, as in Caliendo and Parro (2014); iii) multiple worker groups in production, as in Cravino and Sotelo (2017); and iv) firm heterogeneity as in Chaney (2008). In all these settings, besides the three aggregate mappings that arise in the Generalized Spatial Competitive Economy, there are other functions determining labor and trade outcomes in equilibrium. Thus, counterfactual predictions in the models considered in the extensions require additional data and mappings compared to our baseline model.

3 Counterfactual Changes in Response to Shocks

We now turn to the counterfactual predictions of our model regarding changes in equilibrium variables following exogenous shocks in local shifters of trade costs, productivity and labor supply. Our objective is to trace down the implications of these shocks as they are propagated through cross-market linkages, and assess their effects on employment and wages in different local labor markets. Subsequently, we evaluate the implications of our theoretical results for the design of empirical specifications aimed to measure the impact of these shocks on local labor markets.

3.1 Counterfactual Changes in Response to Shocks

Our starting point is a characterization of the counterfactual predictions of our model following changes in the exogenous shifters of productivity, labor supply and trade costs. In this section, we use $\hat{z}_j = z'_j/z_j$ to denote the ratio between the variables in the new and the initial equilibrium, and use a superscript 0 to denote endogenous variables in the initial equilibrium. To simplify the exposition, we focus on the particular case in which $\{X_{ij}(.), \Phi_i(.), \Psi_i(.)\}_i$ are differentiable functions, and we henceforth assume that there is positive employment in every market.[13]

[13]These assumptions imply that the equilbrium conditions 6–9 can be written with equality. Without such assumptions, our analysis remains unchanged, except for the fact that we these expressions would be written in terms of sets. We also rely on these assumptions to write a first-order approximation of the model's predictions that highlights its general equilibrium properties.

We define the "local labor market module" as the equilibrium conditions (7)–(9) written in changes:

$$L_i^0 \hat{L}_i = \Phi_i \left(\left\{ \left(\nu_i^0 \rho_i^0 \omega_i^0 \right) \left(\hat{\nu}_i \hat{\rho}_i \hat{\omega}_i \right) \right\}_{i \in \mathcal{J}_c} \right) \tag{11}$$

$$\hat{w}_i = \hat{p}_{ii} \hat{\zeta}_i \frac{\Psi_i \left(\left\{ L_j^0 \hat{L}_j \right\}_j \right)}{\Psi_i \left(\left\{ L_j^0 \right\}_j \right)}. \tag{12}$$

We also need to consider the changes in the remaining equilibrium conditions (5), (6), and (10), which we call the "trade module". The market clearing condition in (10) yields

$$w_i^0 L_i^0 \left(\hat{w}_i \hat{L}_i \right) = \sum_j x_{ij}^0 \hat{x}_{ij} \rho_j^0 w_j^0 L_j^0 \left(\hat{\rho}_j \hat{w}_j \hat{L}_j \right), \tag{13}$$

where the changes in spending shares and price indices in (5)–(6) are given by

$$x_{ij}^0 \hat{x}_{ij} = X_{ij} \left(\left\{ \frac{\tau_{oj}^0 p_{oo}^0}{P_j^0} \hat{\tau}_{oj} \hat{p}_{oo} \right\}_o \right) \quad \text{and} \quad \hat{P}_j = P_j \left(\left\{ \frac{\tau_{oj}^0 p_{oo}^0}{P_j^0} \hat{\tau}_{oj} \hat{p}_{oo} \right\}_o \right). \tag{14}$$

In Appendix A.1, we show that it is always possible to choose the location of the preference and productivity shifters in $\{X_{ij}(.), \Phi_i(.), \Psi_i(.)\}_i$ to replicate the initial levels of trade flows and labor market outcomes across markets, while normalizing prices and trade costs to one in the initial equilibrium.[14] Thus, in equations (11)–(14), we choose initial shifters such that

$$\tau_{ij}^0 p_{ii}^0 \equiv 1, \quad P_j^0 \equiv 1, \quad \nu_j^0 \equiv 1 \quad \forall i, j. \tag{15}$$

The labor and trade modules determine the response of the economy to exogenous changes in $\left\{ \hat{\rho}_i, \hat{\nu}_i, \hat{\zeta}_i, \hat{\tau}_{ij} \right\}$. Given the initial conditions $\left\{ w_j^0, L_j^0, \rho_j^0 \right\}_j$, the local labor market module implies a relation between changes in local competitiveness, $\hat{Q}_i$, and changes in local labor market outcomes, $\hat{\omega}_i$ and $\hat{L}_i$, driven by the cross-market linkages in labor mobility and productivity spillovers embedded in the mappings $\Phi_i(\cdot)$ and $\Psi_i(\cdot)$. In contrast, given the initial variables $\left\{ x_{ij}^0, w_j^0, L_j^0, \rho_j^0 \right\}_{i,j}$, the trade module determines a relation between changes in $\hat{Q}_i = \hat{p}_{ii} / \hat{P}_i$ and changes in market size $\hat{w}_i \hat{L}_i$ that depends on the properties of the bilateral trade network in $X_{ij}(\cdot)$.

To establish our first result, recall that, as discussed above, the integrability property of demand guarantees that knowledge of $\{X_{ij}(\cdot)\}_i$ yields knowledge of $P_j(\cdot)$. Thus, equa-

[14] The normalization of bilateral effective prices in the initial equilibrium is analogous to that imposed in neoclassical economies by Adao, Costinot, and Donaldson (2017).

tions (11)–(15) fully characterize the counterfactual predictions in the Generalized Spatial Competitive Economy.

Proposition 1. *Consider the Generalized Spatial Competitive Economy. Conditional on transfer rules and initial levels of endogenous variables, the mappings* $\{X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot)\}_{i,j}$ *are sufficient to characterize counterfactual changes in endogenous outcomes,* $\left\{\hat{p}_{ii}, \hat{P}_i, \hat{L}_i, \hat{\omega}_i\right\}$, *implied by any combination of shocks,* $\left\{\hat{\zeta}_i, \hat{\nu}_i, \hat{\tau}_{ij}\right\}$, *as a solution of (11)–(14), given the normalization in (15).*

The key insight of this proposition is that, in the class of models covered by the Generalized Spatial Competitive Economy, the researcher does not need to take a stance on the underlying microeconomic assumptions that generate the unobservable shocks. Instead, by conditioning on the initial equilibrium outcomes, $\left\{x^0_{ij}, w^0_j, L^0_j, \rho^0_j\right\}_{i,j}$, one can cleanse the model from any microeconomic assumptions. The microeconomic structure matters only insofar it affects the shape of the mappings $(X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot))$, which themselves can be represented with macroeconomic elasticity functions. This insight motivates our empirical methodology in Section 4 whose goal is to directly estimate these macroeconomic elasticity function using variation in trade and labor outcomes across local labor markets.

The counterfactual predictions of our model depend exclusively on the shape of the mappings $(X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot))$ around the initial equilibrium. These mappings put discipline on the cross-market linkages and, therefore, determine the propagation of shocks between markets. We now turn to a careful investigation of such properties of model.

3.2 Cross-market Effects of Local Exposure to Shocks

We now illustrate how the elasticity structure of $(X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot))$ affects the counterfactual predictions of our model. We consider the first-order approximation of the log-change in endogenous variables following shocks in trade costs, $\hat{\tau}_{ij}$. As discussed in Appendix (A.2), all our insights also hold for local shocks in productivity, $\hat{\zeta}_i$, and labor supply, $\hat{\nu}_i$.

We start by introducing additional notation to simplify the exposition of our results. We use bold variables to denote stacked vector, $\boldsymbol{z} \equiv [z_i]_i$, and bar bold variables to denote matrices, $\bar{\boldsymbol{z}} \equiv [z_{ij}]_{i,j}$. Around the initial equilibrium, we summarize cross-market linkages in the mappings $(X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot))$ with the following elasticity matrices,

$$\bar{\boldsymbol{\chi}} \equiv \left[\sum_d y^0_{id}\chi_{jid}\right]_{i,j}, \quad \bar{\boldsymbol{\phi}} \equiv [\phi_{ij}]_{i,j}, \quad \bar{\boldsymbol{\psi}} \equiv [\psi_{ij}]_{i,j}, \tag{16}$$

as well as the following matrices of initial trade network,

$$\bar{\boldsymbol{y}} \equiv \left[y_{ij}^0\right]_{i,j}, \quad \bar{\boldsymbol{x}} \equiv \left[x_{ji}^0\right]_{i,j}, \tag{17}$$

where y_{ij}^0 denotes the share of market i's revenue from sales to market j, and x_{ji}^0 denotes the share of market i's expenditure on goods from market j.

The elasticity matrices defined in (16) capture responses along the three margins of cross-market linkages in our model: $\bar{\boldsymbol{\chi}}$ is the elasticity of total sales to production costs, $\bar{\boldsymbol{\phi}}$ is the elasticity of labor supply to real wages, and $\bar{\boldsymbol{\psi}}$ is the elasticity of productivity to labor supply. Notice that these elasticity matrices are allowed to depend on the endogenous variables in the initial equilibrium – as in the parametric examples introduced in Section 2.3.

We combine these matrices with trade costs shocks to define market i's direct shock exposure as

$$\log \hat{\tau}_i^I \equiv \sum_j x_{ji}^0 \log \hat{\tau}_{ji}, \text{ and } \log \hat{\tau}_i^E \equiv \sum_j y_{ij}^0 \left(\sum_l \chi_{lij} \log \hat{\tau}_{lj} \right). \tag{18}$$

In equation (18), $\hat{\tau}_i^I$ is the change in the cost of the consumption bundle triggered by the trade cost shocks. It is market i's direct import exposure and it depends solely on spending composition in the initial equilibrium. In addition, $\hat{\tau}_i^E$ is the shift in total sales of market i stemming from the change in the cost of exporting. It is a function of i's revenue exposure to a market interacted with the demand response of that market triggered by the trade cost shocks.

We now use these definitions to characterize the implication of cross-market linkages to the effect of trade shocks on labor markets. To this end, we impose the following assumption that guarantees the existence of a unique local solution for the system of equations (11)–(14).

Assumption 2. *Assume that the following matrices exist:*

$$\begin{aligned} \bar{\boldsymbol{\Lambda}} &\equiv \left(I - \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\phi}}\right)^{-1} \bar{\boldsymbol{\phi}} \\ \bar{\boldsymbol{\Gamma}} &\equiv \bar{\boldsymbol{M}}' \left[\bar{\boldsymbol{M}} \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} - \bar{\boldsymbol{\chi}} + (I - \bar{\boldsymbol{y}}) \left(I + \bar{\boldsymbol{\psi}}\right) \bar{\boldsymbol{\Lambda}} (I - \bar{\boldsymbol{x}})\right) \bar{\boldsymbol{M}}'\right]^{-1} \bar{\boldsymbol{M}} \end{aligned} \tag{19}$$

where $\bar{\boldsymbol{M}}$ is the matrix obtained from deleting the m-th row from the identity matrix.

Assumption 2 has two parts. First, the matrix $\bar{\boldsymbol{\Lambda}}$ exists if $I - \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\phi}}$ is non-singular, which guarantees that $\hat{\boldsymbol{\omega}}$ and $\hat{\boldsymbol{L}}$ can be written as a unique function of $\hat{\boldsymbol{Q}}$ using the equations in the labor market module. In addition to that, the existence of matrix $\bar{\boldsymbol{\Gamma}}$ guarantees a local unique solution for the entire system of equations, (11)–(14). It requires the matrix into square

brackets to have only one linearly dependent column, which arises due to the numeraire choice. The matrix $\bar{\boldsymbol{M}}$ imposes the normalization that the production cost in market m is the numeraire, $\hat{p}_{mm} \equiv 1$.

Under Assumption 2, we obtain a precise characterization of the changes in endogenous outcomes caused by exogenous trade cost shocks. Proposition 2 describes the effect of such shocks on local labor market outcomes.

Proposition 2. *Suppose that Assumption 3.a holds. Then, the impact of trade shocks on local labor markets is, up to a first-order approximation, given by*

$$\log \hat{\boldsymbol{L}} = \bar{\boldsymbol{\Lambda}} \log \hat{\boldsymbol{Q}} \tag{20}$$

$$\log \hat{\boldsymbol{\omega}} = \left(\bar{\boldsymbol{I}} + \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\Lambda}}\right) \log \hat{\boldsymbol{Q}} \tag{21}$$

$$\log \hat{\boldsymbol{Q}} = -\bar{\boldsymbol{\Gamma}}^{I} \log \hat{\boldsymbol{\tau}}^{I} + \bar{\boldsymbol{\Gamma}}^{E} \log \hat{\boldsymbol{\tau}}^{E}. \tag{22}$$

where $\bar{\boldsymbol{\Gamma}}^{I} \equiv \left(\bar{\boldsymbol{I}} - \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \left(I - \bar{\boldsymbol{y}}\right) \left(I + \bar{\boldsymbol{\psi}}\right) \bar{\boldsymbol{\Lambda}}\right)$ *and* $\bar{\boldsymbol{\Gamma}}^{E} \equiv \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}}$.

Proof. Appendix A.2.

Proposition 2 implies that, in general, understanding the impact of trade shocks on local labor markets requires not only measuring the impact of local shocks, but also knowledge of how shocks propagate to all other markets in the world economy. With small shocks, matrix inversions of the elasticity matrices $(\boldsymbol{\chi}, \boldsymbol{\phi}, \boldsymbol{\psi})$ are thus sufficient to measure all these various linkages and associated propagation across markets.

In equations (20)–(21), the matrix $\bar{\boldsymbol{\Lambda}}$ controls the effect of endogenous changes in local competitiveness on employment and real wages across markets. This matrix summarizes the spillovers effects of local competitiveness in its off-diagonal entries – that is, how much competitiveness in one market affects labor market outcomes in other markets. Assuming the spillovers are bounded, the matrix can be written as

$$\bar{\boldsymbol{\Lambda}} = \left(I - \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\phi}}\right)^{-1} \bar{\boldsymbol{\phi}} = \sum_{d=0}^{\infty} \left(\bar{\boldsymbol{\psi}}\bar{\boldsymbol{\phi}}\right)^{d} \bar{\boldsymbol{\phi}},$$

which implies that the strenghts of these spillovers depends on the matrix $\bar{\boldsymbol{\psi}}\bar{\boldsymbol{\phi}}$, which combines cross-market linkages in labor mobility and productivity.[15] Intuitively, these two sources of links create feedback responses arising from interconnected endogenous changes in labor supply and productivity across markets for any given change in local competitiveness.

[15] This is possible only if the spillovers are such that $\sum_j |\sum_o \psi_{io}\phi_{oj}| < 1$.

An even more crucial role in our analysis is played by the transmission of trade shocks through $\hat{Q}_i$. The direct exposure to trade shocks, $\left(\hat{\boldsymbol{\tau}}^E, \hat{\boldsymbol{\tau}}^I\right)$, causes changes in competitiveness in both the own market and other markets as captured, respectively, by the diagonal and the off-diagonal entries of the matrices multiplying $\left(\hat{\boldsymbol{\tau}}^E, \hat{\boldsymbol{\tau}}^I\right)$. Cross-market effects arise from the fact that, in response to direct shock exposure, market clearing requires changes in local competitiveness, which by itself triggers further endogenous responses in trade flows due to substitution and income effects in demand. The definition of $\bar{\boldsymbol{\Gamma}}$ in (19) indicates that the strength of cross-market links is controlled by the elasticity of bilateral trade flows, $\bar{\boldsymbol{\chi}}$, and the initial trade network in $\bar{\boldsymbol{y}}$ and $\bar{\boldsymbol{x}}$. In fact, under parameter restrictions the matrix can be written as an infinite series sum of those elasticity and initial trade network capturing the spatial linkages of locations and the percolation of global shocks through the network of locations.

To further understand the role of cross-market spillovers, we define a benchmark case where such cross-market linkages do not play any role in the economy's adjustment to trade shocks. Specifically, we consider the special case of a shock affecting a ''Small Open Economy'', in which changes in wages and prices do not affect endogenous outcomes in other markets.[16] In such a setup, the direct exposure to trade shocks $\left(\log \hat{\tau}_i^I, \log \hat{\tau}_i^E\right)$ affects only local outcomes, $\left(\log \hat{L}_i, \log \hat{\omega}_i\right)$. That is, in this case, local shocks do not spillover to other markets in general equilibrium. Effectively, all the matrices in expressions (20)-(22) have only diagonal elements, as we show below in Proposition 3.

In the following section, we discuss how to use our theoretical framework to empirically assess the effect of trade cost shocks on local labor markets.

3.3 Measuring the Effect of Local Shocks on Labor Markets

Can variation in exposure to local shocks identify the effect of these shocks on local labor markets? To answer this question, we consider the two different environments introduced above: the Small Open Economy and the Generalzied Spatial Competitive Economy. The following proposition shows the effect of local shock exposure on employment in each of these environments.

Proposition 3. *Consider a trade cost shock. The effect on the local employment is, up to a first-order approximation:*

1) In the Small Open Economy:

[16] We define formally a ''Small Open Economy'' in Appendix A. We say that market d is a small open economy if, for all other markets $i \neq d$, (i) d is a trivial part of revenue and expenditure, $y_{id}^0 \approx 0$ and $x_{di}^0 \approx 0$; and (ii) wages and prices in d do not affect endogenous outcomes, $\chi_{dij} = \phi_{id} = \psi_{id} = 0$.

$$\log \hat{L}_i = \alpha_{ii}^E \log \hat{\tau}_i^E - \alpha_{ii}^I \log \hat{\tau}_i^I, \tag{23}$$

where $\alpha_{ii}^E = (1 - x_{ii}^0)\,\lambda_{ii}\gamma_{ii}$ *and* $\alpha_{ii}^I = \lambda_{ii}\left(1 - (1 - x_{ii}^0)\,\gamma_{ii}\,(1 + \psi_{ii})\,(1 - y_{ii}^0)\,\lambda_{ii}\right)$.

2) In the Generalized Spatial Economy:

$$\log \hat{L}_i = \beta_{ii}^E \log \hat{\tau}_i^E - \beta_{ii}^I \log \hat{\tau}_i^I + \sum_{j \neq i} \left(\beta_{ij}^E \log \hat{\tau}_j^E - \beta_{ij}^I \log \hat{\tau}_j^I\right) \tag{24}$$

where $\beta_{ij}^E = \sum_d \lambda_{id}\gamma_{dj}^E$ *and* $\beta_{ij}^I = \sum_d \lambda_{id}\gamma_{dj}^I$ *are computed from the entries of the matrices* $\bar{\mathbf{\Lambda}}$, $\bar{\mathbf{\Gamma}}$, $\bar{\mathbf{\Gamma}}^E$, $\bar{\mathbf{\Gamma}}^I$.

Equation (23) clearly illustrates how, in a small open economy, the relationship between local shock exposure and changes in local employment yields both the differential and the aggregate impact of trade cost shocks on labor market outcomes. This is the main intuition behind the multiple-sector model that motivates the empirical specification in Autor, Dorn, and Hanson (2013), which outlines the impact of foreign productivity shocks in a small open economy. Thus, in response to a shock, local outcomes in the Small Open Economy can adjust to guarantee local market clearing without any impact on goods and labor markets in the world economy. We refer to the responses in Proposition 3 as the ''direct effect'' of a trade cost shock.

Any heterogeneity in expenditures and revenues induces heterogeneity in the indirect effects, so that the analysis complicates substantially when we consider the generalized economy with all the spatial lnks. In fact, equation (24) shows that, in the general case, employment in market i responds not only to local trade shocks, $\left(\hat{\tau}_i^E, \hat{\tau}_i^I\right)$, but also to the trade shocks hitting other markets, $\left(\hat{\tau}_j^E, \hat{\tau}_j^I\right)$ for $j \neq i$. Such indirect spillover effects cannot be disregarded in an empirical analysis of the effect of trade shocks on local labor markets. They constitute an additional exposure channel of market i to trade shocks that, if omitted, introduces bias in the measurement of both the differential and aggregate effects of trade shocks.

A restriction on the elasticity functions to take the form postulated in Assumptions 1a-1c provides an illustration of how the spatial linkages across markets create indirect effects that affect both the differential and the aggregate impact of global shocks across local labor markets.

Corollary 1. *Consider a trade cost shock. The effect on the local employment in the*

parametric gravity model is, up to a first-order approximation:

$$\log \hat{L}_i = \beta^1 \sum_j \gamma_{ij}^E \log \hat{\tau}_j^E + \beta^2 \sum_j \gamma_{ij}^I \log \hat{\tau}_j^I + \sum_j \left(\tilde{\beta}_j^1 \sum_j \gamma_{ij}^E \log \hat{\tau}_j^E + \tilde{\beta}_j^2 \sum_j \gamma_{ij}^I \log \hat{\tau}_j^I \right), \tag{25}$$

where the constants β^i *depend on* $\epsilon, \phi, \psi, \tilde{\phi}, \tilde{\psi}$*, and the constants* β_j^i *depend on* L_j *as well. When* $\tilde{\phi} = \tilde{\psi} = 0$ *we have* $\tilde{\beta}_j^1 = \tilde{\beta}_j^2 = 0$*.*

In this case, the labor supply and productivity links in Assumptions 1b-1c appear through a symmetric component captured by the parenthetical term in equation (25). As the proposition illustrates, the parenthetical term disappears if $\tilde{\phi} = \tilde{\psi} = 0$ and, thus, part of the indirect effects will be differenced out, and thus cannot be measured with regression specifications with time fixed effects. Even in this case, however, asymmetric links across markets imply exposure that affects the measurement of differential effects of trade shocks across markets. Thus, due to between-region trade links, these empirical specifications may still yield an imprecise measurement of the differential effects without the proper cross-market exposure weights, γ_{ij}^E and γ_{ij}^I.

Taken together, our results indicate that any investigation of the complete effect of trade shocks on local labor markets requires a careful assessment of cross-market links in terms of productivity, labor supply and bilateral trade and knowledge of the aggregate mappings controlling cross-market links.

4 Econometric methodology

We now turn to the problem of estimating the central aggregate mappings in our model. In particular, we focus on the estimation of the elasticity structure of labor supply, $\Phi_i(\cdot)$, and agglomeration, $\Psi_i(\cdot)$. Throughout our analysis, we assume that the elasticity for bilateral trade flows, $X_{ij}(.)$, is known, since its estimation has been the goal of an extensive literature in international trade.[17]

We develop our methodology in three steps. First, we propose an econometric version of our model where $\Phi_i(\cdot)$ and $\Psi_i(\cdot)$ depend on a vector of unknown parameters. In this

[17] In single-sector gravity models, the demand for bilateral trade flows only depends on the trade elasticity that has been studied by an extensive empirical literature – for a review, see Head and Mayer (2013). In addition, Caliendo and Parro (2014) and Costinot, Donaldson, and Komunjer (2011) consider multiple-sector gravity models where these functions only depend on the sector-level trade elasticity that is estimated using sector-level bilateral trade flows. More recently, Adao, Costinot, and Donaldson (2017) consider the problem of non-parametrically identifying the functions controlling bilateral trade flows in a competitive environment. It is possible to show that a similar argument holds in our environment, leading to the non-parametric identification of $X_{ij}(.)$.

parametric model, we establish that, conditional on $\Phi_i(\cdot)$ and $\Psi_i(\cdot)$, local unobserved shocks in productivity and labor supply are identified from observable data on trade and labor outcomes. Second, we construct a class of moment conditions using observed trade cost shifters that are orthogonal to local shocks in productivity and labor supply. Finally, we use our general equilibrium model to show that the ''optimal'' instrument in this class is the impact of *observed* trade shocks on the *endogenous* variables predicted by our general equilibrium model – the Model-implied Optimal IV (MOIV).

4.1 Parametric Econometric Model

In every period t, we assume that the world economy is generated by the Generalized Spatial Competitive Economy described in Section 2. Specifically, we impose that the aggregate mappings $\{X_{ij}(\cdot), \Phi_i(\cdot), \Psi_i(\cdot)\}_{i,j}$ are the same in all periods, but markets are subject to changes in the exogenous shifters, $\{\nu_i, \zeta_i, \tau_{ij}\}$. In equilibrium, endogenous trade and labor outcomes are determined by the solution of (6)–(10).

Let y^t denote the value of variable y in period t, and $\hat{y}^t = y^t/y^0$ denote its change between a base period 0 and period t, with $\Delta \log y^t \equiv \log \hat{y}^t$. Assume that we observe bilateral trade flows, $\left\{X_{ij}^t\right\}$, and labor market outcomes, $\{\rho_i^t, \omega_i^t, L_i^t\}$, but we do not observe the shifters, $\left\{\nu_i^t, \zeta_i^t, \tau_{ij}^t\right\}$.

We start by introducing restrictions on the functions controlling labor supply, $\Phi_j(.)$, and agglomeration forces, $\Psi_j(.)$. We assume that these functions are known, except for a vector of unknown parameters, $\boldsymbol{\theta}$. Formally, we impose the following assumption.

Assumption 3a. $\Phi_j(.) = \Phi_j(.|\boldsymbol{\theta})$ *and* $\Psi_j(.) = \Psi_j(.|\boldsymbol{\theta})$ *are differentiable functions on the vector of unknown parameters* $\boldsymbol{\theta} \in \mathbb{R}^s$.

In the rest of this section, we use our general equilibrium model to propose a methodology to consistently estimate $\boldsymbol{\theta}$. To this end, we impose additional restrictions in our model that allow the recovery of the unobserved local shocks, $\left\{\hat{\nu}_i^t, \hat{\zeta}_i^t\right\}$, from the observed changes in trade and labor market outcomes, $\left\{x_{ij}^t, \omega_i^t, L_i^t, \rho_i^t\right\}_i$, and the observed equilibrium variables in the base period, $\boldsymbol{W}^0 \equiv \left\{X_{ij}^0, L_j^0\right\}_{i,j}$. We consider the following invertibility assumption.

Assumption 3b. $\Phi_j(.)$ *and* $X_{ij}(.)$ *are functions such that*

1. $\{L_j\}_{j\in\mathcal{J}_c} = \{\Phi_j(\boldsymbol{\omega}|\boldsymbol{\theta})\}_{j\in\mathcal{J}_c}$ *is invertible,*
2. $\{x_{ij}\}_i = \left\{X_{ij}\left(\{p_{oj}\}_o\right)\right\}_i$ *is invertible (up to a scalar) for all* $j \in \mathcal{J}_c$.

The first part of Assumption 3b imposes that the system $\{L_j\}_{j\in\mathcal{J}_c} = \{\Phi_j(\boldsymbol{\omega}|\boldsymbol{\theta})\}_{j\in\mathcal{J}_c}$ is invertible, so we can write $\omega_j = \Phi_j^{-1}(\boldsymbol{L}|\boldsymbol{\theta})$.[18] The combination of this restriction with the labor supply equation in (7) yields the change in labor supply shifters:

$$\Delta \log \nu_j^t = -\Delta \log \rho_j^t - \Delta \log \omega_j^t + \Delta \log \Phi_j^{-1}\left(\boldsymbol{L}^t|\boldsymbol{\theta}\right). \tag{26}$$

Moreover, our methodology requires a measure of the local competitiveness, $\hat{Q}_i$. This implies either taking an explicit stance on the price data to measure p_{ii}, or using data on trade shares and invert the function $X_{ij}(.)$ to obtain prices. Since our model yields equivalent counterfactual outcomes to a wide class of existing frameworks, one should be cautious in using the first method. In fact, different theories about competition and firm behavior may lead to different disaggregated prices, as argued by Arkolakis, Costinot, and Rodríguez-Clare (2012) and Simonovska and Waugh (2014b), and thus different measured aggregate prices. Thus, in the second part of Assumption 3b, we effectively impose that it is possible to recover relative prices from observed trade flows: $p_{oj}^t/p_{jj}^t = X_{oj}^{-1}\left(\boldsymbol{x}_j^t\right)$ where $\boldsymbol{x}_j^t \equiv \left\{x_{ij}^t\right\}_i$.[19]

Recalling that, as in Section 2, knowledge of $X_{ij}(.)$ yields knowledge of $P_j(.)$, we can write local competitiveness, $Q_i^t \equiv p_{ii}^t/P_i^t$, as a function of the vector of local spending shares:

$$Q_j\left(\boldsymbol{x}_j^t\right) = \left[P_j\left(\left\{p_{ij}^t/p_{jj}^t\right\}_i\right)\right]^{-1} = \left[P_j\left(\left\{X_{ij}^{-1}\left(\boldsymbol{x}_j^t\right)\right\}_i\right)\right]^{-1}.$$

By applying this expression into (9), we obtain the change in productivity shifters:

$$\Delta \log \zeta_j^t = \Delta \log \omega_j^t - \Delta \log Q_j\left(\boldsymbol{x}_j^t\right) - \Delta \log \Psi_j\left(\boldsymbol{L}^t|\boldsymbol{\theta}\right). \tag{27}$$

Given the parameter vector $\boldsymbol{\theta}$, expressions (26)–(27) relate changes in the unobserved shocks, $\left\{\nu_j^t, \zeta_j^t\right\}$, to changes in observed variables, $\left\{\boldsymbol{x}_j^t, \omega_j^t, L_j^t, \rho_j^t\right\}$. We summarize this relationship in the following expression:

[18] The invertibility of the structural residuals is a crucial step in many empirical structural frameworks – see Berry (1994); Berry and Haile (2014). In the Generalized Spatial Competitive Economy, invertibility is guaranteed if the utility function has following separable form:

$$U_c = \sum_j \nu_j C_j + U\left(\{L_j\}\right)$$

with $U(.)$ strictly quasi-concave. In this case, the labor supply function is the unique solution of the system of first-order conditions: $\nu_j \rho_j \omega_j = \frac{\partial U}{\partial L_j}$. We can also consider the weaker condition of invertibility up to a scalar when the labor supply function is also homogeneous of degree zero, as in spatial assignment models.

[19] The demand function $X_{ij}(.)$ is invertible (up to scalar) if it satisfies the connected substitutes property – see Berry, Gandhi, and Haile (2013). Adao, Costinot, and Donaldson (2017) show that the invertibility of $X_{ij}(.)$ is a central property for its non-parametric identification, and that it is guaranteed in a generalized class of Ricardian economies.

$$\begin{bmatrix} \Delta \log \nu_j^t \\ \Delta \log \zeta_j^t \end{bmatrix} = - \begin{bmatrix} \Delta \log \rho_j^t + \Delta \log \omega_j^t \\ \Delta \log Q_j \left(\boldsymbol{x}_j^t \right) - \Delta \log \omega_j^t \end{bmatrix} + \begin{bmatrix} \Delta \log \Phi_j^{-1} \left(\boldsymbol{L}^t | \boldsymbol{\theta} \right) \\ -\Delta \log \Psi_j \left(\boldsymbol{L}^t | \boldsymbol{\theta} \right) \end{bmatrix}. \tag{28}$$

Expression (28) is our main estimating equation: it yields the underlying local shocks from observable data on trade and labor outcomes. This, of course, relies on two premises. First, because our approach is parametric, it requires the correct functional forms of the mappings $\Phi_j(.|\boldsymbol{\theta})$ and $\Psi_j(.|\boldsymbol{\theta})$. Second, since expression (28) follows directly from the equilibrium conditions (7) and (9), these conditions must specify correctly labor supply and production decisions in the economy. Violations of these premises introduce additional endogenous variables in (28) that give rise to usual concerns regarding omitted variable bias in the estimation of $\boldsymbol{\theta}$.[20]

To estimate $\boldsymbol{\theta}$ using expression (28), notice that, in general equilibrium, local trade and local labor outcomes are correlated with the unobserved local shocks. In fact, $\left\{ \boldsymbol{x}_j^t, \omega_j^t, L_j^t, Q_i^t \right\}$ are endogenous variables that depend on exogenous shocks, $\left\{ \hat{\nu}_i^t, \hat{\zeta}_i^t, \hat{\tau}_{ij}^t \right\}$. Thus, we design an instrumental variables methodology that exploits the structure of our model to construct moment conditions for the consistent estimation of $\boldsymbol{\theta}$.

4.2 Model-implied Optimal IV

We now derive moment conditions for the estimation of $\boldsymbol{\theta}$ using the recovered error terms in (28) and observed shifters of bilateral trade costs. To this end, assume that we have an observable variable, $\hat{\boldsymbol{z}}^t \equiv \left\{ \hat{z}_{ij}^t \right\}$, that satisfies the following conditions.

Assumption 3c. *There exists a trade cost shifter,* $\hat{\boldsymbol{z}}^t \equiv \left\{ \hat{z}_{ij}^t \right\}$, *such that*

1. $E \left[\Delta \log \tau_{ij}^t | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 \right] = \kappa \Delta \log z_{ij}^t$,
2. $E \left[\Delta \log \nu_j^t | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 \right] = E \left[\Delta \log \zeta_j^t | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 \right] = 0$.

Assumption 3c imposes two conditions on the observable shifter $\hat{\boldsymbol{z}}^t$. First, changes in trade costs in the model, $\hat{\tau}_{ij}^t$, are log-linearly related to changes in the observable shifter,

[20] Any parametric empirical approach is subject to similar concerns. In our case, the separability of local unobserved shocks in equation (28) allows the nonparametric identification of the functions $(\Psi_j(.), \Phi_j(.))$ as long as instrumental variables satisfy the completeness condition proposed by Newey and Powell (2003) – for similar strategies, see Berry and Haile (2014), and Adao, Costinot, and Donaldson (2017). Expression (28) is subject to misspecification of the channels determining labor supply and production costs in conditions (7) and (9). For instance, as discussed in Appendix D.2, extensions of our model introduce extra variables into these conditions and, therefore, imply modified versions of (28). However, expression (28) is robust to misspecification in other equilibrium conditions of the model and, therefore, relies on weaker assumptions than those necessary to recover local shocks using the economy's full general equilibrium structure – for example, see Bartelme (2015); Allen, Arkolakis, and Takahashi (2014).

$\hat{z}_{ij}^t$. This log-linearity restriction significantly simplifies the conditions for optimality of our methodology. However, it is not necessary to compute our instrument, as we discuss below.

In addition, the second part of the Assumption 3c states that, conditional on the vector of endogenous variables in the initial world equilibrium $\boldsymbol{W}^0$, the trade cost shifter $\hat{\boldsymbol{z}}^t$ is mean-independent from local shocks to productivity and labor supply. It is important to note that this exogeneity condition is similar to the assumption imposed in empirical papers in international trade, such as Topalova (2010), Kovak (2013), and Pierce and Schott (2016). In the empirical application below, we will use worldwide tariff changes as observable shifters of bilateral trade costs that are mean-independent from local shocks.[21]

To construct moment conditions, we introduce a function capturing the exposure of each market to the observable trade cost shock: $H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right)$. By the law of iterated expectations, Assumption 3c immediately implies that, for any function $H_i(\cdot)$,

$$E\left[H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right) e_i^t\right] = 0. \tag{29}$$

where $e_i^t \equiv \left(\Delta \log \nu_j^t, \Delta \log \zeta_j^t\right)'$.

The moment condition in (29) yields the following class of GMM estimators.

Definition 2. *Let $H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right)$ be a $S \times 2$ matrix of functions. The GMM estimator is*

$$\hat{\boldsymbol{\theta}}_H \equiv \arg min_{\boldsymbol{\theta}} \left[\sum_{i,t} H_i(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0) e_i^t(\boldsymbol{\theta})\right]' \left[\sum_{i,t} H_i(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0) e_i^t(\boldsymbol{\theta})\right] \tag{30}$$

where $e_i^t(\boldsymbol{\theta}) \equiv \left(\Delta \log \nu_j^t, \Delta \log \zeta_j^t\right)'$ defined in equation (28).

Definition 2 outlines a standard GMM estimator based on the moment condition in (29). Notice that the dimension of $H_i(\cdot)$ is such that the number of moments is equal to the number of parameter in $\boldsymbol{\theta}$. This does not imply any loss of generality, since it is always possible to define $H_i(\cdot)$ to include the optimal weighting matrix if there are multiple shifters of trade costs. Under standard regularity conditions, $\hat{\boldsymbol{\theta}}_H$ converges in probability to the true parameter vector, $\boldsymbol{\theta}$, and has an asymptotically normal distribution.[22]

[21] Our choice of instrument differs from other work in that we do not use trade and labor outcomes to invert exogenous shifters for estimation. Therefore we avoid imposing that the entire vector of endogenous trade flows is mean independent from unobserved local shocks, as in Autor, Dorn, and Hanson (2013). Similarly, we avoid using the full general equilibrium structure of the model alongside with trade and labor outcomes to invert instrumental variables, as in Monte, Redding, and Rossi-Hansberg (2015); Faber and Gaubert (2016); Allen and Arkolakis (2013b); Bartelme (2015); Allen and Arkolakis (2013a). By relying on the entire general equilibirum structure, misspecification in any part of the model is a threat to identification. In contrast, under Assumption 3c, our approach is only subject to the misspecification of equation (28).

[22] Newey and McFadden (1994) provide regularity conditions for consistency and normality of GMM

Although any exposure function $H_i(\cdot)$ yields a consistent estimator of $\boldsymbol{\theta}$, different exposure functions $H_i(\cdot)$ vary in terms of asymptotic variance — that is, the estimators differ in terms of precision. To choose the exposure function $H_i(\cdot)$, we follow the approach in Chamberlain (1987) and select the one minimizing the asymptotic variance of the estimator. Applying the result in Chamberlain (1987), we show in Appendix B.1 that the most efficient estimator in the class of estimators in Definition 2 is

$$H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0) \equiv E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta}) | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right] \left(\Omega_i^t\right)^{-1} \tag{31}$$

where $\Omega_i^t \equiv E\left[e_i^t(\boldsymbol{\theta}) e_i^t(\boldsymbol{\theta})' | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right]$.

The optimal exposure function in (31) has two components. The matrix Ω_i^t attributes higher weight to observations with a lower variance of local unobserved shocks – if local shocks are homoskedastic, $\Omega_i^t = \Omega$. The optimal exposure function also depends on $E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta}) | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right]$: the expected response of the endogenous variables associated with $\boldsymbol{\theta}$ induced by the exogenous trade cost shifter $\hat{\boldsymbol{z}}^t$. The definition of the structural residuals in (28) implies that, up to a first-order approximation,

$$E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta}) | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right] \approx \nabla_{\boldsymbol{\theta}} \boldsymbol{B}\left(\boldsymbol{\theta}, \boldsymbol{W}^0\right) E\left[\log \hat{\boldsymbol{L}}^t | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right], \tag{32}$$

where $\boldsymbol{B}\left(\boldsymbol{\theta}, \boldsymbol{W}^0\right) \equiv \left(\bar{\boldsymbol{\phi}}^{-1}(\boldsymbol{\theta}, \boldsymbol{W}^0),\ \bar{\boldsymbol{\psi}}(\boldsymbol{\theta}, \boldsymbol{W}^0)\right)$. Note that, in our model, $E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta}) | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right]$ depends only on the expected change in the employment but, in another settings, it could be a function of other endogenous variables.

We use our general equilibrium model to approximate the expected change in employment conditional on the change in the trade cost shifter. The combination of Assumption 3c and the log-linear version of our model in Proposition 2 implies that

$$E\left[\log \hat{\boldsymbol{L}}^t | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right] \approx \kappa \log \hat{\boldsymbol{L}}^p\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 | \boldsymbol{\theta}\right) \tag{33}$$

where $\hat{\boldsymbol{L}}^p(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 | \boldsymbol{\theta})$ is the "exact hat-algebra" solution of the system (11)–(14) given the observed trade cost shock, $\hat{\boldsymbol{\tau}}^t = \hat{\boldsymbol{z}}^t$, while holding constant all other shocks, $\log \hat{\nu}_j^t = \log \hat{\zeta}_j^t = 0$.

The expression in equation (33) is the first-stage implied by our general equilibrium model: $\hat{\boldsymbol{L}}^p(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 | \boldsymbol{\theta})$ is the predicted impact of the trade cost shifter, $\hat{\boldsymbol{z}}^t$, on the endogenous changes in employment, $\hat{\boldsymbol{L}}^t$. Through the lens of our model, $\hat{\boldsymbol{L}}^p(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 | \boldsymbol{\theta})$ is the Optimal IV since

estimators of the form in Definition 2 – for consistency, see Theorems 2.6–2.7 and, for normality, see Theorem 3.4. Such regularity conditions require $\boldsymbol{\theta}$ to be the unique solution of the moment condition in (29). This is implied by usual rank conditions establishing identification in GMM estimators. For instance, if $\Delta \log \nu_i^t(\boldsymbol{\theta})$ and $\Delta \log \zeta_i^t(\boldsymbol{\theta})$ are linear in $\boldsymbol{\theta}$, then uniqueness requires $H_i(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0)$ to be correlated with the endogenous variables multiplying $\boldsymbol{\theta}$.

it is the best predictor of changes in employment induced by $\hat{\boldsymbol{z}}^t$. Intuitively, $\hat{\boldsymbol{L}}^p(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta})$ captures all the channels through which $\hat{\boldsymbol{z}}^t$ affects endogenous labor market outcomes across markets, and thus leads to the most precise estimates.

This discussion is summarized by the following proposition:

Proposition 4. *Suppose the world economy is generated by the model in Section 2, satisfying Assumptions 3a–3c. The function $H_i^*(\cdot)$ in (31) determines the estimator with the minimum asymptotic variance in the class of estimators in Definition 2, and is approximately given by*

$$E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta})|\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right] \approx \kappa \nabla_{\boldsymbol{\theta}} \boldsymbol{B}\left(\boldsymbol{\theta}, \boldsymbol{W}^0\right) \log \hat{\boldsymbol{L}}^p\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}\right). \tag{34}$$

We now characterize a simple two-stage procedure that is asymptotically equivalent to the MOIV estimator. Using a guess of the structural parameters, we compute the predicted changes in endogenous variables, $\log \hat{\boldsymbol{L}}^p\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}^0\right)$, and the the instrumental variable, $H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}^0\right)$. In the first-stage, we use this instrument to estimate the model, leading to estimates $\hat{\boldsymbol{\theta}}^1$. Since the instrument is a function of $\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right)$, the first-stage estimator is a consistent estimator of $\boldsymbol{\theta}$. In the second-stage, we use $\hat{\boldsymbol{\theta}}^1$ to compute the instrument, $H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\hat{\boldsymbol{\theta}}^1\right)$, and estimate the model again, leading to estimates $\hat{\boldsymbol{\theta}}^2_{MOIV}$. The following proposition establishes that the asymptotic properties of this two-stage estimator are identical to those of the estimator based on (34).

Proposition 5. *The Model-implied Optimal IV estimator is asymptotically equivalent to the estimator obtained from the following two-stage procedure.*

Stage 1. *For an initial guess for the parameter vector $\boldsymbol{\theta}^0$,*

$$\hat{\boldsymbol{\theta}}^1 = \arg min_{\tilde{\boldsymbol{\theta}}} \left[\sum_{i,t} H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}^0\right) e_i^t\left(\tilde{\boldsymbol{\theta}}\right)\right]' \left[\sum_{i,t} H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}^0\right) e_i^t\left(\tilde{\boldsymbol{\theta}}\right)\right] \tag{35}$$

Stage 2. *Using the first-stage estimates $\hat{\boldsymbol{\theta}}^1$,*

$$\hat{\boldsymbol{\theta}}^2_{MOIV} = \arg min_{\tilde{\boldsymbol{\theta}}} \left[\sum_{i,t} H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\hat{\boldsymbol{\theta}}^1\right) e_i^t\left(\tilde{\boldsymbol{\theta}}\right)\right]' \left[\sum_{i,t} H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\hat{\boldsymbol{\theta}}^1\right) e_i^t\left(\tilde{\boldsymbol{\theta}}\right)\right]. \tag{36}$$

Proof. *Appendix B.2.*

Having characterized the optimal estimator for our Generalized Economy, we now use such estimator and our framework to empirically assess the impact of trade shocks on local labor markets.

5 Estimation of cross-market linkages in the US

Our theoretical results describe the importance of cross-market linkages in measuring the effect of shocks on local labor markets. We have developed a methodology to estimate such linkages from cross-market variation in trade and labor outcomes induced by exogenous shocks to observable trade cost shocks. We now use our theoretical results to evaluate the importance of cross-market effects of local shocks among states in the United States.

5.1 Multiple-Sector Spatial Model

The first step of our methodology is the parametrization of the Generalized Spatial Competitive Economy. We introduce a multiple-sector version of the spatial gravity model discussed in Section 2.3. Our parametric model is flexible enough to capture the main forces in the quantitative spatial models reviewed by Redding and Rossi-Hansberg (2016). Yet, it remains tractable: the strength of cross-market linkages depend on a parsimonious number of parameters. In this section, we use k, s to denote sectors, and o, d to denote regions.

Bilateral Trade Flows. We follow the extensive literature on recent quantitative trade models reviewed in Costinot and Rodriguez-Clare (2013) by assuming that bilateral trade flows exhibit a multiple-sector gravity structure. We impose the nested CES preferences for goods from different sectors and regions presented in Appendix C.1. Thus, conditional on bilateral good prices $\{p_{ko,sd}\}_{ko}$, the share of expenditure on goods from sector k of region o by workers in sector s of region d is

$$x_{ko,sd}^t = \left(\frac{p_{ko,sd}^t}{P_{k,sd}^t}\right)^{1-\chi_k} \frac{\left(P_{k,sd}^t\right)^{1-\chi}}{\sum_{k'}\left(P_{k',sd}^t\right)^{1-\chi}} \tag{37}$$

where the sector-level price index is $P_{k,sd}^t = \left[\sum_o \left(p_{ko,sd}^t\right)^{1-\chi_k}\right]^{\frac{1}{1-\chi_k}}$. In equation (6), χ_k is the elasticity of substitution across origins in sector k, and χ is the elasticity of substitution across sectors.

An important step of our empirical strategy is the inversion of local competitiveness from observed trade outcomes. The expenditure shares in (37) implies that

$$Q_{ko}^t = \left(x_{kr,kr}^t\right)^{\frac{1}{1-\chi_k}} \left(\sum_o x_{ko,kr}^t\right)^{\frac{1}{1-\chi}-\frac{1}{1-\chi_k}}. \tag{38}$$

A necessary and sufficient condition for the invertibility of the trade demand in (37) is $\chi_k \neq 1$ and $\chi \neq 1$. This rules out an unitary elasticity of substitution across good from different sectors and region, in which case spending shares are not sensitive to changes in

bilateral prices. It is important to notice that the workhorse multiple-sector gravity model imposes exactly this knife-edge case: the upper-level preference aggregator is Cobb-Douglas. This is a simplifying assumption that, as argued in Costinot and Rodriguez-Clare (2013), has important quantitative implications. In our empirical application, we follow a strand of recent papers that estimate the elasticity of substitution across sectors to be below one – e.g., Herrendorf, Rogerson, and Valentinyi (2013), Comin, Lashkari, and Mestieri (2015), and Cravino and Sotelo (2017).

Labor Supply. We assume that, as described in Appendix C.1, the representative household has a nested utility function over the number of worked hours in a sector-region pair, h_{sd}^t, and the allocation of labor across sector-region pairs, l_{sd}^t. We incorporate non-employment in the model by introducing a home sector, $s = 0$, that yields an identical level of outside utility in all regions. Such a preference structure implies three margins of employment responses:

$$\Delta \log h_{sd}^t = \phi_h \Delta \log \left(\rho_{sd}^t \omega_{sd}^t\right) + \nu_{sd}^t, \tag{39}$$

$$\Delta \log \left(l_{sd}^t / l_{0d}^t\right) = \phi_e \Delta \log \left(\rho_{sd}^t \omega_{sd}^t\right) + \tilde{\nu}_{sd}^t, \tag{40}$$

$$\Delta \log \left(n_d^t / n_o^t\right) = \phi_m \Delta \log \left(l_{0d}^t / l_{0o}^t\right) + \tilde{\nu}_d^t, \tag{41}$$

where $n_d^t \equiv \sum_s l_{sd}^t$ is the share of national population in region d, and l_{0d}^t is the share of non-employed individuals residing in region d.

Equations (39) and (40) specify the choices of hours and relative employment in the different sectors of region d. Following changes in the real wage, the parameters ϕ_h and ϕ_e regulate labor supply responses in hours worked and sectoral employment shares. These parameters capture two main margins of labor supply adjustment emphasized in the literature reviewed by Chetty et al. (2013): ϕ_h is the intensive labor supply elasticity and ϕ_e is the extensive labor supply elasticity. In addition, equation (41) characterizes the allocation of labor across regions. The parameter ϕ_m regulates the elasticity of relative labor supply to changes in relative non-participation in any two regions of the country.[23]

Technology. We impose two main restrictions on the production structure of the Generalized Spatial Competitive Economy. The first assumption is that the cost of shipping

[23]Our parametric specification encompasses recent models with endogenous choices of employment across sectors and regions. Whenever $\phi_m = 0$, our labor supply structure is isomorphic to that implied by a Roy models with Frechet distribution of sector-specific efficiency and prefereces – such as Galle, Rodriguez-Clare, and Yi (2015). In addition, if $\phi_m \to \infty$, our labor supply structure is equivalent to a static version of the model in Caliendo, Dvorkin, and Parro (2015a), which impose the same elasticity of relative employment across sectors and regions. It is also possible to show that our environment is isomorphic to a model where heterogeneous individuals maximize a separable quasi-linear utility function in consumption goods and labor, and draw idiosyncratic preferences for sectors and regions from a Generalized Extreme Value distribution, as in McFadden (1980).

goods from sector k of region o to *any* sector of region d is the same. That is, we assume that

$$\tau^t_{ko,sd} = \tau^t_{ko,d} \quad \forall s.$$

When combined with the bilateral trade demand in (37), this restriction implies that, for *any* sector s of region d, spending shares are identical: $x^t_{ko,sd} = x^t_{ko,d}$ for all s. This is an important assumption for our empirical application given the lack of expenditure data for individuals employed in different sectors of regional labor markets.

We also need to specify the function form governing agglomeration and congestion forces in the model. Specifically, we assume that productivity in sector k of region o is

$$\Psi_{sd} = \left(h^t_{sd} l^t_{sd}\right)^{\psi_e} \left(n^t_d\right)^{\psi_m}.$$

The combination of this expression with equation (9) yields

$$\Delta \log\left(\omega^t_{sd}/Q^t_{sd}\right) = \psi_e \Delta \log\left(h^t_{sd} l^t_{sd}\right) + \psi_m \Delta \log\left(n^t_d\right) + \zeta_{sd} \tag{42}$$

where Q^t_{sd} is given by (38).

Expression (42) entails two parameters to separately control the strength of agglomeration/congestion forces at the sector and region levels. While the parameter ψ_e regulates the effect arising from sectoral employment concentration in the region, the parameter ψ_m regulates the effect of spatial population concentration. As discussed in Section 2.3, our specification combines two important features of spatial models through these parameters. Specifically, it captures local productivity gains arising from knowledge spillovers, increasing economies of scale, and firm entry. These agglomeration benefits are balanced by congestion costs arising from the effect of employment concentration on the price of local factors in limited supply – for example, land or infrastructure, as highlighted in Redding and Rossi-Hansberg (2016).

5.2 Data

To apply our methodology, we construct a dataset of trade and labor outcomes for two sectors in US states between 1993 and 2012. We now describe the main variables in our analysis, and discuss the details of the data construction in Appendix C.2.

Trade Cost Shock. Our methodology requires trade cost shifters that are orthogonal to local shocks in productivity and labor supply affecting regional markets in the United States. In our empirical application, we use worldwide changes in import tariffs between 1993 and 2012. Tariff declines are arguably exogenous to regional shocks for two main reasons. First,

as argued by Caliendo et al. (2015b), the Uruguay Round of multilateral trade negotiations was followed by significant reductions in protection levels in all sectors and countries. These trends were particularly stronger in developing countries with initially more protectionist trade policies. Second, the period was also marked by a radical change in the international political landscape that allowed the integration into the word economy of countries like Russia and China. To the extent that both of these phenomena were driven by a worldwide trend towards eliminating import tariffs, the liberalization process was unlikely to be affected by changes in productivity in particular regions of the United States.

We obtain tariff schedules for 53 countries from the TRAINS database in the World Bank's WITS website. For each country pair, we create ad-valorem import tariffs at the 2-digit HS level by taking the simple average of tariff lines within each HS code. We then construct the bilateral tariff between each US state and each foreign country. For each state, we construct the weighted average of HS-2 tariffs between the US and foreign countries using as weights trade flows of each US state and the foreign country in the initial period.

Labor Market Data. We extract employment and wage information from the CPS-MORG. We compute state-level outcomes using reported weekly earnings and weekly hours. We classify individuals into three aggregate sectors: (i) the home sector that includes non-employed individuals; (ii) the manufacturing sector that includes all manufacturing and agriculture industries; and (iii) the nonmanufacturing sector that includes all remaining industries.

We construct state-level price indices using data from the Cost of Living Index, published by the Council for Community and Economic Research (C2ER). This is a well-known source of living cost differentials among cities in the U.S – see Moretti (2013). Our state-level price index is the average of the price index changes of urban areas in the state weighted by the area's population.[24]

Trade Data. Our main source of international trade data is the Eora Multi Region Input Output – for details, see Lenzen et al. (2013). We construct bilateral trade flows for 59 countries in two sectors, manufacturing and nonmanufacturing. We merge this matrix of international trade flows with a matrix of bilateral flows between the 50 US states. To construct this *regional* matrix of trade, we use between-states shipments data from the Commodity Flow Survey, which is conducted every five years by the US Census Bureau. Finally, we use Census data at the port of entry/exit level to obtain trade flows between each

[24] As highlighted by Feenstra (1994), our price index does not fully capture the effect of new varieties in the cost of living. One possible solution is to use the local prices in the Nielsen Homescan Dataset to construct state-level price indices, as in Handbury and Weinstein (2014). However, this would significantly reduce our sample since bar-code price data is only available for a subset of states after 2004. A deeper investigation of these issues is a fruitful agenda of future research.

of the US states and each of the 59 countries in the MRIO. Merging these three matrices of trade flows leads to a matrix of bilateral trade flows between the 50 US states and 59 countries. Appendix C describes the details of the trade data construction.

5.3 Results

We now turn to the estimation of the structural parameters $\boldsymbol{\theta} \equiv (\phi_h, \phi_e, \phi_m, \psi_e, \psi_m)$ using the methodology derived in Section 4. For each period, we obtain the structural residual $e_i^t(\boldsymbol{\theta}) = (\nu_{sd}^t, \tilde{\nu}_{sd}^t, \nu_d^t, \zeta_d^t)$ by combining equations (39)–(42) with data on trade and labor outcomes for US states. Following Proposition 5, we implement the Two-Step GMM where, in each step, the moment conditions use the model's predicted values of $H_i^*\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0 | \boldsymbol{\theta}^p\right) = \nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta}^p)$ given the observed changes in import tariffs, $\hat{\boldsymbol{z}}^t$.

We report estimates obtained from the estimation of the model including period and sector dummies in the stacked sample of state-sector pairs for the periods of 1993-1997, 1997-2002, 2002-2007 and 2007-2012. In all specifications, we calibrate the elasticity of substitution across origins to $\chi_k = 5$ for both sectors, as in Simonovska and Waugh (2014a), and the elasticity of substitution across sectors to $\chi = 0.3$, as in Buera, Kaboski, and Rogerson (2015).[25]

Table 1 shows the estimates of the structural parameters along with the standard errors clustered at the state-level. In the estimation, we impose that $\psi_e = -\psi_m = \psi$, so that the agglomeration and congestion forces are combined into a single parameter. As shown in Appendix C.4 , we are unable to jointly estimate these parameters, but their separate estimation indicates that this is a good approximation for the forces in the data. Our point estimate of ψ is 0.35, which multiplied by the trade elasticity implies an agglomeration elasticity close to the local agglomeration forces implied by models with increasing returns to scale and firm entry as in Krugman (1980). These sectoral agglomeration forces are balanced by the regional congestion forces stemming from lower cost advantage as population becomes more concentrated in space. [26]

The other three columns report our estimates of the structural parameters governing

[25] Appendix C.4 shows that our results are robust to a number of specification choices. First, we obtain similar results by including a set of regional controls that absorb state-specific shocks potentially correlated with exposure to tariff shocks. Specifically, we include period dummies interacted with sector dummies and local outcomes in 1993 (manufacturing employment share, college graduates labor force share, and female labor force participation). Second, our results are robust to the sample choice. We obtain similar estimates if we restrict the sample to manufacturing and to the period prior to the great recession (i.e., exclude 2007-2012). Third, we obtain similar results when states are weighted by their share in national population in 1993.

[26] The estimates of agglomeration forces in the literature present large variation. Kline and Moretti (2014) estimate the local agglomeration elasticity, i.e. the elasticity of county productivity with respect to manufacturing density, to be 0.4. Peters (2017) also finds a large agglomeration elasticity.

regional labor supply. First, our point estimate for the intensive margin, i.e. the elasticity of hours worked with respect to the real wage, equals 0.35. This is in the range of the estimates provided by Chetty (2012). Second, the extensive margin elasticity is $\phi_e = 0.69$, implying that a 1% increase in the sector relative wage triggers a 0.69% increase in the sector relative employment share. In terms of employment shares, if real wages in sector k increase by 1%, the employment share in sector k increases by $\phi_e l_{ko}(1 - l_{ko})$. Using the national employment composition in 2007, this implies an extensive margin elasticity of 0.06% for the manufacturing sector and 0.17% for the non-manufacturing. Our estimates are on the lower bound of the estimates of the extensive margin elasticity in the literature reviewed by Chetty (2012). Finally, we obtain an imprecise estimate of the migration elasticity. This result is somewhat consistent with recent empirical literature finding very weak responses of migration flows to international trade shocks (see Autor et al. (2013), Kovak (2013) and Dix Carneiro and Kovak (2016)).

We now turn to an investigation of the model ability to generate predicted responses that are consistent with the actual change in trade and labor market outcomes across US states. To this end, we compute predicted changes in endogenous variables triggered by tariff shocks with the estimates in Table 1. Table 2 reports the relationship between the actual changes in endogenous variables and the predicted changes in the same endogenous variables implied by our general equilibrium model. For all variables, there is a positive and statistically significant relationship between actual and predicted changes across US states. This indicates that, following trade costs shocks, the general equilibrium responses implied by our model are consistent with the observed pattern of responses in labor market outcomes across US states. We take this as evidence that the adjustment margins in our model provide a good approximation for those at work in the US.

6 Application: Effect of trade shocks on US local labor markets

The goal of this section is to assess the empirical relevance of the general equilibrium effects arising from trade shocks highlighted in the theoretical model. We consider the following counterfactual exercise: starting from the initial equilibrium in 2007, we revert the manufacturing trade costs between all countries and US states to their level in 1993. This was a period of rising trade integration across countries that ended in the trade collapse triggered by the great recession.

We proceed in two steps. First, we use the parametric gravity model described in the

previous section to recover changes in trade costs from the data. Second, we investigate the impact of the shock on US local labor markets, by decomposing the total shock in a direct effect and an indirect effect, arising from spillover exposure from shocks to other states. Finally, we decompose further the effects to study which countries originated them.

6.1 Recover Changes in Bilateral Trade Costs

The first step of the analysis is to use the model to recover trade shocks from the data. Recovering the shocks through the lens of our general equilibrium model allows us to exploit not only tariff changes but all other sources of trade shocks driving international market integration in the period. Specifically, we exploit the spending shares in in equation (37) to write changes in trade costs as:

$$\hat{\tau}_{ko,sd} = \left(\frac{\hat{\bar{x}}_{k,sd}}{\hat{\bar{x}}_{k,ko}} \right)^{\frac{1}{1-\chi} - \frac{1}{1-\chi_k}} \left(\frac{\hat{x}_{ko,sd}}{\hat{x}_{ko,ko}} \right)^{\frac{1}{1-\chi_k}} \frac{\hat{P}_{sd}}{\hat{P}_{ko}}. \tag{43}$$

where $\bar{x}_{k,sd} \equiv \sum_o x_{ko,sd}$ is the total expenditure share on sector k by individuals in sector d of region d.

Note that equation (43) does not assume that trade costs are symmetric, and thus can be seen as a generalization of the approach proposed by Head and Ries (2001). The changes in trade shares $\hat{x}_{ko,sd}$ and expenditure shares $\hat{\pi}_{k,sd}$ are computed directly from the EORA dataset. Following Eaton, Kortum, Neiman, and Romalis (2016), we measure the aggregate price indices $\hat{P}_{sd}$ with the price index of expenditures available from the Penn World Tables. We keep all the other trade costs and shocks fixed, and use the estimated parameters reported in Table 1. Given our imprecise estimates of the migration elasticity, our baseline results are computed under the assumption of $\phi_m = 0$. Section 6.3 evaluates the robustness of our results to changes in the strength of migration responses across markets.

Figure 1 plots the magnitude of the implied trade costs shock faced by US states in our counterfactual exercise. The shock is quite heterogenous across states and is large: the average change in the cost of exporting from US states to other countries is -13.83%, while the average trade costs of importing from other countries increase by 42.48%. The implied trade shock with China, one of most relevant trading partners of US states, is even larger: -27.63% for US exports and 129% for US imports from China.

6.2 Impact on Local Labor Markets

Figures 2 and 3 illustrate the impact of the trade shock on labor market outcomes in both manufacturing and non-manufacturing sectors. The effects are very heterogeneous, owing to

the different exposure of the regions to foreign markets, directly or indirectly, and to their different degrees of specialization on manufacturing and non-manufacturing. Interestingly, in almost all states, workers in manufacturing gain both in terms of real wages and employment, as a result of the lower exposure to foreign competition. Instead, non-manufacturing workers generally lose and non-manufacturing employment decreases for most of the states. However, the increase of the real wages and employment in the manufacturing industry has an effect on the non-manufacturing sectors because of demand spillovers. As a result, few states experience an increase in non-manufacturing real wages and employment.

Direct and Indirect Effects. Our framework allows us to decompose the total effect of the shock in a direct and indirect component. In particular, we use the results in Proposition 3 on the response of employment in a Small Open Economy to compute the "direct effect" of the trade shock to each of the US states. In such a setup, the direct exposure to trade shocks only affects local outcomes, and does not spillover to other markets in general equilibrium. We then define the indirect effect as the simple difference between the total effect and the direct effect.

Figure 4 plots the total and the direct effects of the trade shock on US states for real wages and employment in manufacturing. We can see that the two are highly correlated - the correlation is around 0.9 for both variables - and for almost all states the direct exposure to the shock is sizeable relative to the total effect. Interestingly, the states that experience a reduction in manufacturing employment and real wages, are also the ones for which the direct effects are weaker. This suggests that, although the direct exposure is the main force affecting the responses, the indirect effects arising from cross-market spillovers may revert the direction of the effect. This intuition is confirmed by Figure 5, which suggests that the direct and indirect effects of the trade shock are strongly negatively correlated, with the indirect spillovers being much smaller in magnitude. In contrast, Figure 6 shows that in the non-manufacturing sector the total effect is higher than the direct, but still the correlation between them is very high.

The main message from these results is that the direct effect is a good predictor of the overall general equilibrium effects of a trade shock. The indirect effects are one order of magnitude smaller than the direct effect of local shocks, which suggests that the cross-market general equilibirum linkages do not seem to be a first-order driver of labor market outcomes with our baseline estimates. This takes on even greater importance in light of the generality of our parametric empirical model, which captures the main forces present in existing quantitative spatial models.[27]

[27]Our results are consistent with the findings in Caliendo et al. (2014). In their quantitative model, featuring free mobility of workers across regions and input-output linkages, they show that the effects to

Sources of Trade Shocks. Finally, the flexibility of our framework allows us to decompose the trade shocks across origin countries. We focus on the responses after a trade shock with China, the main US trading partner, and compare them with the responses after a shock to all other countries. Figure 7 plots the US state-level response of labor market outcomes after the China shock with the corresponding impact following a shock to all countries except China. The plot is striking: the responses of the endogenous variables, in particular real wages, are strongly correlated. This suggests that simple reduced form approaches that try to estimate the impact of the China shock on US labor markets may confound the China shock with shocks from other countries, since these are correlated with the China one. In contrast, our counterfactual exercise, by isolating each of these shocks and by taking into account for the general equilibrium interactions, is able to separate the different shocks.

6.3 Sensitivity Analysis

Sensitivity analysis. Figure 8 plots the changes in employment in manufacturing after the trade shock under alternative parametrizations. Specifically, we plot the responses assuming i) a positive migration elasticity of $\phi_m = 0.3$; ii) positive elasticities of migration and agglomeration across regions of $\phi_m = \psi_m = 0.3$; iii) positive elasticity of agglomeration across sectors within a region, $\psi_e = 0.3$. The graphs show that, if we allow for movement of workers and for agglomeration spillovers *across* US states, the indirect effects are much stronger than in the baseline counterfactual, but the direct effect is still highly correlated with the total effect. Therefore, if agglomeration forces across regions are strong, a diff-in-diff empirical analysis, that relies on regional variation to identify the effect of a trade shock on local labor markets, would get right the differential effects of the shock, but would miss the aggregate effect.

7 Conclusions

In this paper, we bring endogenous labor supply and agglomeration forces to the forefront of the analysis regarding the relationship between trade and labor markets. Our analysis stresses the need to directly estimate the relationship between trade and labor market outcomes at regional level, using it to discipline the structural parameters of the model. To this end, we propose a new empirical methodology that uses as an instrument the impact of

other states of a productivity shock in a US state are one order of magnitude smaller than the effects of the state receiving the shock.

exogenous trade shocks on changes of the endogenous variables predicted by our general equilibrium model. This yields the most efficient estimator of the structural elasticities, i.e. a Model-implied Optimal IV. We then apply our methodology to evaluate the aggregate impacts of trade shocks affecting regional labor markets in the U.S. Interesting avenues for future research emerge from our study. We hope that our novel methodology, easily implementable with a simple GMM procedure, can be used for the estimation of structural parameters in a wide class of general equilibrium models.

References

Acemoglu, D., Restrepo, P., et al., 2017. Robots and jobs: Evidence from us labor markets. Tech. rep., National Bureau of Economic Research, Inc.

Adão, R., 2015. Worker heterogeneity, wage inequality, and international trade: Theory and evidence from brazil. Unpublished paper, MIT.

Adao, R., Costinot, A., Donaldson, D., 2017. Nonparametric counterfactual predictions in neoclassical models of international trade. The American Economic Review 107 (3), 633--689.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., Wolf, N., 2015. The economics of density: Evidence from the berlin wall. Econometrica 83 (6), 2127--2189.

Allen, T., Arkolakis, C., 2013a. A note on the existence and uniqueness of trade equilibria. mimeo, Northwestern and Yale Universities.

Allen, T., Arkolakis, C., 2013b. Trade and the topography of the spatial economy. mimeo, Northwestern University.

Allen, T., Arkolakis, C., 2014. Trade and the topography of the spatial economy. The Quarterly Journal of Economics 129 (3), 1085--1140.

Allen, T., Arkolakis, C., Li, X., 2015. Optimal city structure. Yale University, mimeograph.

Allen, T., Arkolakis, C., Takahashi, Y., 2014. Universal gravity. Tech. rep., Mimeo Northwestern and Yale University.

Anderson, J. E., 1979. A theoretical foundation for the gravity equation. American Economic Review 69 (1), 106--116.

Arkolakis, C., Costinot, A., Donaldson, D., Rodríguez-Clare, A., 2017. The elusive pro-competitive effects of trade.

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

Armington, P. S., 1969. A theory of demand for products distinguished by place of production. International Monetary Fund Staff Papers 16, 159--178.

Autor, D., Dorn, D., 2013. The growth of low-skill service jobs and the polarization of the us labor market. The American Economic Review 103 (5), 1553--1597.

Autor, D., Dorn, D., Hanson, G. H., 2013. The china syndrome: Local labor market effects of import competition in the united states. The American Economic Review 103 (6), 2121--2168.

Autor, D. H., Katz, L. F., Kearney, M. S., 2008. Trends in us wage inequality: Revising the revisionists. The Review of economics and statistics 90 (2), 300--323.

Balistreri, E. J., Hillberry, R. H., Rutherford, T. F., 2010. Trade and welfare: Does industrial organization matter? Economics Letters 109 (2), 85--87.

Bartelme, D., 2015. Trade costs and economic geography: evidence from the us. Work. Pap., Univ. Calif., Berkeley.

Beraja, M., Hurst, E., Ospina, J., 2016. The aggregate implications of regional business cycles. Tech. rep., National Bureau of Economic Research.

Berry, S., 1994. Estimating discrete-choice models of product differentiation. RAND Journal of Economics 25 (2), 242--262.

Berry, S., Gandhi, A., Haile, P., 2013. Connected substitutes and invertibility of demand. Econometrica 81 (5), 2087--2111.

Berry, S., Levinsohn, J., Pakes, A., 1999. Voluntary export restraints on automobiles: Evaluating a trade policy. American Economic Review, 400--430.

Berry, S. T., Haile, P. A., 2014. Identification in differentiated products markets using market level data. Econometrica 82 (5), 1749--1797.

Bryan, G., Morten, M., 2015. Economic development and the spatial allocation of labor: Evidence from indonesia. Manuscript, London School of Economics and Stanford University.

Buera, F. J., Kaboski, J. P., Rogerson, R., 2015. Skill biased structural change. Tech. rep., National Bureau of Economic Research.

Burstein, A., Morales, E., Vogel, J., 2016. Changes in between-group inequality: computers, occupations, and international trade. Tech. rep., mimeo, September 13.

Bustos, P., Caprettini, B., Ponticelli, J., et al., 2016. Agricultural productivity and structural transformation: Evidence from brazil. American Economic Review 106 (6), 1320--1365.

Caliendo, L., Dvorkin, M., Parro, F., 2015a. The impact of trade on labor market dynamics. Tech. rep., National Bureau of Economic Research.

Caliendo, L., Feenstra, R. C., Romalis, J., Taylor, A. M., 2015b. Tariff reductions, entry, and welfare: Theory and evidence for the last two decades. Tech. rep., National Bureau of Economic Research.

Caliendo, L., Parro, F., 2014. Estimates of the trade and welfare effects of nafta. The Review of Economic Studies, rdu035.

Caliendo, L., Parro, F., Rossi-Hansberg, E., Sarte, P.-D., 2014. The impact of regional and sectoral productivity changes on the us economy. Tech. rep., National Bureau of Economic Research.

Chamberlain, G., 1987. Asymptotic efficiency in estimation with conditional moment restrictions. Journal of Econometrics 34 (3), 305--334.

Chaney, T., 2008. Distorted gravity: The intensive and extensive margins of international trade. American Economic Review 98 (4), 1707--1721.

Chetty, R., 2012. Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica 80 (3), 969--1018.

Chetty, R., Guren, A., Manoli, D., Weber, A., 2013. Does indivisible labor explain the difference between micro and macro elasticities? a meta-analysis of extensive margin elasticities. NBER Macroeconomics Annual 27 (1), 1--56.

Comin, D. A., Lashkari, D., Mestieri, M., 2015. Structural change with long-run income and price effects. Tech. rep., National Bureau of Economic Research.

Costinot, A., Donaldson, D., Komunjer, I., 2011. What goods do countries trade? a quantitative exploration of ricardo's ideas. The Review of Economic Studies, rdr033.

Costinot, A., Rodriguez-Clare, A., 2013. Trade theory with numbers: Quantifying the consequences of globalization. Tech. rep., National Bureau of Economic Research.

Cravino, J., Sotelo, S., 2017. Trade induced structural change and the skill premium. Tech. rep., National Bureau of Economic Research.

Dekle, R., Eaton, J., Kortum, S., 2007. Unbalanced trade. Tech. rep., National Bureau of Economic Research.

Dix Carneiro, R., Kovak, B. K., 2016. Trade liberalization and regional dynamics.

Eaton, J., Kortum, S., 2001. Technology, trade, and growth: A unified framework. European Economic Review 45 (4-6), 1741--1780.

Eaton, J., Kortum, S., 2002. Technology, geography and trade. Econometrica 70 (5), 1741--1779.

Eaton, J., Kortum, S., Neiman, B., Romalis, J., 2016. Trade and the global recession. The American Economic Review 106 (11), 3401--3438.

Ethier, W. J., 1982a. Decreasing costs in international trade and frank graham's argument for protection. Econometrica: Journal of the Econometric Society, 1243--1268.

Ethier, W. J., 1982b. National and international returns to scale in the modern theory of international trade. American Economic Review 72 (3), 389--405.

Faber, B., Gaubert, C., 2016. Tourism and economic development: Evidence from mexico's coastline. Tech. rep., National Bureau of Economic Research.

Feenstra, R., 1994. New product varieties and the measurement of international prices. American Economic Review 84 (1), 157--177.

Fujita, M., Krugman, P., Venables, A. J., 1999. The Spatial Economy: Cities, Regions, and International Trade. MIT Press, Boston, Massachussetts.

Galle, S., Rodriguez-Clare, A., Yi, M., 2015. Slicing the pie: Quantifying the aggregate and distributional effects of trade. Unpublished manuscript, Univ. Calif., Berkeley.

Hakobyan, S., McLaren, J., 2016. Looking for local labor market effects of nafta. Review of Economics and Statistics 98 (4), 728--741.

Handbury, J., Weinstein, D. E., 2014. Goods prices and availability in cities. The Review of Economic Studies 82 (1), 258--296.

Head, K., Mayer, T., 2013. Gravity equations: Workhorse, toolkit, and cookbook. Centre for Economic Policy Research.

Head, K., Ries, J., 2001. Increasing returns versus national product differentiation as an explanation for the pattern of u.s.-canada trade. American Economic Review 91 (4), 858--876.

Helpman, E., 1998. The size of regions. Topics in Public Economics. Theoretical and Applied Analysis, 33--54.

Herrendorf, B., Rogerson, R., Valentinyi, A., 2013. Two perspectives on preferences and structural transformation. American Economic Review 103 (7), 2752--89.

Kline, P., Moretti, E., 2014. Local economic development, agglomeration economies and the big push: 100 years of evidence from the tennessee valley authority. Quarterly Journal of Economics 129, 275--331.

Kortum, S., 1997. Research, patenting, and technological change. Econometrica 65 (6), 1389--1419.

Kovak, B. K., 2013. Regional effects of trade reform: What is the correct measure of liberalization? The American Economic Review 103 (5), 1960--1976.

Krugman, P., 1980. Scale economies, product differentiation, and the pattern of trade. American Economic Review 70 (5), 950--959.

Krugman, P., 1991. Increasing returns and economic geography. The Journal of Political Economy 99 (3), 483--499.

Krugman, P., Venables, A. J., 1995. Globalization and the inequality of nations. Quarterly Journal of Economics 110 (4), 857--880.

Kucheryavyy, K., Lyn, G., Rodríguez-Clare, A., 2016. Grounded by gravity: A well-behaved trade model with industry-level economies of scale. Tech. rep., National Bureau of Economic Research.

Lee, E., 2015. Trade, inequality, and the endogenous sorting of heterogeneous workers. Tech. rep., mimeo.

Lenzen, M., Moran, D., Kanemoto, K., Geschke, A., 2013. Building eora: a global multi-region input--output database at high country and sector resolution. Economic Systems Research 25 (1), 20--49.

Lucas, R. E., Rossi-Hansberg, E., 2003. On the internal structure of cities. Econometrica 70 (4), 1445--1476.

McFadden, D., 1980. Econometric models for probabilistic choice among products. Journal of Business, S13--S29.

Mian, A., Sufi, A., 2014. What explains the 2007--2009 drop in employment? Econometrica 82 (6), 2197--2223.

Milgrom, P., Segal, I., 2002. Envelope theorems for arbitrary choice sets. Econometrica 70 (2), 583--601.

Monte, F., Redding, S. J., Rossi-Hansberg, E., 2015. Commuting, migration and local employment elasticities. Tech. rep., National Bureau of Economic Research.

Moretti, E., 2013. Real wage inequality. American Economic Journal: Applied Economics 5 (1), 65--103.

Muendler, M., 2017. Trade, technology, and prosperity: An account of evidence from a labor-market perspective.

Nakamura, E., Steinsson, J., 2014. Fiscal stimulus in a monetary union: Evidence from us regions. The American Economic Review 104 (3), 753--792.

Newey, W. K., McFadden, D., 1994. Large sample estimation and hypothesis testing. Handbook of econometrics 4, 2111--2245.

Newey, W. K., Powell, J. L., 2003. Instrumental variable estimation of nonparametric models. Econometrica 71 (5), 1565--1578.

Peters, M., 2017. Refugees and local agglomeration: Evidence from germany's post-war population expulsions. Unpublished manuscript.

Pierce, J. R., Schott, P. K., 2016. The surprisingly swift decline of us manufacturing employment. The American Economic Review 106 (7), 1632--1662.

Redding, S. J., 2012. Goods trade, factor mobility and welfare. mimeo.

Redding, S. J., Rossi-Hansberg, E., 2016. Quantitative spatial economics. Tech. rep., National Bureau of Economic Research.

Reynaert, M., Verboven, F., 2014. Improving the performance of random coefficients demand models: the role of optimal instruments. Journal of Econometrics 179 (1), 83--98.

Simonovska, I., Waugh, M. E., 2014a. The elasticity of trade: Estimates and evidence. Journal of International Economics 92 (1), 34--50.

Simonovska, I., Waugh, M. E., 2014b. Trade models, trade elasticities, and the gains from trade. Tech. rep., National Bureau of Economic Research.

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.

8 Tables and Figures

Table 1: Estimation of structural parameters

	Labor Supply Elasticity			Agglomeration Elasticity
	Hours	Between sectors	Between regions	
	ϕ_h	ϕ_e	ϕ_m	$\psi_e = -\psi_m$
	0.358***	0.686**	0.458	0.354**
	(0.137)	(0.313)	(0.483)	(0.174)
F-stat	23.6	23.6	4.8	19.3

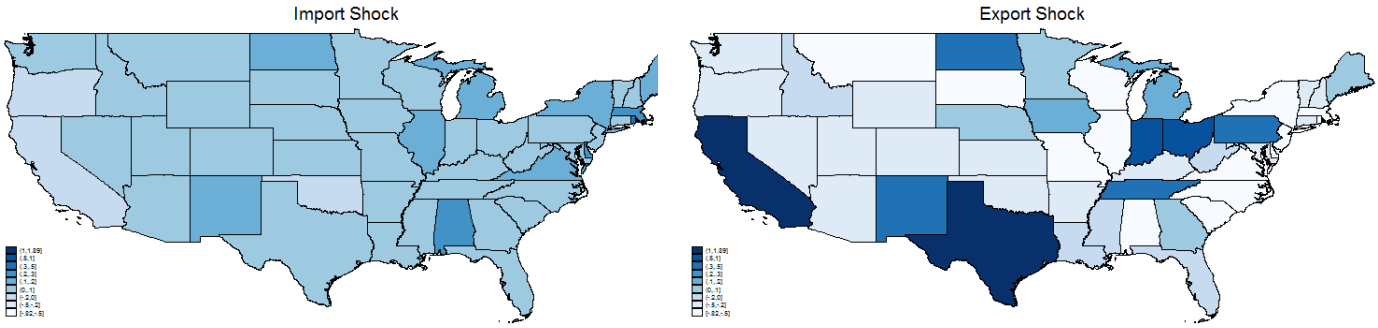
Note. Sample of state-sector pairs (2 sectors x 50 states x 4 time periods 1993-1997, 1997-2002, 2002-2007, 2007-2012). All regressions include period and sector dummies. Standard errors in parenthesis are clustered at the state-level. *** p < .10, ** p < .05, * p < .01

Table 2: Model Fit

	Real Wage	Hours	Sectoral relative employment	Non-Participation Share
	(1)	(2)	(3)	(4)
Predicted change	2.699***	2.787***	1.815**	3.702**
	(0.556)	(0.720)	(0.877)	(1.688)

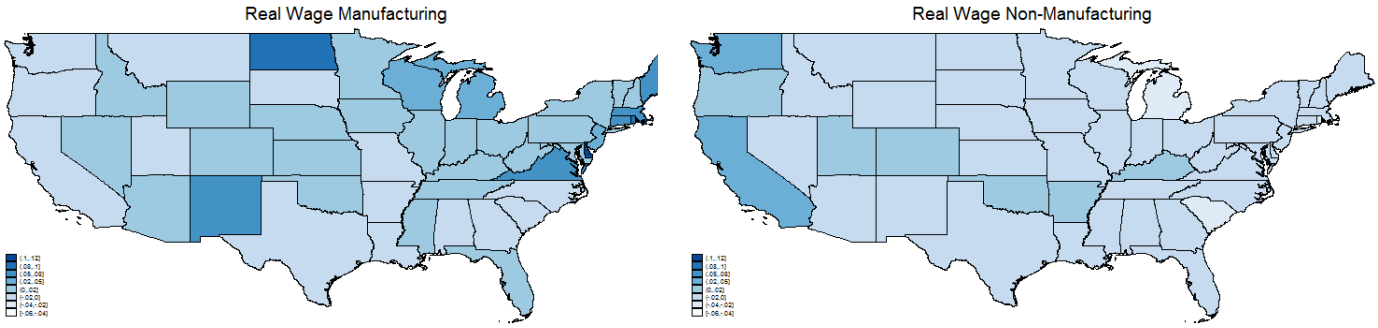
Note. Sample of state-sector pairs (2 sectors x 50 states x 4 time periods 1993-1997, 1997-2002, 2002-2007, 2007-2012). All regressions include period and sector dummies. Standard errors in parenthesis are clustered at the state-level. *** p < .10, ** p < .05, * p < .01

Figure 1: Trade shock



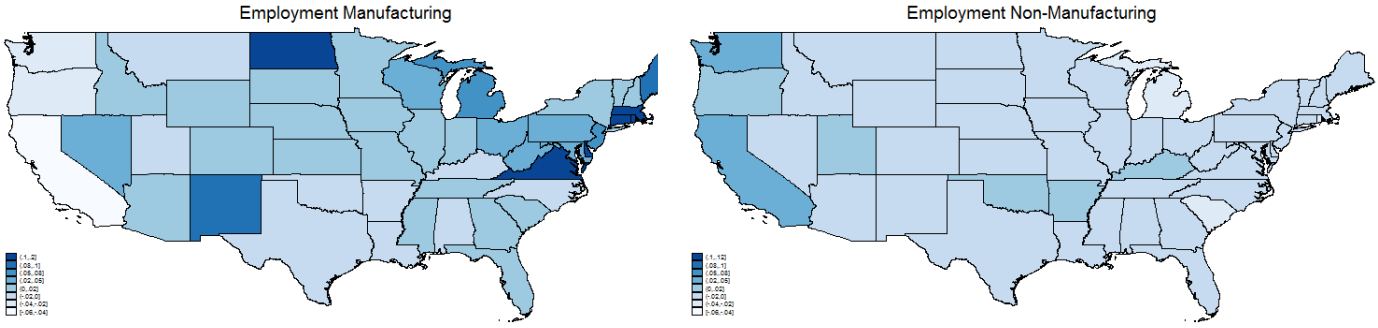
The map on the left shows, for each US state, the log-change from 2007 to 1993 in the average trade costs of importing from foreign countries, weighted by the import share. Values not reported for Alaska and Hawaii are 0.04 for both. The map on the right instead shows the log-change from 2007 to 1993 in the average trade costs of exporting to foreign countries, weighted by the export share. Values not reported for Alaska and Hawaii are -0.3 for both.

Figure 2: Impact of trade shock, real wage



The map on the left shows, for each US state, the log-change in the real wage in the manufacturing sector after the counterfactual trade shock shown in Figure 1. Values not reported for Alaska and Hawaii are 0.002 and 0.018, respectively. The map on the right instead shows the corresponding change in the non manufacturing sector. Values not reported for Alaska and Hawaii are -0.006 and 0.002, respectively.

Figure 3: Impact of trade shock, employment



The map on the left shows, for each US state, the log-change in total employment in the manufacturing sector after the counterfactual trade shock shown in Figure 1. Values not reported for Alaska and Hawaii are 0.007 and 0.028, respectively. The map on the right instead shows the corresponding change in the non manufacturing sector. Values not reported for Alaska and Hawaii are -0.006 and 0.001, respectively.

Figure 4: Direct vs total effect

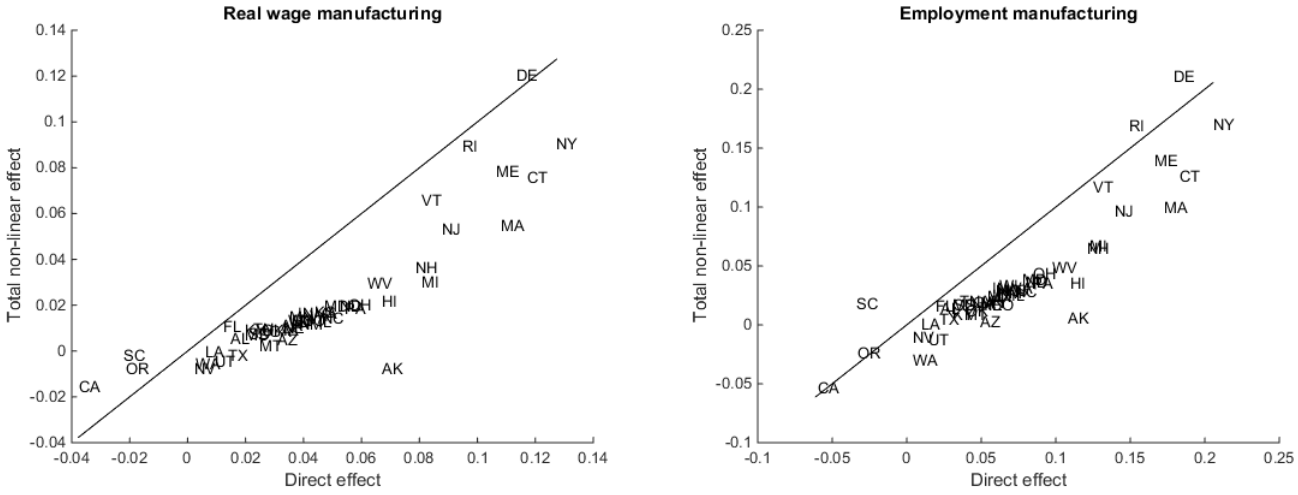


Figure 5: Direct vs indirect effect

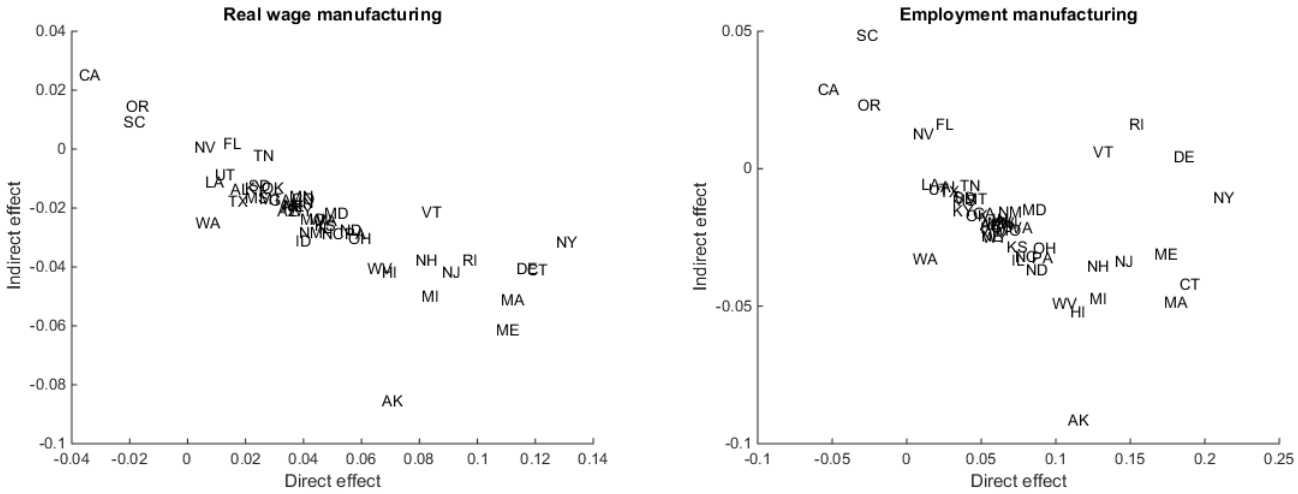


Figure 6: Direct vs total effect, non-manufacturing

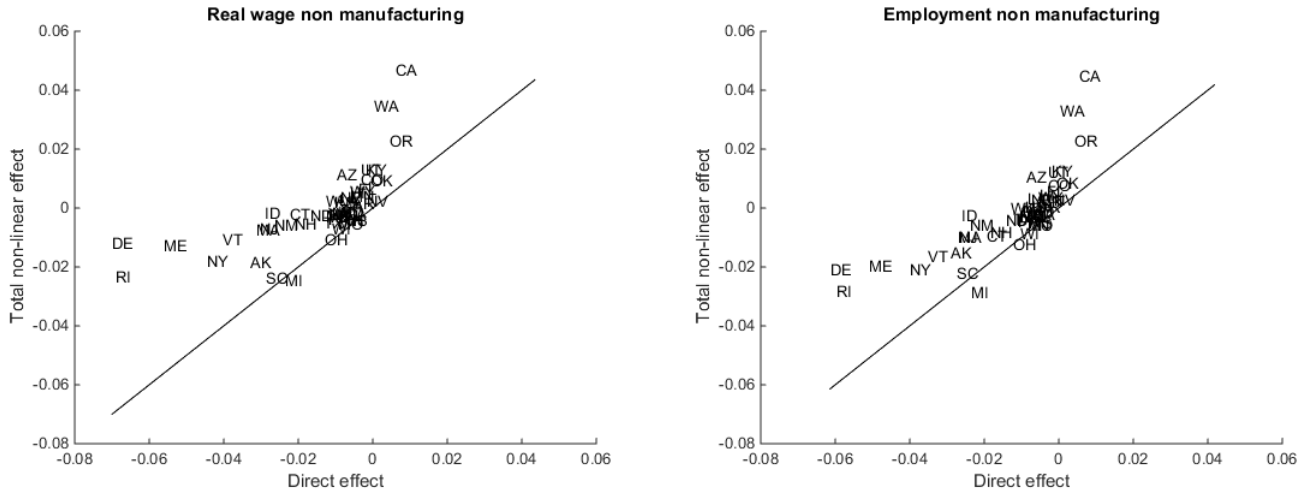


Figure 7: Correlation between China shock and other shocks

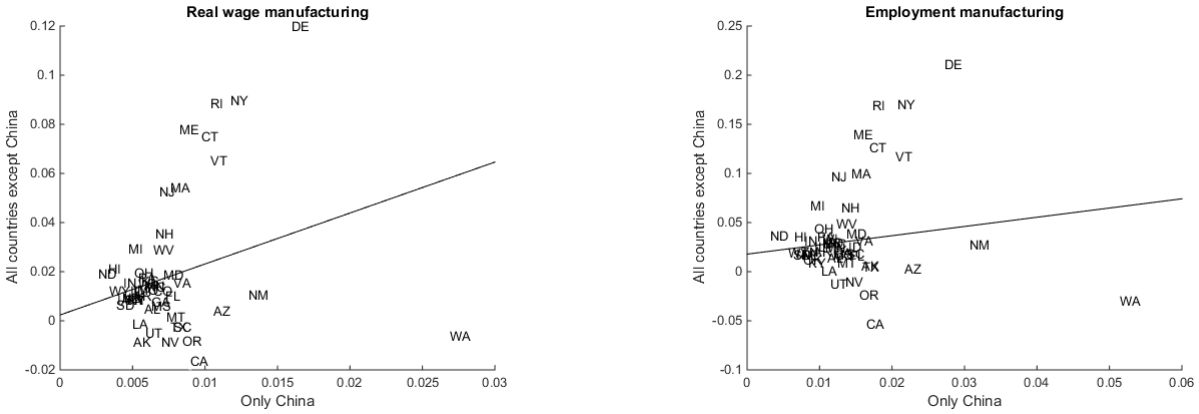
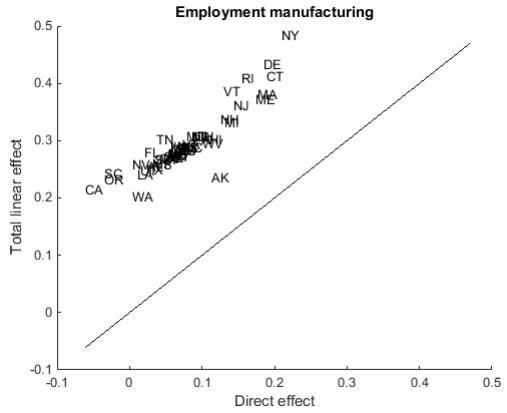
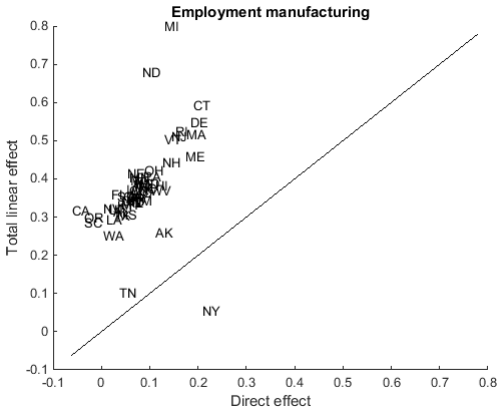


Figure 8: Sensitivity analysis

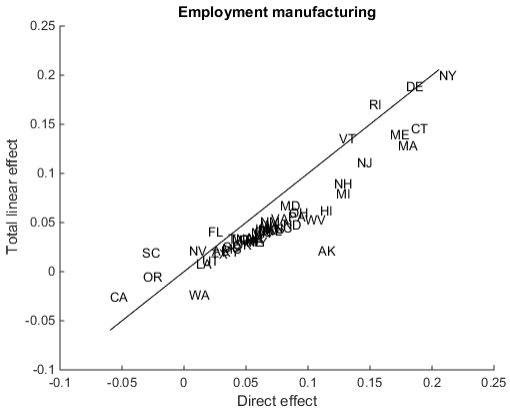
Panel A



Panel B



Panel C



Panel A: baseline parameters plus migration across regions. Panel B: baseline parameters plus migration and agglomeration across regions. Panel C: baseline parameters plus agglomeration across sectors.

A Proofs for Section 3

A.1 Proposition 1

We obtain (11)–(14)immediately by writing equations (6)–(10) in changes. We now show that, by specifying preferences and technology, we obtain an equilibrium with identical trade and labor outcomes as the initial equilibrium under the normalization in (15). To see this, consider an initial equilibrium such that

$$L_j^0 = \Phi_j\left(\left\{\nu_i^0\rho_i^0\frac{w_i^0}{P_i^0}\right\}_{i\in\mathcal{J}_c}\right)$$

$$p_{ii}^0 = \frac{w_i^0}{P_i^0}\frac{1}{\zeta_i^0\Psi_i^0\left(\left\{L_j^0\right\}_j\right)}$$

$$w_i^0 L_i^0 = \sum_j x_{ij}^0\rho_i^0 w_j^0 L_j^0$$

$$x_{ij}^0 = X_{ij}\left(\left\{\tau_{oj}^0 p_{oo}^0\right\}_o\right) \quad \text{and} \quad P_j^0 = P_j\left(\left\{\tau_{oj}^0 p_{oo}^0\right\}_o\right).$$

Let us construct an alternative economy without trade costs ($\tilde{\tau}_{ij}\equiv 1$), where technology is given by

$$\tilde{\Psi}_i\left(\left\{L_j\right\}_j\right) \equiv \frac{w_i^0}{\zeta_i^0}\frac{\Psi_i\left(\left\{L_j\right\}_j\right)}{\Psi_i\left(\left\{L_j^0\right\}_j\right)},$$

and preferences are given by

$$\tilde{U}_c\left(\left\{C_j\right\}_{j\in\mathcal{J}_c},\left\{L_j\right\}_{j\in\mathcal{J}_c}\right) \equiv U_c\left(\left\{\frac{\nu_j^0}{P_j^0}C_j\right\}_{j\in\mathcal{J}_c},\left\{L_j\right\}_{j\in\mathcal{J}_c}\right)$$

$$\tilde{V}_j\left(\left\{c_{ij}\right\}_i\right) \equiv V_j\left(\left\{c_{ij}\frac{P_j^0}{\tau_{ij}^0 p_{ii}^0}\right\}_i\right).$$

In this case, we immediately get that

$$\tilde{\Phi}_j\left(\left\{\rho_i\omega_i\right\}_{i\in\mathcal{J}_c}\right) = \Phi_j\left(\left\{\frac{\nu_i^0}{P_i^0}\rho_i\omega_i\right\}_{i\in\mathcal{J}_c}\right)$$

$$\tilde{X_{ij}}\left(\left\{p_{ij}\right\}_i\right) = X_{ij}\left(\left\{\frac{\tau_{ij}^0 p_{ii}^0}{P_j^0}p_{ij}\right\}_i\right) \quad \text{and} \quad \tilde{P}_j\left(\left\{p_{ij}\right\}_i\right) = P_j\left(\left\{\frac{\tau_{ij}^0 p_{ii}^0}{P_j^0}p_{ij}\right\}_i\right).$$

In this economy, the equilibrium entails $\tilde{w}_i = w_i^0$, $\tilde{L}_i = L_i^0$, $\tilde{p}_{ii} = \tilde{P}_i = 1$, and $\tilde{x}_{ij} = x_{ij}^0$.

To see this, notice that

$$\tilde{p}_{ii} = \frac{\tilde{w}_i}{\zeta_i^0 \tilde{\Psi}_i\left(\left\{L_j^0\right\}_j\right)} = 1$$

$$\tilde{\Phi}_j\left(\left\{\rho_i^0 \tilde{w}_i\right\}_{i\in\mathcal{J}_c}\right) = \Phi_j\left(\left\{\frac{\nu_i^0}{P_i^0}\rho_i^0 w_i^0\right\}_{i\in\mathcal{J}_c}\right) = L_j^0$$

$$w_i^0 L_i^0 = \sum_j x_{ij}^0 \rho_i^0 w_j^0 L_j^0$$

$$\tilde{X_{ij}}\left(\{\tilde{\tau}_{ij}\tilde{p}_{ii}\}_i\right) = X_{ij}\left(\left\{\frac{\tau_{ij}^0 p_{ii}^0}{P_j^0}\right\}_i\right) = x_{ij}^0 \quad \text{and} \quad \tilde{P}_j\left(\{\tilde{\tau}_{ij}\tilde{p}_{ii}\}_i\right) = P_j\left(\left\{\frac{\tau_{ij}^0 p_{ii}^0}{P_j^0}\right\}_i\right) = 1 \quad \blacksquare$$

A.2 Proposition 2

Part 1: Equations (20)–(21). We totally differentiate equations 7–9:

$$\log \hat{L}_i = \sum_j \phi_{ij}\left(\log \hat{\omega}_j + \log \hat{\rho}_j + \log \hat{\nu}_j\right).$$

$$\log \hat{\omega}_i = \log \hat{Q}_i + \log \hat{\zeta}_i + \sum_j \psi_{ij} \log \hat{L}_j$$

In matrix notation, these expressions can be written as

$$\log \hat{\boldsymbol{L}} = \bar{\boldsymbol{\phi}}\left(\log \hat{\boldsymbol{\omega}} + \log \hat{\boldsymbol{\rho}} + \log \hat{\boldsymbol{\nu}}\right)$$

$$\log \hat{\boldsymbol{\omega}} = \log \hat{\boldsymbol{Q}} + \log \hat{\boldsymbol{\zeta}} + \bar{\boldsymbol{\psi}} \log \hat{\boldsymbol{L}}.$$

Plugging the second expression into the first expression,

$$\left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{\phi}}\bar{\boldsymbol{\psi}}\right) \log \hat{\boldsymbol{L}} = \bar{\boldsymbol{\phi}}\left(\log \hat{\boldsymbol{Q}} + \log \hat{\boldsymbol{\rho}} + \log \hat{\boldsymbol{\zeta}} + \log \hat{\boldsymbol{\nu}}\right).$$

Under Assumption 2, the matrix $\bar{\boldsymbol{\Lambda}} \equiv \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{\phi}}\bar{\boldsymbol{\psi}}\right)^{-1} \bar{\boldsymbol{\phi}}$ is well defined.. Thus,

$$\log \hat{\boldsymbol{L}} = \bar{\boldsymbol{\Lambda}}\left(\log \hat{\boldsymbol{Q}} + \log \hat{\boldsymbol{\rho}} + \log \hat{\boldsymbol{\zeta}} + \log \hat{\boldsymbol{\nu}}\right) \tag{44}$$

Substituting it into the expression for real wages,

$$\log \hat{\boldsymbol{\omega}} = \left(\bar{\boldsymbol{I}} + \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\Lambda}}\right)\left(\log \hat{\boldsymbol{Q}} + \log \hat{\boldsymbol{\zeta}}\right) + \bar{\boldsymbol{\psi}}\bar{\boldsymbol{\Lambda}}\left(\log \hat{\boldsymbol{\rho}} + \log \hat{\boldsymbol{\nu}}\right). \tag{45}$$

Equations (44)–(45) immediately yield equations (20)–(21) by setting $\log \hat{\boldsymbol{\rho}} = \log \hat{\boldsymbol{\zeta}} = \log \hat{\boldsymbol{\nu}} = 0$.

Part 2: Equation (22). We totally differentiating the definition of $Q_i = p_{ii}/P_i$. By Shepard's lemma,

$$\log \hat{Q}_j = \log \hat{p}_{jj} - \sum_i x_{ij}^0 \left(\log \hat{\tau}_{ij} + \log \hat{p}_{ii}\right).$$

In matrix notation, the expression above becomes

$$\log \hat{\boldsymbol{Q}} = -\log \hat{\boldsymbol{\tau}}^I + \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \log \hat{\boldsymbol{p}}. \tag{46}$$

By substituting (9) into (10), we get

$$p_{ii}\zeta_i \Psi_i\left(\boldsymbol{L}\right) L_i = \sum_j x_{ij}\rho_j p_{jj}\zeta_j \Psi_j\left(\boldsymbol{L}\right) L_j,$$

with $x_{ij} = X_{ij}\left(\{\tau_{oj}p_{oo}\}_o\right)$.

The total differentiation of this expression implies that

$$\sum_j \left(\mathbb{I}_{[j=i]} - y_i^0 - \sum_d y_{id}^0 \chi_{jid} \right) \log \hat{p}_{jj} = \log \hat{\tau}_i^E + \sum_j y_{ij}^0 \log \hat{\rho}_j - \sum_j \left(\mathbb{I}_{[j=i]} - y_{ij}^0 \right) \log \hat{e}_j, \tag{47}$$

where $\mathbb{I}_{[o=i]}$ is an indicator function that equals 1 whenever $o = i$, and

$$\log e_j = \log \zeta_j + \log \hat{L}_j + \sum_o \psi_{jo} \log \hat{L}_o.$$

In matrix form,

$$\left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} - \bar{\boldsymbol{\chi}}\right) \log \hat{\boldsymbol{p}} = \log \hat{\boldsymbol{\tau}}^E - \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}}\right) \log \hat{\boldsymbol{e}} + \bar{\boldsymbol{y}} \log \hat{\boldsymbol{\rho}}$$

$$\log \hat{\boldsymbol{e}} = \log \hat{\boldsymbol{\zeta}} + \left(\bar{\boldsymbol{I}} + \bar{\boldsymbol{\psi}}\right) \log \hat{\boldsymbol{L}}.$$

Define the matrix $\bar{\bar{\boldsymbol{\Lambda}}} \equiv \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}}\right)\left(I + \bar{\boldsymbol{\psi}}\right)\bar{\boldsymbol{\Lambda}}$. Substituting (44) and (46),

$$\left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} - \bar{\boldsymbol{\chi}} + \bar{\bar{\boldsymbol{\Lambda}}}(I - \bar{\boldsymbol{x}})\right) \log \hat{\boldsymbol{p}} = \log \hat{\boldsymbol{\tau}}^E + \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\tau}}^I - \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} + \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\zeta}} + \left(\bar{\boldsymbol{y}} - \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\rho}} - \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\nu}}.$$

Notice that the matrix $\left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} - \bar{\boldsymbol{\chi}} + \bar{\bar{\boldsymbol{\Lambda}}}(I - \bar{\boldsymbol{x}})\right)$ is singular because, by definition, the transfer rule guarantees trade balance at the world level. So, in the system above, we use the matrix $\bar{\boldsymbol{M}}$ to set $\hat{p}_{mm} = 0$ and eliminate the m-th equation. Under Assumption 2,

$$\log \hat{\boldsymbol{p}} = \bar{\boldsymbol{\Gamma}} \left(\log \hat{\boldsymbol{\tau}}^E + \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\tau}}^I - \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} + \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\zeta}} + \left(\bar{\boldsymbol{y}} - \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\rho}} - \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\nu}} \right),$$

where we remind the reader the definition of $\bar{\boldsymbol{\Gamma}}$ from equation (19)

$$\bar{\boldsymbol{\Gamma}} \equiv \bar{\boldsymbol{M}}' \left[\bar{\boldsymbol{M}} \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} - \bar{\boldsymbol{\chi}} + \left(I - \bar{\boldsymbol{y}}\right)\left(I + \bar{\boldsymbol{\psi}}\right) \bar{\boldsymbol{\Lambda}}(I - \bar{\boldsymbol{x}}) \right) \bar{\boldsymbol{M}}' \right]^{-1} \bar{\boldsymbol{M}}.$$

Substituting this expression into (46),

$$\log \hat{\boldsymbol{Q}} = -\log \hat{\boldsymbol{\tau}}^I + \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \left(\log \hat{\boldsymbol{\tau}}^E + \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\tau}}^I - \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}} + \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\zeta}} + \left(\bar{\boldsymbol{y}} - \bar{\bar{\boldsymbol{\Lambda}}}\right) \log \hat{\boldsymbol{\rho}} - \bar{\bar{\boldsymbol{\Lambda}}} \log \hat{\boldsymbol{\nu}} \right). \tag{48}$$

By setting $\log \hat{\boldsymbol{\rho}} = \log \hat{\boldsymbol{\zeta}} = \log \hat{\boldsymbol{\nu}} = 0$, we immediately obtain (22):

$$\log \hat{\boldsymbol{Q}} = \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \log \hat{\boldsymbol{\tau}}^E - \left(\bar{\boldsymbol{I}} - \left(I - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \left(I - \bar{\boldsymbol{y}}\right)\left(I + \bar{\boldsymbol{\psi}}\right) \bar{\boldsymbol{\Lambda}}\right) \log \hat{\boldsymbol{\tau}}^I. \quad \blacksquare$$

A.3 Proposition 3

A.3.1 Part 1) Small Open Economy

We say that a market is a small open economy if its outcomes do not impact other markets in the world economy. Formally, we define a small open economy as follows.

Definition 3. Market d is a Small Open Economy if, for all other markets $i \neq d$,
1) The share of revenue and spending associated with d is approximately zero: $y_{id}^0 \approx 0$ and $x_{di}^0 \approx 0$ for all $i \neq d$.
2) Spending shares, employment, and productivity do not respond to outcomes in market d: $\chi_{dij} = \phi_{id} = \psi_{id} = 0$ for all $i, j \neq d$.

We use this definition to derive the impact of trade cost shocks for a small open economy. The equilibrium satisfies the following system of equations:

$$\log \hat{L}_i = \sum_j \phi_{ij} \left(-\log \hat{\tau}_j^I + \log \hat{p}_{jj} - \sum_o x_{oj}^0 \log \hat{p}_{oo} + \sum_l \psi_{jl} \log \hat{L}_l \right),$$

$$\sum_o \left(\mathbb{I}_{[o=i]} - y_{io}^0 - \bar{\chi}_{io} \right) \log \hat{p}_{oo} = \log \hat{\tau}_i^E - \sum_j \left(\mathbb{I}_{[j=i]} - y_{ij}^0 \right) \left(\log \hat{L}_j + \sum_o \psi_{jo} \log \hat{L}_o \right).$$

Consider a shock only to the small open economy d: $\log \hat{\tau}_{ij} = 0$ if $i \neq d$ and $j \neq d$. By the definition of a Small Open Economy, we have that $x_{di}^0 = 0$ and $\chi_{dij} = 0$ all $i, j \neq d$, which implies that

$$\log \hat{\tau}_i^I \equiv x_{di}^0 \log \hat{\tau}_{di} = 0, \text{ and } \log \hat{\tau}_i^E \equiv \sum_j y_{ij}^0 \left(\chi_{dij} \log \hat{\tau}_{dj} \right) = 0.$$

Applying $\log \hat{\tau}_i^I = \log \hat{\tau}_i^E = 0$ for all $i \neq d$ to the system above, we obtain that

$$\log \hat{L}_j = \sum_{j \neq d} \phi_{ji} \left(\log \hat{p}_{jj} - \sum_{i \neq d} x_{ij}^0 \log \hat{p}_{ii} + \sum_{l \neq d} \psi_{il} \log \hat{L}_l \right)$$

$$\sum_{o \neq d} \left(\mathbb{I}_{[o=i]} - y_{io}^0 - \sum_{j \neq d} y_{ij}^0 \chi_{oij} \right) \log \hat{p}_{oo} = - \sum_{j \neq d} \left(\mathbb{I}_{[j=i]} - y_{ij}^0 \right) \left(\log \hat{L}_j + \sum_{o \neq d} \psi_{jo} \log \hat{L}_o \right)$$

where the first equation uses $x_{id}^0 = \phi_{id} = \psi_{id} = 0$ and the second equation uses $\psi_{id} = y_{id} = \chi_{dij} = 0$.

Thus, we have $\log \hat{p}_{ii} = \log \hat{L}_i = 0$ for all $i \neq d$, and we only have to solve the following system for market d:

$$\log \hat{L}_d = \phi_{dd} \left(-\log \hat{\tau}_d^I + \left(1 - x_{dd}^0\right) \log \hat{p}_{dd} + \psi_{dd} \log \hat{L}_d \right)$$

$$\left(1 - y_{dd}^0 - \bar{\chi}_{dd}\right) \log \hat{p}_{dd} = \log \hat{\tau}_d^E - \left(1 - y_{dd}^0\right) \left(1 + \psi_{dd}\right) \log \hat{L}_d.$$

Define $\lambda_{dd} \equiv \left(1 - \phi_{dd}\psi_{dd}\right)^{-1} \phi_{dd}$. Rearranging the first expression in this system, we get

$$\log \hat{L}_d = \lambda_{dd} \left(-\log \hat{\tau}_d^I + \left(1 - x_{dd}^0\right) \log \hat{p}_{dd} \right), \tag{49}$$

which immediately yields expression (**??**).

Substituting (**??**) into the second equation in the system above, we get that

$$\left(1-y_{dd}^{0}-\bar{\chi}_{dd}\right)\log\hat{p}_{dd}=\log\hat{\tau}_{d}^{E}-\left(1-y_{dd}^{0}\right)\left(1+\psi_{dd}\right)\lambda_{dd}\left(-\log\hat{\tau}_{d}^{I}+\left(1-x_{dd}^{0}\right)\log\hat{p}_{dd}\right)$$

Define $\gamma_{dd}\equiv\left(1-y_{dd}^{0}-\bar{\chi}_{dd}+\lambda_{dd}\left(1+\psi_{dd}\right)\left(1-y_{dd}^{0}\right)\left(1-x_{dd}^{0}\right)\right)^{-1}$. Thus

$$\log\hat{p}_{dd}=\gamma_{dd}\left(\log\hat{\tau}_{d}^{E}+\lambda_{dd}\left(1-y_{dd}^{0}\right)\left(1+\psi_{dd}\right)\log\hat{\tau}_{d}^{I}\right).$$

Combining this expression together with (49) completes the proof of the first part of Proposition (3).■

A.3.2 Part 2) General Model

Inserting expression (44) into (48) we obtain

$$\log\hat{\boldsymbol{L}}=\bar{\boldsymbol{\Lambda}}\left(-\log\hat{\boldsymbol{\tau}}^{I}+\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{x}}\right)\bar{\boldsymbol{\Gamma}}\left[\log\hat{\boldsymbol{\tau}}^{E}+\bar{\bar{\boldsymbol{\Lambda}}}\log\hat{\boldsymbol{\tau}}^{I}-\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}+\bar{\bar{\boldsymbol{\Lambda}}}\right)\log\hat{\boldsymbol{\zeta}}+\left(\bar{\boldsymbol{y}}-\bar{\bar{\boldsymbol{\Lambda}}}\right)\log\hat{\boldsymbol{\rho}}-\bar{\bar{\boldsymbol{\Lambda}}}\log\hat{\boldsymbol{\nu}}\right]+\log$$

Using the definitions of $\bar{\boldsymbol{\Gamma}}^{I},\bar{\boldsymbol{\Gamma}}^{E}$, we obtain

$$\log\hat{\boldsymbol{L}}=\bar{\boldsymbol{\Lambda}}\left(\bar{\boldsymbol{\Gamma}}^{E}\log\hat{\boldsymbol{\tau}}^{E}-\bar{\boldsymbol{\Gamma}}^{I}\log\hat{\boldsymbol{\tau}}^{I}+\left[\bar{\boldsymbol{I}}+\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{x}}\right)\bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{y}}-\bar{\bar{\boldsymbol{\Lambda}}}\right)\right]\log\hat{\boldsymbol{\rho}}+\left[\bar{\boldsymbol{I}}-\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{x}}\right)\bar{\boldsymbol{\Gamma}}\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}+\bar{\bar{\boldsymbol{\Lambda}}}\right)\right]\log\hat{\boldsymbol{\zeta}}-$$

In addition, setting the non-trade related shocks to zero we have

$$\log\hat{\boldsymbol{L}}=\bar{\boldsymbol{\Lambda}}\left(\bar{\boldsymbol{\Gamma}}^{E}\log\hat{\boldsymbol{\tau}}^{E}+\bar{\boldsymbol{\Gamma}}^{I}\log\hat{\boldsymbol{\tau}}^{I}\right),$$

which in variables notation is equivalen to expression (24).■

A.4 Corollary 1

We start with the Local Labor Market Module. The total differentiation of equations (7) and (9) under Assumption 2.2-2.3 implies that

$$\begin{aligned}\log\hat{L}_{i} &= \phi\left(\log\hat{Q}_{i}+\log\hat{\zeta}_{i}+\psi\log\hat{L}_{i}-\tilde{\psi}\textstyle\sum_{j}\Psi_{j}\log\hat{L}_{j}\right)\\ &- \tilde{\phi}\textstyle\sum_{j}L_{j}\left(\log\hat{Q}_{j}+\log\hat{\zeta}_{j}+\psi\log\hat{L}_{j}-\tilde{\psi}\sum_{o}\Psi_{o}\log\hat{L}_{o}\right)\end{aligned}$$

Defining $\alpha_{j}^{1}\equiv\phi\tilde{\psi}\Psi_{j}+\tilde{\phi}\psi L_{j}-\tilde{\phi}\tilde{\psi}\Psi_{j}\sum_{o}L_{o}$, this expression implies

$$\left(1-\psi\phi\right)\log\hat{L}_{i}+\sum_{j}\alpha_{j}^{1}\log\hat{L}_{j}=\phi\left(\log\hat{Q}_{i}+\log\hat{\zeta}_{i}\right)-\tilde{\phi}\sum_{j}L_{j}\left(\log\hat{Q}_{j}+\log\hat{\zeta}_{j}\right)$$

Again, there are different micro foundations leading to this functional form. In Krugman (1980), firm entry and increasing returns to scale in production give rise to agglomeration forces that depend on the size of the employment in each market – in this case, $\psi=1/\left(\sigma-1\right)$ and $\tilde{\psi}=0$ for $i\neq j$, where σ is the elasticity of substitution across varieties. In the same spirit, the specification in Allen and Arkolakis (2014) corresponds to the case of $\psi=\alpha$ and

$\tilde{\psi} = 0$. Marshallian external economies of scale, as introduced in Ethier (1982b), correspond to similar specifications, with i denoting a sector in a region. This specification also captures congestion forces arising from other factors in production, such as housing or capital. In this case, production costs are increasing in other market's employment, and thus $\tilde{\psi} > 0$.

Furthermore, we show that the combination of the mappings $X_{ij}(\cdot)$ and $\Psi_i(.)$ implies that our framework is observationally equivalent to a general one-factor neoclassical economy with a flexible form of market-level external economies of scale.[28] Let us multiply both sides by $\tilde{\lambda}_i$ and sum across all i:

$$\left(1 - \psi\phi + \sum_o \alpha_o^1\right) \sum_i \alpha_i^1 \log \hat{L}_i = \sum_i \left(\phi\alpha_i^1 - \tilde{\phi} L_i \sum_o \alpha_o^1\right) \left(\log \hat{Q}_i + \log \hat{\zeta}_i\right).$$

Substituting this expression back into the expression above,

$$(1 - \psi\phi) \log \hat{L}_i = \phi \left(\log \hat{Q}_i + \log \hat{\zeta}_i\right) - \sum_j \left(\tilde{\phi} L_j + \frac{\left(\phi\alpha_j^1 - \tilde{\phi} L_j \sum_o \alpha_o^1\right)}{(1 - \psi\phi + \sum_o \alpha_o^1)}\right) \left(\log \hat{Q}_j + \log \hat{\zeta}_j\right)$$

$$(1 - \psi\phi) \log \hat{L}_i = \phi \left(\log \hat{Q}_i + \log \hat{\zeta}_i\right) - \phi \sum_j \left(\frac{1}{\phi} \frac{(1 - \psi\phi)\tilde{\phi} L_j + \phi\alpha_j^1}{1 - \psi\phi + \sum_o \alpha_o^1}\right) \left(\log \hat{Q}_j + \log \hat{\zeta}_j\right)$$

Thus,

$$\log \hat{L}_i = \alpha^2 \left(\log \hat{Q}_i + \log \hat{\zeta}_i\right) - \alpha^2 \sum_j \alpha_j^3 \left(\log \hat{Q}_j + \log \hat{\zeta}_j\right) \tag{50}$$

where

$$\alpha^2 \equiv \frac{\phi}{1 - \psi\phi}$$

$$\alpha_j^3 \equiv \frac{1}{\phi} \frac{(1 - \psi\phi)\tilde{\phi} L_j + \phi\alpha_j^1}{1 - \psi\phi + \sum_o \alpha_o^1} = \frac{1}{\phi} \frac{\tilde{\phi} L_j + \phi\tilde{\psi}\Psi_j \left(\phi - \tilde{\phi} \sum_o L_o\right)}{1 - \left(\psi - \tilde{\psi} \sum_o \Psi_o\right) \left(\phi - \tilde{\phi} \sum_o L_o\right)}.$$

Next we continue with the Trade Module. Consider $\log \hat{e}_i = \log \hat{\zeta}_i + \log \hat{L}_i + \bar{\boldsymbol{\psi}} \log \hat{\boldsymbol{L}}$, which under Assumption 2.2 is

$$\log \hat{e}_i = \log \hat{\zeta}_i + (1 + \psi) \log \hat{L}_i - \tilde{\psi} \sum_j \Psi_j \log \hat{L}_j$$

[28] Technology levels that are proportional to population are also postulated in Kortum (1997); Eaton and Kortum (2001). Models with spatial diffusion of knowledge, as in Fujita, Krugman, and Venables (1999) and Lucas and Rossi-Hansberg (2003), specify cross-location spillovers of knowledge so that $\psi_{ij} > 0$ even if $i \neq j$. Our environment can also easily accommodate models with multiple sectors that differ in terms of market structure and strength of economies of scale – e.g., Krugman and Venables (1995), Balistreri, Hillberry, and Rutherford (2010), Kucheryavyy, Lyn, and Rodríguez-Clare (2016).

Substituting expression (50),

$$\begin{aligned}\log \hat{e}_i &= \log \hat{\zeta}_i + (1+\psi)\alpha^2 \left(\log \hat{\zeta}_i + \log \hat{Q}_i - \textstyle\sum_j \alpha_j^3 \left(\log \hat{Q}_j + \log \zeta_j\right)\right) \\ &- \quad \tilde{\psi}\alpha^2 \textstyle\sum_j \Psi_j \left(\log \hat{\zeta}_j + \log \hat{Q}_j - \sum_o \alpha_o^3 \left(\log \hat{Q}_o + \log \zeta_o\right)\right)\end{aligned}$$

Thus,

$$\log \hat{e}_i = (1 + (1+\psi)\alpha^2) \log \hat{\zeta}_i + (1+\psi)\alpha^2 \log \hat{Q}_i - \tilde{\xi} \tag{51}$$

where

$$\tilde{\xi} \equiv \alpha^2 \sum_j \left((1+\psi)\lambda_j + \tilde{\psi}\Psi_j - \tilde{\psi}\lambda_j \sum_o \Psi_o\right) \left(\log \hat{Q}_j + \log \zeta_j\right).$$

To solve for the change in relative competitiveness $\hat{Q}_i$, we totally differentiate the market clearing condition under Assumption 2.1:

$$\log \hat{p}_{ii} + \log \hat{e}_i = \log \hat{\tau}_i^E - \epsilon \sum_j y_{ij} \left(\log \hat{p}_{ii} - \sum_o x_{oj} \log \hat{p}_{oo}\right) + \sum_j y_{ij} \left(\log \hat{p}_{jj} + \log \hat{e}_{jj}\right)$$

where

$$\log \hat{\tau}_i^E = -\epsilon \sum_j y_{ij} \left(\log \hat{\tau}_{ij} - \sum_o x_{oj} \log \hat{\tau}_{oj}\right).$$

Substituting $\log \hat{e}_i$ with expression (51),

$$\begin{aligned}(1+\epsilon) \log \hat{p}_{ii} - \textstyle\sum_j \left(y_{ij} + \epsilon \sum_l y_{il} x_{jl}\right) \log \hat{p}_{jj} &= \log \hat{\tau}_i^E - (1 + (1+\psi)\alpha^2) \left(\log \hat{\zeta}_i - \textstyle\sum_j y_{ij} \log \hat{\zeta}_j\right) \\ &- \quad (1+\psi)\alpha^2 \left(\log \hat{Q}_i - \textstyle\sum_j y_{ij} \log \hat{Q}_j\right)\end{aligned}$$

Again, there are different micro foundations leading to this functional form. In Krugman (1980), firm entry and increasing returns to scale in production give rise to agglomeration forces that depend on the size of the employment in each market – in this case, $\psi = 1/(\sigma - 1)$ and $\tilde{\psi} = 0$ for $i \neq j$, where σ is the elasticity of substitution across varieties. In the same spirit, the specification in Allen and Arkolakis (2014) corresponds to the case of $\psi = \alpha$ and $\tilde{\psi} = 0$. Marshallian external economies of scale, as introduced in Ethier (1982b), correspond to similar specifications, with i denoting a sector in a region. This specification also captures congestion forces arising from other factors in production, such as housing or capital. In this case, production costs are increasing in other market's employment, and thus $\tilde{\psi} > 0$.

Furthermore, we show that the combination of the mappings $X_{ij}(\cdot)$ and $\Psi_i(.)$ implies that our framework is observationally equivalent to a general one-factor neoclassical economy with a flexible form of market-level external economies of scale.[29]

[29]Technology levels that are proportional to population are also postulated in Kortum (1997); Eaton and Kortum (2001). Models with spatial diffusion of knowledge, as in Fujita, Krugman, and Venables (1999) and Lucas and Rossi-Hansberg (2003), specify cross-location spillovers of knowledge so that $\psi_{ij} > 0$ even if $i \neq j$. Our environment can also easily accommodate models with multiple sectors that differ in terms of market structure and strength of economies of scale – e.g., Krugman and Venables (1995), Balistreri, Hillberry, and Rutherford (2010), Kucheryavyy, Lyn, and Rodríguez-Clare (2016).

$$\begin{aligned}(1+\epsilon)\log\hat{p}_{ii}-\textstyle\sum_j\left(y_{ij}+\epsilon\sum_l y_{il}x_{jl}\right)\log\hat{p}_{jj} &= \log\hat{\tau}_i^E-\left(1+(1+\psi)\alpha^2\right)\left(\log\hat{\zeta}_i-\textstyle\sum_j y_{ij}\log\hat{\zeta}_j\right)\\ &- \qquad (1+\psi)\alpha^2\left(\log\hat{Q}_i-\textstyle\sum_j y_{ij}\log\hat{Q}_j\right)\end{aligned}$$

Thus,

$$\begin{gathered}(1+\epsilon+(1+\psi)\lambda)\log\hat{p}_{ii}-\textstyle\sum_j\left((1+\psi)\alpha^2x_{ji}+\left(1+(1+\psi)\alpha^2\right)y_{ij}+\left(\epsilon-(1+\psi)\alpha^2\right)\sum_l y_{il}x_{jl}\right)\log\hat{p}_{jj}=\\ \left[\log\hat{\tau}_i^E+(1+\psi)\alpha^2\left(\log\hat{\tau}_i^I-\textstyle\sum_j y_{ij}\log\hat{\tau}_j^I\right)\right]-\left(1+(1+\psi)\alpha^2\right)\left(\log\hat{\zeta}_i-\textstyle\sum_j y_{ij}\log\hat{\zeta}_j\right)\end{gathered}$$

Define $\alpha^4\equiv(1+\epsilon+(1+\psi)\alpha^2)^{-1}$. Thus,

$$\begin{aligned}\log\hat{p}_{ii}-\textstyle\sum_j b_{ij}\log\hat{p}_{jj} &= \alpha^4\left[\log\hat{\tau}_i^E+(1+\psi)\alpha^2\left(\log\hat{\tau}_i^I-\textstyle\sum_j y_{ij}\log\hat{\tau}_j^I\right)\right]\\ &- \qquad \alpha^4(1+(1+\psi)\alpha^2)\left(\log\hat{\zeta}_i-\textstyle\sum_j y_{ij}\log\hat{\zeta}_j\right)\end{aligned}\tag{52}$$

where

$$b_{ij}\equiv\alpha^4\alpha^2(1+\psi)x_{ji}+\alpha^4\left(1+(1+\psi)\alpha^2\right)y_{ij}+\alpha^4\alpha^2\left(\epsilon-(1+\psi)\alpha^2\right)\sum_l y_{il}x_{jl}$$

Define the matrix

$$\bar{\boldsymbol{B}}=\bar{\boldsymbol{M}}\left(\alpha^4\alpha^2(1+\psi)\bar{\boldsymbol{x}}+\alpha^4\left(1+(1+\psi)\alpha^2\right)\bar{\boldsymbol{y}}+\alpha^4\alpha^2\left(\epsilon-(1+\psi)\alpha^2\right)\bar{\boldsymbol{y}}\bar{\boldsymbol{x}}\right)\bar{\boldsymbol{M}}'\tag{53}$$

Expression (52) can be written as

$$\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{B}}\right)\log\hat{\boldsymbol{p}}=\alpha^4\log\hat{\boldsymbol{\tau}}^E-\alpha^4(1+\psi)\alpha^2\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}\right)\log\hat{\boldsymbol{\tau}}^I-\alpha^4(1+(1+\psi)\alpha^2)\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}\right)\log\hat{\boldsymbol{\zeta}}.$$

If $\epsilon>(1+\psi)\alpha^2$ and $\phi\psi<1$, then $\left(I-\bar{\boldsymbol{B}}\right)$ is non-singurlar and $\left(I-\bar{\boldsymbol{B}}\right)^{-1}=\sum_{d=0}^{\infty}\bar{\boldsymbol{B}}^d$. To see this, notice that, in this case, $b_{ij}>0$, which implies that, for all i,

$$\sum_{j\neq m}|b_{ij}|=\alpha^4\alpha^2(1+\psi)(1-x_{mi})+\alpha^4\left(1+(1+\psi)\alpha^2\right)(1-y_{im})+\alpha^4\left(\epsilon-(1+\psi)\alpha^2\right)(1-x_{mi}),$$

$$\sum_{j\neq m}|b_{ij}|=\alpha^4\left(1+(1+\psi)\alpha^2\right)(1-y_{im})+\kappa\epsilon(1-x_{mi}),$$

$$\sum_{j\neq m}|b_{ij}|=1-\alpha^4\left(\alpha^2\left(1+(1+\psi)\lambda\right)y_{im}+\epsilon x_{mi}\right)<1,$$

where the inequality follows from the fact that gravity structure implies $x_{ji}>0$ for all i and j. We define

$$\bar{\boldsymbol{\Gamma}}\equiv\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{B}}\right)^{-1}.$$

This implies that

$$\log\hat{\boldsymbol{p}}=\bar{\boldsymbol{\Gamma}}\left(\alpha^4\log\hat{\boldsymbol{\tau}}^E+\alpha^4(1+\psi)\lambda\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}\right)\log\hat{\boldsymbol{\tau}}^I-\alpha^4(1+(1+\psi)\alpha^2)\left(\bar{\boldsymbol{I}}-\bar{\boldsymbol{y}}\right)\log\hat{\boldsymbol{\zeta}}\right).$$

Let us set $\log \hat{\boldsymbol{\zeta}} = 0$. Thus,

$$\log \hat{\boldsymbol{Q}} = \alpha^4 \bar{\boldsymbol{\Gamma}}^E \log \hat{\boldsymbol{\tau}}^E - \bar{\boldsymbol{\Gamma}}^I \log \hat{\boldsymbol{\tau}}^I. \tag{54}$$

where

$$\bar{\boldsymbol{\Gamma}}^E \equiv \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \quad \text{and} \quad \bar{\boldsymbol{\Gamma}}^I \equiv \bar{\boldsymbol{I}} - \alpha^4(1+\psi)\alpha^2 \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{x}}\right) \bar{\boldsymbol{\Gamma}} \left(\bar{\boldsymbol{I}} - \bar{\boldsymbol{y}}\right).$$

Applying (54) into (50),

$$\log \hat{L}_i = \alpha^2\alpha^4 \sum_j \gamma_{ij}^E \log \hat{\tau}_j^E + \alpha^2 \sum_j \gamma_{ij}^I \log \hat{\tau}_j^I + \xi\left(\hat{\boldsymbol{\tau}}^I, \hat{\boldsymbol{\tau}}^E\right)$$

where

$$\xi\left(\hat{\boldsymbol{\tau}}^I, \hat{\boldsymbol{\tau}}^E\right) \equiv -\alpha^2 \sum_j \alpha_j^3 \left(\alpha^4 \sum_j \gamma_{ij}^E \log \hat{\tau}_j^E + \sum_j \gamma_{ij}^I \log \hat{\tau}_j^I\right)$$

Thus, to get expression (25), we just have to define $\beta^1 \equiv \alpha^2\alpha^4$, $\beta^2 \equiv \alpha^2$, $\tilde{\beta}_j^1 \equiv \alpha^2\alpha_j^3\alpha^4$, $\tilde{\beta}_j^2 \equiv \alpha^2\alpha_j^3$, $\beta^3 \equiv \alpha^4\alpha^2(1+\psi)$, $\beta^4 \equiv \alpha^4\left(1+(1+\psi)\alpha^2\right)$, and $\beta^5 \equiv \alpha^4\alpha^2\left(\epsilon - (1+\psi)\alpha^2\right)$. Finally, notice that $\alpha_j^3 = 0 \implies \xi\left(\hat{\boldsymbol{\tau}}^I, \hat{\boldsymbol{\tau}}^E\right) = 0$, if $\tilde{\phi}$,$\tilde{\psi} = 0$.■

B Proofs for Section 4

B.1 Proof of the Optimal IV in Expression (31)

The asymptotic variance of the GMM estimator for any function $H_i(\cdot)$ is

$$V(H) = \left(E\left[H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right) G_i\right]\right)^{-1} \left(E\left[H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right) e_i e_i' H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right)'\right]\right) \left(E\left[H_i\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right) G_i\right]\right)^{-1\prime} \tag{55}$$

where $G_i \equiv E\left[\nabla_{\boldsymbol{\theta}} e_i(\boldsymbol{\theta}) | \hat{\boldsymbol{z}}^t, \boldsymbol{W}^{t_0}\right]$.

The asymptotic variance of the Optimal IV estimator in (31) is

$$V(H^*) = E\left[G_i' \left(\Omega\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right)\right)^{-1} G_i\right] \tag{56}$$

with $\Omega\left(\hat{\boldsymbol{z}}^{\boldsymbol{t}}, W^{t_0}\right) = E\left[e_i(\boldsymbol{\theta}) e_i(\boldsymbol{\theta})' | \hat{\boldsymbol{z}}^{\boldsymbol{t}}, W^{t_0}\right]$.

We now show that $V(H) - V(H^*)$ is positive semi-definite for any $H_i(\cdot)$:

$$\begin{aligned} V(H) - V(H^*) &= \left(E\left[H_i^t G_i^t\right]\right)^{-1} \left(E\left[\left(H_i^t e_i^t\right)\left(H_i^t e_i^t\right)'\right]\right) \left(E\left[H_i^t G_i^t\right]\right)^{-1\prime} - \left(E\left[G_i^{t\prime} \Omega^{-1} G_i^t\right]\right)^{-1} \\ &= \left(E\left[H_i^t G_i^t\right]\right)^{-1} \left(E\left[\left(H_i^t e_i^t\right)\left(H_i^t e_i^t\right)'\right] - E\left[H_i^t G_i^t\right] \left(E\left[G_i^{t\prime} \Omega^{-1} G_i^t\right]\right)^{-1} E\left[H_i^t G_i^t\right]'\right) \left(E\left[H_i^t G_i^t\right]\right)^{-1}. \end{aligned}$$

Let us define

$$U_i^t \equiv H_i^t e_i^t - E\left[\left(H_i^t e_i^t\right)\left(G_i^{t\prime} \Omega^{-1} e_i^t\right)'\right] \left(E\left[G_i^{t\prime} \Omega^{-1} G_i^t\right]\right)^{-1} G_i^{t\prime} \Omega^{-1} e_i^t,$$

which implies that

$$E\left[U_i^t U_i^{t\prime}\right] = E\left[\left(H_i^t e_i^t\right)\left(H_i^t e_i^t\right)'\right] - E[H_i^t G_i^t] \left(E\left[G_i^t \Omega^{-1} G_i^t\right]\right)^{-1} E[H_i^t G_i^t]'.$$

Therefore,

$$V(H) - V(H^*) = \left(E\left[H_i^t G_i^t\right]\right)^{-1} \left(E\left[U_i^t U_i^{t\prime}\right]\right) \left(E\left[H_i^t G_i^t\right]\right)^{-1\prime}.$$

Since $E\left[U_i^t U_i^{t\prime}\right]$ is positive semi-definite, $V(H) - V(H^*)$ is also positive semi-definite. Therefore, the asymptotic variance is minimized at H^*.■

B.2 Proof of the Two-Stage Procedure in Proposition 5

We use the strategy in Section 6.1 of Newey and McFadden (1994) to establish asymptotic properties of two-step estimators. To this end, we define the joint moment equation for the two estimating steps:

$$\left(\hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_1\right) \equiv \arg\min_{\theta_2, \theta_1} \left(\sum_{i,t} v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right)\right)' \left(\sum_{i,t} v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right)\right) \tag{57}$$

where

$$v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right) \equiv \left[\begin{array}{cc} H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}_1) e_i^t(\boldsymbol{\theta}_2) & H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}_0) e_i^t(\boldsymbol{\theta}_1) \end{array}\right]$$

$$H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\tilde{\boldsymbol{\theta}}) \equiv E\left[\nabla_{\boldsymbol{\theta}} e_i^t(\tilde{\boldsymbol{\theta}})|\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0\right] \left(\Omega_i^t\right)^{-1}.$$

We have that $\left(\hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_1\right) \xrightarrow{p} (\boldsymbol{\theta}, \boldsymbol{\theta})$, with an asymptotic variance given by

$$Var\left(\hat{\boldsymbol{\theta}}_2, \hat{\boldsymbol{\theta}}_1\right) = \left(\tilde{G}'\tilde{\Omega}^{-1}\tilde{G}\right)^{-1}$$

where $\tilde{G} \equiv \left[\nabla_{(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1)} v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right)\right]$ and $\tilde{\Omega} \equiv E\left[\left(v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right)\right)\left(v_i^t\left(\boldsymbol{\theta}_2, \boldsymbol{\theta}_1\right)\right)'\right]$.

Define $h_i^t \equiv H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}) e_i^t(\boldsymbol{\theta})$ and $\bar{h}_i^t \equiv H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}_0) e_i^t(\boldsymbol{\theta})$. Thus, $\tilde{G}$ and $\tilde{\Omega}$ are given by

$$\tilde{\Omega} = E\left[\begin{array}{cc} h_i^t h_i^{t\prime} & h_i^t \bar{h}_i^{t\prime} \\ \bar{h}_i^t h_i^{t\prime} & \bar{h}_i^t \bar{h}_i^{t\prime} \end{array}\right] \quad \text{and} \quad \tilde{G} = \left[\begin{array}{cc} G & G_1 \\ 0 & G_2 \end{array}\right]$$

where

$$G \equiv E\left[H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta})\right]$$

$$G_1 \equiv E\left[\nabla_{\boldsymbol{\theta}} H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}) e_i^t(\boldsymbol{\theta})\right]$$

$$G_2 \equiv E\left[H_i^*(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^0|\boldsymbol{\theta}_0) \nabla_{\boldsymbol{\theta}} e_i^t(\boldsymbol{\theta})\right].$$

By Assumption 4.3, any function of $\left(\hat{\boldsymbol{z}}^t, \boldsymbol{W}^{t_0}\right)$ is orthogonal to $e_i^t(\boldsymbol{\theta})$, which implies that $G_1 = 0$. Thus, $\left(\tilde{G}'\tilde{\Omega}^{-1}\tilde{G}\right)^{-1}$ is block diagonal and the marginal distribution of $\hat{\boldsymbol{\theta}}_2$ is asymptotically normal with variance

$$Var\left(\hat{\boldsymbol{\theta}}_2\right) = \left(G'\Omega^{-1}G\right)^{-1},$$

which is equivalent to the asymptotic distribution of the Optimal IV in (56). ■

C Empirical Application Appendix

C.1 Derivations in Section 5.1

Bilateral Trade Demand. In sector s of region d, we impose that preferences have the following nested CES form:

$$C_{sd} = \left[\sum_k \left(C_{k,sd}\right)^{\frac{\chi-1}{\chi}}\right]^{\frac{\chi}{\chi-1}} \quad \text{where} \quad C_{k,sd} = \left[\sum_o \left(C_{ko,sd}\right)^{\frac{\chi_k-1}{\chi_k}}\right]^{\frac{\chi_k}{\chi_k-1}}$$

where $\chi_k, \chi > 0$.

The first-stage cost minimization problem is

$$\min_{C_{ko,sd}} \sum p_{ko,sd} C_{ko,sd} \quad \text{s.t.} \quad \left[\sum_k \left(\sum_o \left(C_{ko,sd}\right)^{\frac{\chi_k-1}{\chi_k}}\right)^{\frac{\chi_k}{\chi_k-1}\frac{\chi-1}{\chi}}\right]^{\frac{\chi}{\chi-1}} = 1.$$

The first-order condition implies that

$$\left[\sum_k \left(\sum_o \left(C_{ko,sd}\right)^{\frac{\chi_k-1}{\chi_k}}\right)^{\frac{\chi_k}{\chi_k-1}\frac{\chi-1}{\chi}}\right]^{\frac{\chi}{\chi-1}-1} \left(\sum_o \left(C_{ko,sd}\right)^{\frac{\chi_k-1}{\chi_k}}\right)^{\frac{\chi_k}{\chi_k-1}\frac{\chi-1}{\chi}-1} \left(C_{ko,sd}\right)^{-\frac{1}{\chi_k}} = \lambda p_{ko,sd}$$

Thus,

$$\left(\frac{C_{ko,sd}}{C_{ko',sd}}\right) = \left(\frac{p_{ko,sd}}{p_{ko',sd}}\right)^{-\chi_k}$$

and

$$\frac{\left(p_{ko',sd}\right)^{\chi_k}}{\left(p_{k'o',sd}\right)^{\chi_{k'}}} \frac{\left(\sum_o \left(p_{ko,sd}\right)^{1-\chi_k}\right)^{-\frac{\chi-\chi_k}{\chi_k-1}}}{\left(\sum_o \left(p_{k'o,sd}\right)^{1-\chi_k}\right)^{-\frac{\chi-\chi_{k'}}{\chi_{k'}-1}}} = \left(\frac{C_{ko',sd}}{C_{k'o',sd}}\right)^{-1}.$$

Thus,

$$x_{k'o',sd} = \frac{p_{k'o',sd} C_{k'o',sd}}{\left[\sum_k \left(\left(\sum_o \left(p_{ko,sd}\right)^{1-\chi_k}\right)^{\frac{1}{1-\chi_k}}\right)^{1-\chi}\right]^{\frac{1}{1-\chi}}} = \frac{\left(p_{k'o',sd}\right)^{1-\chi_{k'}}}{\sum_o \left(p_{k'o,sd}\right)^{1-\chi_k}} \frac{\left(\sum_o \left(p_{k'o,sd}\right)^{1-\chi_k}\right)^{\frac{1-\chi}{1-\chi_{k'}}}}{\sum_k \left(\left(\sum_o \left(p_{ko,sd}\right)^{1-\chi_k}\right)^{\frac{1}{1-\chi_k}}\right)^{1-\chi}}$$

Defining $P_{k,sd} = \left[\sum_o \left(p_{ko,sd}\right)^{1-\chi_k}\right]^{\frac{1}{1-\chi_k}}$, we get that

$$x_{k'o',sd} = \left(\frac{p_{k'o',sd}}{P_{k',sd}}\right)^{1-\chi_k} \frac{\left(P_{k',sd}\right)^{1-\chi}}{\sum_k \left(P_{k,sd}\right)^{1-\chi}}.$$

Labor Supply. Assume that the representative household preferences have the following

nested structure, if $\sum_{ko} l_{ko} = 1$,

$$U_c = \left[\sum_d \nu_d \left(\sum_s V_{sd} \left(l_{sd}\right)^{\frac{\phi_e - 1}{\phi_e}}\right)^{\frac{\phi_m \phi_e}{\phi_m \phi_e - 1}}\right]^{\frac{\phi_m \phi_e - 1}{\phi_m \phi_e - \phi_e}} \quad \text{where} \quad V_{sd} = \left(\nu_{sd} \frac{C_{sd}}{l_{sd}} - \left(h_{sd}\right)^{\frac{1+\phi_h}{\phi_h}}\right)^{\frac{1}{1+\phi_h}}$$

and $U_c = -\infty$ whenever $\sum_{ko} l_{ko} \neq 1$.

In sector s of region d, the total labor supply is $L_{sd} = h_{sd} l_{sd}$ and, using the spending shares in (37), the real consumption is $C_{sd} = \omega_{sd} h_{sd} l_{sd}$. Thus, the representative household solves the following second-stage problem:

$$\max_{\{l_{kd}, h_{sd}\}} \left[\sum_d \nu_d \left(\left(l_{0d}\right)^{\frac{\phi_e - 1}{\phi_e}} + \sum_s V_{sd} \left(l_{sd}\right)^{\frac{\phi_e - 1}{\phi_e}}\right)^{\frac{\phi_m \phi_e}{\phi_m \phi_e - 1}}\right]^{\frac{\phi_m \phi_e - 1}{\phi_m \phi_e - \phi_e}}$$

subject to

$$V_{sd} = \nu_{sd} \rho_{sd} \omega_{sd} h_{sd} - \left(h_{sd}\right)^{\frac{1+\phi_h}{\phi_h}}$$

$$\sum_{ko} l_{ko} = 1.$$

The first-order condition for h_{sd} implies

$$h_{sd} = \left(\tilde{\nu}_{sd} \rho_{sd} \omega_{sd}\right)^{\phi_h},$$

so that

$$V_{sd} = \tilde{\nu}_{sd} \rho_{sd} \omega_{sd}.$$

let μ be Lagrange multiplier of the additify constraint. The first-order condition for l_{sd} is

$$\kappa_c \tilde{w}_d \left(l_{0d}\right)^{-\frac{1}{\phi_e}} = \mu$$

$$\kappa_c \tilde{w}_d \tilde{\nu}_{sd} \omega_{sd} \left(l_{sd}\right)^{-\frac{1}{\phi_e}} = \mu$$

where

$$\kappa_c \equiv \left[\sum_d \nu_d \left(\left(l_{0d}\right)^{\frac{\phi_e - 1}{\phi_e}} + \sum_s \left(\nu_{sd} \rho_{sd} \omega_{sd} h_{sd} - \left(h_{sd}\right)^{\frac{1+\phi_h}{\phi_h}}\right)^{\frac{1}{1+\phi_h}} \left(l_{sd}\right)^{\frac{\phi_e - 1}{\phi_e}}\right)^{\frac{\phi_m \phi_e}{\phi_m \phi_e - 1}}\right]^{\frac{\phi_m \phi_e - 1}{\phi_m \phi_e - \phi_e} - 1}$$

$$\tilde{w}_d \equiv \nu_d \left(\left(l_{0d}\right)^{\frac{\phi_e - 1}{\phi_e}} + \sum_s \left(\nu_{sd} \rho_{sd} \omega_{sd} h_{sd} - \left(h_{sd}\right)^{\frac{1+\phi_h}{\phi_h}}\right)^{\frac{1}{1+\phi_h}} \left(l_{sd}\right)^{\frac{\phi_e - 1}{\phi_e}}\right)^{\frac{\phi_m \phi_e}{\phi_m \phi_e - 1} - 1}.$$

Thus,

$$\frac{l_{sd}}{l_{0d}} = \left(\tilde{\nu}_{sd} \omega_{sd}\right)^{\phi_e}$$

and

$$\frac{l_{0d}}{l_{0d}+\sum_s l_{sd}} = \frac{1}{1+\left(\tilde{\nu}_{sd}\omega_{sd}\right)^{\phi_e}}.$$

The first-order condition for l_{sd} also implies that

$$\frac{\tilde{w}_d}{\tilde{w}_0} = \left(\frac{l_{0d}}{l_{00}}\right)^{\frac{1}{\phi_e}}$$

such that

$$\tilde{w}_d = \nu_d \left(\left(l_{0d}\right)^{\frac{\phi_e-1}{\phi_e}} + \sum_s \tilde{\nu}_{sd}\omega_{sd}\left(l_{sd}\right)^{\frac{\phi_e-1}{\phi_e}} \right)^{\frac{\phi_m\phi_e}{\phi_m\phi_e-1}-1}$$

$$\tilde{w}_d = \nu_d \left(\left(l_{0d}\right)^{-\frac{1}{\phi_e}} \left(l_{0d} + \sum_s l_{sd} \right) \right)^{\frac{1}{\phi_m\phi_e-1}}.$$

Thus,

$$\frac{\nu_d}{\nu_0} \left(\left(\frac{l_{0d}}{l_{00}}\right)^{-\frac{1}{\phi_e}} \left(\frac{l_{0d}+\sum_s l_{sd}}{l_{00}+\sum_s l_{s0}}\right) \right)^{\frac{1}{\phi_m\phi_e-1}} = \left(\frac{l_{0d}}{l_{00}}\right)^{\frac{1}{\phi_e}}$$

$$\left(\frac{l_{0d}+\sum_s l_{sd}}{l_{00}+\sum_s l_{s0}}\right) = \left(\frac{\nu_d}{\nu_0}\right)^{(\phi_m\phi_e-1)} \left(\frac{l_{0d}}{l_{00}}\right)^{\phi_m}.$$

C.2 Data Construction

C.2.1 World Trade Matrix

We construct a matrix of bilateral sector-level trade flows among 50 US states and 59 foreign countries for 1993, 1997, 2002, 2007 and 2012. To this end, we merge information on bilateral trade flows of 59 countries extracted from the Multi Region Input Output database (MRIO) and information on domestic and foreign trade flows of 50 US states extracted from the Commodity Flow Survey (CFS) and US Census Foreign Trade Database (FTD). We then aggregate the data at the manufacturing and non-manufacturing level. In order to avoid zeros in the trade matrix, we merge DC and Maryland into a single state, and consider the set of countries in Table 3. We proceed in four steps.

First, we construct foreign trade flows of US states for each sector, year and foreign country. Let $\left(Z_{dj}^{kt}, Z_{jd}^{kt}\right)$ denote the trade flows between each of the 40 US custom districts, d, and foreign country, j, by sector k and year t. We obtain $\left(Z_{dj}^{kt}, Z_{jd}^{kt}\right)$ from the US Merchandise Trade Files released annually by the US Census between 1990 and 2016. The exports and imports of state i to foreign country j are

$$X_{ij}^{kt} = \sum_d a_i^{dj,kt} \cdot Z_{dj}^{kt}$$

$$X_{ji}^{kt} = \sum_d b_i^{dj,kt} \cdot Z_{jd}^{kt}$$

where $a_i^{dj,kt}$ and $b_i^{dj,kt}$ correspond to the share of total exports and imports in district d whose respective origin and destination are state i. We normalize the size of international trade flows so that the total value of export shipments in the CFS and total value of exports in the FTD are equal in 2012.

Second, we construct bilateral trade flows between US states for each sector and year. Let $\tilde{X}_{ir}^{kt}$ denote the value of shipments from state i to state r of goods in sector k at year t. We obtain $\tilde{X}_{ir}^{kt}$ from the Commodity Flow Survey released by the US Census in 1993, 1997, 2002, 2007 and 2012. The trade flow between state i to state r are

$$X_{ir}^{kt} = \tilde{X}_{ir}^{kt} - \sum_{d,j} \left(\tilde{a}_{ir}^{dj,kt} \cdot Z_{dj}^{kt} + \tilde{b}_{ir}^{dj,kt} \cdot Z_{jd}^{kt} \right)$$

where $\tilde{a}_{ir}^{dj,kt}$ and $\tilde{b}_{ir}^{dj,kt}$ correspond respectively to the share of total exports and imports in district d transiting between states i and r.

Third, we adjust domestic sales of the residual sector to include local spending in services:

$$X_{ii}^{NT,t} = \left(\sum_{k \neq NT} \sum_{r} X_{ri}^{kt} \right) e_i^t$$

where e_i^t is the expenditure ratio between non-tradeable and tradeable goods of state i at year t obtained from the BEA state-level accounts. This adjustment is equivalent to ignoring trade between states in goods and services excluded from the CFS, $X_{ir}^{NT,t} = 0$ for $i \neq r$.

Fourth, we merge the trade bilateral trade flows of US states with the bilateral trade flows of the US and other countries in the MRIO database. To this end, we use US domestic sales in the MRIO to normalize total expenditures of US states on goods produced from other US states. We also distribute the bilateral trade flows of the US in the MRIO among US states using each state share in total trade flows to/from other foreign countries obtained in step 1.

To compute the variables above, we assume that the transit route is the same for all export and import of all sectors with identical state of origin/destination, port of exit/entry, and foreign country of origin/destination. Using the US Census data on state of origin exports by port and destination, we compute the following variables:

$$a_i^{dj,kt} = b_i^{dj,kt} = \frac{\text{exports}_i^{dj,t}}{\sum_l \text{exports}_l^{dj,t}} \quad \text{and} \quad \tilde{a}_{ir}^{dj,kt} = \tilde{b}_{ir}^{dj,kt} = \frac{\text{exports}_{ij}^{dj,t}}{\sum_{r,l} \text{exports}_{rl}^{dj,t}}.$$

C.2.2 Labor Market Data

We use the Current Population Surveys - Merged Outgoing Rotation Groups (CPS-MORG) to construct labor market outcomes in 50 US states and 2 sectors. We consider the sample all individuals aged between 16 and 64 years in the survey, and we follow the cleaning procedure of Autor, Katz, and Kearney (2008) to adjust for top censoring, outliers, and time consistency of variables. For each state and sector, we compute the nominal hourly wage as the weighted average of the weekly earnings divided by the number of weekly hours across employed individuals, where individual weights correspond to the number of worked hours times the sampling weights. For each state and sector, we also compute the average hours worked as the weighted average number of

Table 3: Sample of Countries

Sample of Countries			
ARG	Argentina	KAZ	Kazakhstan
AUS	Australia	KOR	South Korea
AUT	Austria	MEX	Mexico
BAL	Baltic Republics	MYS	Malaysia
BGR	Bulgaria	NAF	North Africa
BLR	Belarus	NLD	Netherlands
BLX	Benelux	NOR	Norway
BRA	Brazil	NZL	New Zealand
CAN	Canada	PAK	Pakistan
CHE	Switzerland	PHL	Philippines
CHL	Chile	POL	Poland
CHN	China	PRT	Portugal
COL	Colombia	ROAF	Rest of Africa
CZE	Czech Republic	ROAS	Rest of Asia
DEU	Germany	ROC	Rest of Central America
DNK	Denmark	ROE	Rest of Europe
ESP	Spain	ROS	Rest of South America
FIN	Finland	ROU	Romania
FRA	France	RUS	Russia
GBR	United Kingdom	SAU	Saudi Arabia
GRC	Greece	SGP	Singapore
HRV	Croatia	SVK	Slovakia
HUN	Hungary	SWE	Sweden
IDN	Indonesia	THA	Thailand
IND	India	TWN	Taiwan
IRL	Ireland	UKR	Ukraine
IRN	Iran	URY	Uruguay
ISL	Iceland	VEN	Venezuela
ITA	Italy	ZAF	South Africa
JPN	Japan		

Notes: Baltic Republics includes Estonia, Lithuania and Latvia; North Africa includes Algeria, Egypt, Ethiopia, Morocco, Tunisia; Rest of Africa includes all the remaining African countries; Rest of Asia includes all the countries from Asia and Oceania not listed in the table; Rest of Central America includes all countries from Central America not listed in the table; Rest of Europe includes all countries from Europe not listed in the table; Rest of South America includes all countries from South America not listed in the table.

weekly hours of employed individuals where individual weights correspond to sampling weights. Finally, we use individual sampling weights to compute total number of employed individuals in each sector and state. The total sector-level employment is the average number of hours worked times the total number of employed individuals, and the total home-sector employment is the total number of individuals either unemployed or out of the labor force.

C.3 Algorithm to solve for the equilibrium in changes

For every year t, we have to solve for the endogenous variables of the model, conditional on the shocks $\hat{\tau}_{ij}^{k,t}$, $\hat{\xi}_r^{k,t}$, $\hat{\nu}_r^{k,t}$, and $\hat{\zeta}_r^{k,t}$, as well as on the initial observable conditions, which are: Y_i^{k,t_0}, the total production in country or region i; x_{ij}^{k,t_0}, the bilateral trade flows; E_i^{k,t_0}, the expenditure in country or region i; $n_r^{t_0}$ the share of working age population in region r; l_r^{k,t_0}, the employment share

in sector k and region r.

The algorithm we implement to solve the multiple sectors model in changes is as follows.

1) Guess $\hat{w}_r^{k,t}$, $\hat{P}_i^{k,t}$. Create $\hat{P}_r^t = \Pi_k \left(\hat{P}_r^{k,t}\right)^{e_r^{k,t_0}}$, where $e_i^{k,t0} \equiv E_i^{k,t0} / \sum_k E_i^{k,t0}$ is the expenditure share of sector k in region i, and the real wage $\hat{\omega}_r^{k,t} = \frac{\hat{w}_r^{k,t}}{\hat{P}_r^t}$.

2) Create the labor market variables $\hat{L}_i^{k,t}$, $\hat{h}_r^{k,t}\hat{n}_r^t$, $\hat{l}_r^{k,t}$:

$$\hat{h}_r^{k,t} = \left(\hat{\omega}_r^{k,t}\hat{\zeta}_r^{k,t}\right)^{\phi}$$

$$\hat{l}_r^{k,t} = \frac{\left(\hat{\nu}_r^k \hat{\omega}_r^k\right)^{\kappa}}{\sum_{s=0}^{K} l_r^{s,t_0} \left(\hat{\nu}_r^s \hat{\omega}_r^s\right)^{\kappa}}$$

$$\hat{n}_r^t = \frac{\left(\hat{l}_r^{0,t}\right)^{-\frac{\eta}{\kappa}}}{\sum_j n_j^{t_0} \left(\hat{l}_j^{0,t}\right)^{-\frac{\eta}{\kappa}}}$$

$$\hat{L}_r^k = \hat{h}_r^k \cdot \hat{l}_r^k \cdot \hat{n}_r$$

3) Create expenditures. First we obtain the initial trade imbalances in region i as:

$$T_i^{t_0} = E_i^{t_0} - Y_i^{t_0}$$

noting that

$$\sum_i T_i^{k,t_0} = 0$$

for all sectors k. We then hold T_i constant in terms of the numeraire (which is the price in sector 1 in country 1). The expenditure change is

$$\hat{E}_i = \frac{Y_i^{t_0}\hat{Y}_i + T_i^{t_0}\hat{T}_i}{E_i^{t_0}}$$

where

$$\hat{Y}_i = \sum_s y_i^{s,t_0} \cdot \left(\hat{w}_i^s \hat{L}_i^s\right)$$

where $y_i^{k,t} \equiv Y_i^{k,t} / \sum_k Y_i^{k,t}$ is the output share of sector k in region i. Define $\delta_i^{k,t_0} \equiv Y_i^{k,t_0}/E_i^{k,t_0}$. Thus,

$$\hat{E}_i^{k,t} = \delta_i^{k,t_0} \cdot \left(\sum_s y_i^{s,t_0} \cdot \left(\hat{w}_i^s \hat{L}_i^s\right)\right) + \left(1 - \delta_i^{k,t_0}\right) \hat{T}_i^{k,t}$$

and we impose $\hat{T}_i^{k,t} = 1$. With Cobb-Douglas preferences,

$$\hat{E}_i^{k,t} = \hat{E}_i^t$$

Under the assumption of trade balance, $\delta_i^{k,t_0} = 1$ and the expression above becomes

$$\hat{E}_i^k = \sum_s y_i^{s,t_0} \hat{w}_i \hat{L}_i$$

4) Create domestic trade shares

$$\hat{x}_{rr}^{k,t} = \left(\frac{\hat{P}_r^{k,t}}{\hat{P}_r^t} \right)^{\theta^k} \left(\hat{\omega}_r^{k,t} \right)^{-\theta^k} \left(\hat{L}_r^{k,t} \right)^{\psi} \hat{\xi}_r^{k,t}$$

5) Write the labor market clearing and budget balance equations as excess demand functions:

$$F_i^1 \equiv \left(\hat{w}_i^{k,t} \right) - \left(\hat{P}_i^{k,t} \right)^{-\theta^k} \sum_j K_{ij}^{1,k} \left(\hat{P}_j^{k,t} \right)^{\theta^k}$$

$$F_i^2 \equiv \left(\hat{P}_i^{k,t} \right)^{-\theta^k} - \sum_j K_{ji}^{2,k} \left(\hat{P}_j^{k,t} \right)^{-\theta^k}$$

where $K_{ij}^{1,k} \equiv \hat{x}_{ii}^{k,t} x_{ij}^{k,t_0} \left(\hat{\tau}_{ij}^{k,t} \right)^{-\theta^k} \frac{E_j^{k,t_0} \hat{E}_j^k}{\hat{L}_i^{k,t} Y_i^{k,t_0}}$ and $K_{ji}^{2,k} \equiv \hat{x}_{jj}^{k,t} \cdot x_{ji}^{k,t_0} \cdot \left(\hat{\tau}_{ji}^{k,t} \right)^{-\theta^k}$.

6) Iterate over $\hat{w}_i^{k,t}$ and $\hat{P}_i^{k,t}$ until the excess demand functions are zero for all sectors and regions.

C.4 Additional Results

Table 4: Agglomeration Functional Forms - Robustness

	ψ_e	ψ_m
Panel A: $\psi_e = -\psi_m$		
	0.354**	-0.354**
	(0.174)	(0.174)
Panel B: $\psi_m = 0$		
	1.734*	
	(0.962)	
Panel C: $\psi_e = 0$		
		-0.649**
		(0.267)

Note. Sample of state-sector pairs (2 sectors x 50 states x 4 time periods 1993-1997, 1997-2002, 2002-2007, 2007-2012). All regressions include period and sector dummies. Standard errors in parenthesis are clustered at the state-level. *** p < .10, ** p < .05, * p < .01

Table 5: Estimation of structural parameters - Robustness

Labor Supply Elasticity			Agglomeration Elasticity
Hours	Between sectors	Between regions	
ϕ_h	ϕ_e	ϕ_m	ψ
Panel A: Time and sector dummies			
0.358***	0.686**	0.458	0.354**
(0.137)	(0.313)	(0.483)	(0.174)
Panel B: Include initial state demographic controls			
0.325**	0.605*	0.646	0.438*
(0.127)	(0.323)	(0.583)	(0.262)
Panel C: Include initial state demographic controls interacted with period dummies			
0.319***	0.863**	0.528	0.379*
(0.117)	(0.376)	(0.666)	(0.196)
Panel D: Weight states by population in 1992			
0.319***	0.863**	0.528	0.379*
(0.117)	(0.376)	(0.666)	(0.196)
Panel E: Include only manufacturing sector			
0.434***	1.29**	0.626	0.256
(0.155)	(0.584)	(0.498)	(0.200)
Panel F: Drop great recession period of 2007-2012			
0.306***	1.067**	0.634	0.428*
(0.109)	(0.447)	(0.502)	(0.220)
Panel G: CFS data without inputation for supressed values			
0.331***	1.072**	0.667	0.353
(0.108)	(0.460)	(0.510)	(0.241)

Note. Sample of state-sector pairs (2 sectors x 50 states x 4 time periods 1993-1997, 1997-2002, 2002-2007, 2007-2012). All regressions include period and sector dummies. Initial state demographic controls: female labor force participation, share of college graduates in labor force, share of manufacturing employees in labor force, share of non-manufacturing employees in labor force. Standard errors in parenthesis are clustered at the state-level. *** p < .10, ** p < .05, * p < .01

Table 6: Model Fit - Robustness

Labor Supply Elasticity			Agglomeration Elasticity
Hours	Between sectors	Between regions	
ϕ_h	ϕ_e	ϕ_m	ψ
Panel A: Time and sector dummies			
0.358***	0.686**	0.458	0.354**
(0.137)	(0.313)	(0.483)	(0.174)
Panel B: Include initial state demographic controls			
0.325**	0.605*	0.646	0.438*
(0.127)	(0.323)	(0.583)	(0.262)
Panel C: Include initial state demographic controls interacted with period dummies			
0.319***	0.863**	0.528	0.379*
(0.117)	(0.376)	(0.666)	(0.196)
Panel D: Weight states by population in 1992			
0.319***	0.863**	0.528	0.379*
(0.117)	(0.376)	(0.666)	(0.196)
Panel E: Include only manufacturing sector			
0.434***	1.29**	0.626	0.256
(0.155)	(0.584)	(0.498)	(0.200)
Panel F: Drop great recession period of 2007-2012			
0.306***	1.067**	0.634	0.428*
(0.109)	(0.447)	(0.502)	(0.220)
Panel G: CFS data without inputation for supressed values			
0.331***	1.072**	0.667	0.353
(0.108)	(0.460)	(0.510)	(0.241)

Note. Sample of state-sector pairs (2 sectors x 50 states x 4 time periods 1993-1997, 1997-2002, 2002-2007, 2007-2012). All regressions include period and sector dummies. Initial state demographic controls: female labor force participation, share of college graduates in labor force, share of manufacturing employees in labor force, share of non-manufacturing employees in labor force. Standard errors in parenthesis are clustered at the state-level. *** p < .10, ** p < .05, * p < .01

D Proofs of Equivalences and Extensions

In the first sub-section we provide the proofs of the equivalence of exisiting models with our Generalized Spatial Economy. We show the equivalence with, respectively: Neoclassical models with economies of scale, New Trade Theory models, New Economic Geography models, Spatial Assignment models, and Spatial Assignment models with others factors of production

D.1 Equivalences

D.1.1 Neoclassical Economy

Environment. Consider a neoclassical economy with a single factor of production. As in the Generalized Spatial Economy, each country has a representative agent with preferences for consumption and labor supply in different markets, with utility function given by

$$U_c\left(\left\{\nu_j C_j\right\}_j, \left\{L_j\right\}_j\right).$$

The main difference is that we explicitly allow for preferences over goods, z:

$$C_j \equiv \bar{V}\left(\left\{c_{z,ij}\right\}_{z,i}\right)$$

where $V(.)$ is twice differentiable, quasi-concave, homothetic, and increasing in all arguments. Notice that $V(.)$ allows for the possibility that goods from different origins to be imperfect substitutes.

The representative household's budget constraint is

$$\sum_i \sum_z p_{z,ij} c_{z,ij} = \rho_j w_j L_j.$$

There are many perfectly competitive firms supplying each good in any market. The production technology uses only labor and entails external economies of scale at the market level. In particular, the technology of producing good z in i and delivering to j is given by

$$Y_{z,ij} = \Psi_i\left(\boldsymbol{L}\right)\zeta_i \frac{L_{z,ij}}{\tau_{ij}\alpha_{z,ir}},$$

where α_{ir}^z is good-specific productivity shifter of producing in i and delivering in j.

Equilibrium. We use the fact that $V(.)$ is homothetic to derive the price index in market j:

$$P_i = \bar{P}_i\left(\left\{p_{z,ij}\right\}_{i,z}\right) \equiv \min_c \sum_{i,z} p_{z,ij} c_{z,ij} \quad \text{s.t.} \quad V\left(\left\{c_{z,ij}\right\}_{z,i}\right) \geq 1 \tag{58}$$

where the associated spending share on good z from i is

$$x_{ij,z} \in \bar{X}_{ij,z}\left(\left\{p_{z,ij}\right\}_{i,z}\right). \tag{59}$$

Conditional on prices, the representative household solves the utility maximization problem that yields the labor supply in market j:

$$L_j = \Phi_j\left(\left\{\nu_i \rho_i \omega_i\right\}_i\right). \tag{60}$$

Profit maximization implies that

$$p_{z,ij} = \tau_{ij} p_{ii} \alpha_{z,ij} \tag{61}$$

where

$$p_{ii} = \frac{w_i}{\Psi_i(\boldsymbol{L})\,\zeta_i} \tag{62}$$

Finally, the labor market clearing condition is

$$w_i L_i = \sum_j \sum_z x_{z,ij} \cdot \rho_j w_j L_j.$$

Using the price equation 62, this becomes

$$p_{ii} e_i = \sum_j x_{ij} \cdot \rho_j p_{jj} e_j \tag{63}$$

where $e_i \equiv \zeta_i L_i \Psi_i(\boldsymbol{L})$ is the effective labor supply in i, and $x_{ij} \equiv \sum_z x_{z,ij}$ is the spending share of j on all goods produced in i.

The competitive equilibrium corresponds to $\left\{\{p_{z,ij}\}_{z,j}, w_i, L_i, P_i\right\}$ such that equations 61–63 hold. Thus, the equilibrium can be written as $\{p_{ii}, \omega_i, L_i, Q_i\}$ solving 7–**??** with $\Phi_j(.)$, $\Psi_j(.)$, and

$$\tilde{X}_{ij}\left(\{\tau_{ij} p_{ii}\}_i\right) \equiv \left\{ x_{ij} = \sum_z x_{ij,z} : \; x_{ij,z} \in \bar{X}_{ij,z}\left(\{\tau_{ij} p_{ii} \alpha_{z,ij}\}_{i,z}\right) \right\}$$

such that

$$P_i\left(\{\tau_{ij} p_{ii}\}_i\right) = \bar{P}_i\left(\{\tau_{ij} p_{ii} \alpha_{z,ij}\}_{i,z}\right).$$

Equivalence. We now construct a Generalized Spatial Competitive Economy that, for any $(\boldsymbol{\tau}, \boldsymbol{\zeta}, \boldsymbol{\nu})$, entails identical equilibrium levels of the endogenous variables, $\{\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{P}\}$, as those obtained in the equilibrium of the Neoclassical Economy just described. Since the equilibrium of the Neoclassical Economy consists of vectors $\{p_{ii}, Q_i, L_i, \omega_i\}_i$ solving (6)–(10) with $\left(\Phi_j(.), \Psi_j(.), \tilde{X}_{ij}(.)\right)$, we only need to show that there are preferences and technology in the Generalized Spatial Competitive Economy that are consistent with $\left(\Phi_j(.), \Psi_j(.), \tilde{X}_{ij}(.)\right)$. We assume that the agglomeration function and the labor supply function for the Generalized Economy are the same as above. We also assume that the production function of the market-specific composite good is

$$Y_{ij} = \Psi_i(\boldsymbol{L})\,\zeta_i \frac{L_{ij}}{\tau_{ij}}.$$

In addition, consider the preferences in Section 2 with

$$V_j(\{c_{ij}\}) \equiv \max_{\{c_{z,ij}\}_{z,i}} \bar{V}\left(\{c_{z,ij}\}_{z,i}\right) \quad \text{s.t.} \quad \textstyle\sum_z \alpha_{z,ij} c_{z,ij} = c_{ij} \;. \tag{64}$$

Intuitively, the preference structure in (64) implies that, if the representative household acquires c_{ij} units of $i's$ composite good for j's consumption, then it optimally allocates the composite good into the production of different goods, given the exogenous weights $\alpha_{z,ij}$ that are now embedded into

the representative agent's preferences. Since the relative price of goods in market i only depends on $\alpha_{z,ij}$, this decision yields allocations that are identical to those in the competitive equilibrium of the decentralized economy. To see this, denote the spending shares associated with the cost minimization problem with $V_j(.)$ by $x_{ij} \in X_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right)$. Thus, the equivalence follows from

$$X_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right) = \tilde{X}_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right) \quad \forall \boldsymbol{\tau}. \tag{65}$$

To show this, we consider the dual utility maximization problem. We proceed in two steps. First, we show that $x_{ij} \in X_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right) \implies \exists x_{ij,z} \in \bar{X}_{ij,z}\left(\{\tau_{ij}p_{ii}\alpha_{z,ij}\}_{i,z}\right)$ with $x_{ij} = \sum_z x_{ij,z}$. Let $\left\{c^*_{z,ij}\right\}_{z,i}$ be the solution of the consumer's maximization problem in the decentralized economy:

$$\left\{c^*_{z,ij}\right\}_{z,i} \in \arg\max_{\boldsymbol{c}} \bar{V}\left(\{c_{z,ij}\}_{z,i}\right) \quad \text{s.t.} \quad \sum_i \tau_{ij}p_{ii}\sum_z \alpha_{z,ij}c_{z,ij} = 1. \tag{66}$$

Define $c^*_{ij} \equiv \sum_z \alpha_{z,ij}c^*_{z,ij}$. We want to show that

$$\left\{c^*_{ij}\right\} \in \arg\max V_j\left(\{c_{ij}\}\right) \quad \text{s.t.} \quad \sum_i \tau_{ij}p_{ii}c_{ij} \leq 1,$$

which implies that $x^*_{ij} \equiv p_{ij}c^*_{ij} \in X_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right)$. We proceed by contradiction. Suppose there exists a $\{\tilde{c}_{ij}\}$ such that

$$V_j\left(\{\tilde{c}_{ij}\}\right) > V_j\left(\left\{c^*_{ij}\right\}\right)$$

$$\sum_i p_{ij}\tilde{c}_{ij} \leq \sum_i p_{ij}c^*_{ij} = 1$$

By the definition of $V_j(.)$ in (64),

$$V_j\left(\{\tilde{c}_{ij}\}\right) > V_j\left(\left\{c^*_{ij}\right\}\right) \geq \bar{V}\left(\left\{c^*_{z,ij}\right\}_{z,i}\right) \tag{67}$$

To establish the contradiction, let $\{\tilde{c}_{z,ij}\}$ be the good allocation associated with $\{\tilde{c}_{ij}\}$ in the definition of $V_j(.)$ in (64). Thus,

$$\sum_i p_{ij}\sum_z \alpha_{z,ij}\tilde{c}_{z,ij} = \sum_i p_{ij}\tilde{c}_{ij} \leq 1$$

which implies that $\{\tilde{c}_{ij}\}$ is in the feasibility set of problem (66) and, therefore,

$$\bar{V}\left(\left\{c^*_{z,ij}\right\}_{z,i}\right) \geq \bar{V}\left(\{\tilde{c}_{z,ij}\}_{z,i}\right) = V_j\left(\{c_{ij}\}\right) \tag{68}$$

This contradicts the inequality in (67).

Second, we show that $x_{ij} = \sum_z x_{ij,z},\ x_{ij,z} \in \bar{X}_{ij,z}\left(\{\tau_{ij}p_{ii}\alpha_{z,ij}\}_{i,z}\right) \implies x_{ij} \in X_{ij}\left(\{\tau_{ij}p_{ii}\}_i\right)$. Let $\left\{c^*_{ij}\right\}$ be the optimal solution of the consumer's maximization problem

$$\left\{c^*_{ij}\right\} \in \arg\max V_j\left(\{c_{ij}\}\right) \quad \text{s.t.} \quad \sum_i p_{ij}c_{ij} = 1. \tag{69}$$

Also, let $\left\{c^*_{z,ij}\right\}_{z,i}$ be the solution associated with $\left\{c^*_{ij}\right\}$ in the definition of $V_j(.)$ in (66). We want to show that

$$\left\{c^*_{z,ij}\right\} \in \arg\max_{\boldsymbol{c}} \bar{V}\left(\left\{c_{z,ir}\right\}_{z,i}\right) \quad \text{s.t.} \quad \sum_i p_{ij} \sum_z \alpha_{z,ij} c_{z,ij} = 1,$$

which implies $x^*_{ij} = \sum_z x_{ij,z}$, with $x_{ij,z} \in \bar{X}_{ij,z}\left(\left\{\tau_{ij} p_{ii} \alpha_{z,ij}\right\}_{i,z}\right)$.

We proceed by contradiction. Suppose that there exists $\left\{\tilde{c}_{z,ij}\right\}_{z,i}$ such that

$$V\left(\left\{\tilde{c}_{z,ij}\right\}_{z,i}\right) > \bar{V}\left(\left\{c^*_{z,ij}\right\}_{z,i}\right) \tag{70}$$

$$\sum_i p_{ij} \sum_z \alpha_{z,ij} \tilde{c}_{z,ij} \leq \sum_i p_{ij} \sum_z \alpha_{z,ij} c^*_{z,ij} = \sum_i p_{ij} c^*_{ij} = 1.$$

Notice however that $\tilde{c}_{ij} \equiv \sum_z \alpha_{z,ij} \tilde{c}_{z,ij}$ is a feasible allocation in problem (69) and, therefore,

$$\bar{V}\left(\left\{c^*_{z,ij}\right\}_{z,i}\right) = V_j\left(\left\{c^*_{ij}\right\}\right) \geq V_j\left(\left\{\tilde{c}_{ij}\right\}\right) \geq \bar{V}_j\left(\left\{\tilde{c}_{z,ij}\right\}_{z,i}\right).$$

This contradicts the inequality in (70). Since we have found preferences and technology that imply the mappings $\left(\Phi_j(.), \Psi_j(.), \tilde{X}_{ij}(.)\right)$, we have proven the equivalence.

D.1.2 New Trade Theory

Environment. As in the Generalized Spatial Economy, the representative agent in each country has preferences for consumption and labor supply in different markets (i.e., sector-region pairs), with utility function given by

$$U_c\left(\left\{\nu_j C_j\right\}_j, \left\{L_j\right\}_j\right).$$

We assume that C_j has a nested preference structure across sectors, $C_j = V\left(\left\{C_{s,j}\right\}_s\right)$ with $V(.)$ strictly quasi-concave and homogeneous of degree one. Sectors are divided into two groups: competitive sectors, $s \in K^C$, and monopolistic competitive sectors, $k \in K^M$.

In any sector $s \in K^C$, firms in each country produce one homogeneous good with the production technology in (4). We assume that $C_{s,j}$ is an aggregator of goods from different origins, $C_{s,j} \equiv V_{s,j}\left(\left\{c_{s,rj}\right\}_r\right)$, where $V_{s,j}(.)$ is twice differentiable, increasing, quasi-concave, and homogeneous of degree one. Notice that the utility function allows the goods produced in different regions to be perfect substitutes and, therefore, it covers homogeneous goods.

In any sector $k \in K^M$, there is large mass of potential entrants in each region that produce a differentiated good, indexed by z, and operate in monopolistic competition. We assume that all potential entrants in sector-region $i \equiv (k, r)$ have access to the same increasing returns technology, with the cost of producing $Y_i(z)$ units of the differentiated good z given by

$$\bar{c}_i\left(Y(z)\right) = w_i \cdot \left(\frac{Y_i(z)}{\zeta_i \cdot \Psi^P_i\left(\boldsymbol{L}\right)} + \mu_i \cdot \Psi^E_i\left(\boldsymbol{L}\right)\right)$$

where w_i is the wage rate faces by firms in sector-region i, $\mu_i \cdot \Psi^E_i\left(\boldsymbol{L}\right)$ is a fixed entry cost, and $\zeta_i \cdot \Psi^P_i\left(\boldsymbol{L}\right)$ is the productivity in region-sector i. We explicitly allow $\Psi^P_i\left(\boldsymbol{L}\right)$ and $\Psi^E_i\left(\boldsymbol{L}\right)$ to depend on employment, but we assume that firms perceive $\Psi^P_i\left(\boldsymbol{L}\right)$ and $\Psi^E_i\left(\boldsymbol{L}\right)$ as given. So, these function incorporate external agglomeration and congestion forces at the market level.

We also assume that, for $k \in K^M$, preferences are CES across the available differentiated goods with elasticity $\sigma > 1$:

$$C_{k,j} = \left[\int_{z \in Z_{k,j}} (c(z))^{\frac{\sigma}{\sigma-1}} \, dz \right]^{\frac{\sigma-1}{\sigma}},$$

where $Z_{k,j}$ is the set of k sector goods available in market j.

Equilibrium. First, consider a competitive sector $s \in K^C$. Since $V_{s,j}(.)$ is homogeneous of degree one, we can define the price index:

$$P_{s,j}\left(\{p_{s,oj}\}_o\right) \equiv \min \sum_r c_{s,rj} p_{s,rj} \quad \text{s.t.} \quad V_{s,j}\left(\{c_{s,ij}\}_i\right) = 1 \tag{71}$$

with associated spending shares

$$\bar{x}_{s,rj} \in \bar{X}_{s,rj}\left(\{p_{s,oj}\}_o\right). \tag{72}$$

In addition, profit maximization implies that

$$p_{s,rj} = \tau_{s,rj} \frac{w_{s,r}}{\zeta_{s,r} \Psi^P_{s,r}(\boldsymbol{L})} \tag{73}$$

In sector-region $i = (s,r)$, the labor market clearing condition implies that the total labor payments are equal to total revenue,

$$w_i L_i = \sum_j \left(\bar{x}_{ij} g_{s,j}\right) \cdot \rho_j w_j L_j \tag{74}$$

where $g_{s,j}$ is market $j's$ spending share on goods of sector s.

Second, consider a monopolistic competitive sector $k \in K^M$. We focus on the symmetric equilibrium where all firms of sector k in region r choose the same price, $p_{ij}(z) = \frac{\sigma}{\sigma-1} \frac{\tau_{ij} w_i}{\zeta_i \cdot \Psi^M_i(\boldsymbol{L})}$, with $\Psi^M_i(\boldsymbol{L})$ potentially different from $\Psi^P_i(\boldsymbol{L})$. Thus, the price index of goods of sector k for consumers in j is

$$P^{1-\sigma}_{k,j} = \sum_r p^{1-\sigma}_{k,rj}$$

where, for the sector-region pair $i = (k,r)$,

$$p^{1-\sigma}_{ij} = M_i \left(\frac{\sigma}{\sigma-1} \frac{\tau_{ij} w_i}{\zeta_i \cdot \Psi^P_i(\boldsymbol{L})} \right)^{1-\sigma},$$

with M_i denoting the mass of operational firms in market i.

To characterize M_i, consider the labor market clearing condition in $i \equiv (k,r)$,

$$w_i L_i = M_i \cdot \left(\frac{w_i}{\zeta_i \cdot \Psi^M_i(\boldsymbol{L})} \cdot \sum_j \tau_{ij} \left(\frac{\sigma}{\sigma-1} \frac{1}{\zeta_i \cdot \Psi^P_i(\boldsymbol{L})} \tau_{ij} w_i \right)^{-\sigma} \frac{E_{k,j}}{(P_{k,j})^{1-\sigma}} \right) + M_i \cdot \mu_i \cdot \Psi^E_i(\boldsymbol{L}),$$

and the free entry condition in market in $i \equiv (k,r)$,

$$\frac{1}{\sigma} \sum_j \left(\frac{\sigma}{\sigma-1} \frac{1}{\zeta_i \cdot \Psi^P_i(\boldsymbol{L})} \tau_{ij} w_i \right)^{1-\sigma} \frac{E_{k,j}}{(P_{k,j})^{1-\sigma}} = w_i \mu_i \cdot \Psi^E_i(\boldsymbol{L}).$$

By substituting the free entry condition into the labor market clearing conditions, we obtain

that $M_i = \frac{1}{\sigma\mu_i} \cdot \frac{L_i}{\Psi_i^E(\boldsymbol{L})}$. Thus, in sector $k \in K^M$, the spending share of j in goods of region r is

$$\bar{x}_{k,rj} = \bar{X}_{k,rj}\left(\{p_{k,oj}\}_o\right) = \left(\frac{p_{k,rj}}{P_{k,j}\left(\{p_{k,oj}\}_o\right)}\right)^{1-\sigma} \tag{75}$$

where

$$P_{k,j}\left(\{p_{k,oj}\}_o\right) = \left(\sum_o p_{k,oj}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{76}$$

$$p_{k,rj} = \tau_{k,rj}\frac{w_{k,r}}{\zeta_{k,r}\tilde{\Psi}_{k,r}(\boldsymbol{L})} \tag{77}$$

with

$$\tilde{\Psi}_i(\boldsymbol{L}) \equiv \left(\frac{\sigma}{\sigma-1}\frac{1}{\Psi_i^P(\boldsymbol{L})}\right)\left(\frac{1}{\sigma\mu_i}\cdot\frac{L_i}{\Psi_i^E(\boldsymbol{L})}\right)^{\frac{1}{1-\sigma}}. \tag{78}$$

Using these expressions, it is straightforward to show that the labor market clearing condition in sector k of region r, $i = (k, r)$, is

$$w_i L_i = \sum_j \left(\bar{x}_{ij} g_{k,j}\right) \cdot \rho_j w_j L_j. \tag{79}$$

where $g_{k,j}$ is market $j's$ spending share on goods of sector k.

Third, consider the representative household's choice of spending in different sectors. Conditional on sector-level price indices, the cost minimization problem is

$$P_j\left(\{P_{s,j}\}_s\right) \equiv \min_{C_{s,j}} \sum_s P_{s,j}C_{s,j} \quad \text{s.t.} \quad V\left(C_{s,j}\right) = 1 \tag{80}$$

with associated spending shares given by

$$g_{s,j}\left(\{P_{s,j}\}_s\right) = \frac{\partial \ln P_j\left(\{P_{s,j}\}_s\right)}{\partial \ln P_{s,j}}. \tag{81}$$

We denote the spending of j on goods from sector s of region r as

$$\tilde{X}_{s,rj}\left(\{p_{ij}\}_i\right) \equiv \bar{X}_{s,rj}\left(\{p_{s,oj}\}_o\right) g_{s,j}\left(\left\{P_{k,j}\left(\{p_{k,oj}\}_o\right)\right\}_k\right).$$

Lastly, as in the Generalized Spatial Competitive Economy, the representative household's problem yields the labor supply in region-sector j,

$$L_j \in \tilde{\Phi}_j\left(\{\nu_i\rho_i\omega_i\}_i\right). \tag{82}$$

The equilibrium of this economy is characterized by $\left\{\{p_{ij}\}_{ij}, w_i, L_i, P_i\right\}$ that solve (80)–(82) for all j, (71)–(74) for $s \in K^C$, (75)–(79) for $s \in K^M$. Thus, the equilibrium consists of vectors $\{p_{ii}, Q_i, L_i, \omega_i\}_i$ solving (6)–(10) with $\Phi_j(.) = \tilde{\Phi}_j(.)$, $\Psi_j(.) = \tilde{\Psi}_j(.)$, and $X_{ij}(.) = \tilde{X}_{s,rj}(.)$.

Equivalence. We now construct a Generalized Spatial Competitive Economy that, for any $(\boldsymbol{\tau}, \boldsymbol{\zeta}, \boldsymbol{\nu})$, the equilibrium entails identical levels of the endogenous variables, $\{\boldsymbol{w}, \boldsymbol{L}, \boldsymbol{P}\}$, as those obtained in the equilibrium of the economy above. Since the equilibrium of the economy above corresponds to $\{p_{ii}, \omega_i, L_i, Q_i\}$ solving (6)–(10) with $\left(\tilde{X}_{ij}(.), \tilde{\Phi}_j(.), \tilde{\Psi}_j(.)\right)$, we only need to show

that there are preferences and technology in the Generalized Spatial Competitive Economy that are consistent with $\left(\tilde{X}_{ij}(.), \tilde{\Phi}_j(.), \tilde{\Psi}_j(.)\right)$.

To this end, consider the environment in Section 2 with the nested preference structure: $U_c\left(\{\nu_j C_j\}_j, \{L_j\}_j\right)$ where $C_j = V\left(\{C_{s,j}\}_s\right)$. Sectors are divided into two groups: K^C and K^M. For $s \in K^C$, we assume that preferences and production are identical to those described above. For $k \in K^M$, we assume that there is a representative competitive firm with a production technology in (4) where the agglomeration function is $\Psi_{k,r}(\boldsymbol{L}) = \tilde{\Psi}_{k,r}(\boldsymbol{L})$ as in (78). In addition, we impose that every market j has CES preferences across k sector goods produced in different regions,

$$V_{k,j}\left(\{c_{k,rj}\}_r\right) \equiv \left[\sum_r (c_{k,rj})^{\frac{\sigma}{\sigma-1}}\right]^{\frac{\sigma-1}{\sigma}}.$$

Thus, the consumer maximization problem implies that the spending share on of j on goods from $i \equiv (k,r)$ is $x_{ij} = \bar{x}_{k,rj} g_{k,j} \in X_{ij}\left(\{p_{k,oj}\}_o\right)$ where $\bar{x}_{k,rj}$ is given by (75), and $g_{k,j}$ is given by (81). Since we have found preferences and technology that imply the mappings $\left(\tilde{X}_{ij}(.), \tilde{\Phi}_j(.), \tilde{\Psi}_j(.)\right)$, we have proven the equivalence.

D.1.3 New Economic Geography

Environment. Consider an economy with production structure and preference for goods identical to those in the New Trade Theory Economy of Section D.1.2. We assume that each country c is populated by a continuum of individuals with identical preferences for goods. These individuals differ in terms of mobility across markets. There are two groups of markets in each country, J_c^I and J_c^M. Market $j \in J_c^I$ is populated by a subset of completely immobile individuals such that

$$L_j = \bar{L}_j \quad \forall j \in J_c^I, \tag{83}$$

Alternatively, there is a mass $\bar{L}_c$ of individuals that is completely mobile across markets $j \in J_c^M$ such that

$$\sum_{j \in J_c^M} L_j = \bar{L}_c. \tag{84}$$

Mobile individuals have identical preferences for being employed in $j \in J_c^M$:

$$U_j = \nu_j \omega_j L_j^{\beta}.$$

where ω_j is the real wage in market j.

Equilibrium. We restrict attention to equilibria with positive employment in every $j \in J^M$, and analyze separately the cases of $\beta \neq 0$ and $\beta = 0$. If $\beta = 0$, any employment allocation is feasible as long as $\nu_i \omega_i = \bar{u}$. Thus, the labor supply is

$$\boldsymbol{L} = \bar{\Phi}(\{\omega_i\}) = \begin{cases} L_j = \bar{L}_c,\ L_i = 0 & \text{if} \quad \nu_j \omega_j > \nu_i \omega_i\ \forall i \in J^M \\ \left\{L_j :\ \sum_j L_j = \bar{L}_c\right\} & \text{if} \quad \nu_i \omega_i = \bar{u}\ \forall i \in J^M \end{cases}. \tag{85}$$

If $\beta \neq 0$, in this case, any $j \in J^M$ with positive employment must have

$$\nu_j \omega_j L_j^{\beta} = \bar{u} \quad \Longrightarrow\ L_j = \left(\frac{\bar{u}}{\nu_j \omega_j}\right)^{\frac{1}{\beta}}$$

From equation (84),

$$L_j = \bar{\Phi}_j\left(\{\nu_i\omega_i\}_i\right) \equiv L_c \frac{\left(\nu_j\omega_j\right)^{-\frac{1}{\beta}}}{\sum_{i\in J_c^M}\left(\nu_i\omega_i\right)^{-\frac{1}{\beta}}} \tag{86}$$

The equilibrium of this economy is $\{p_{ii}, \omega_i, L_i, Q_i\}$ solving (6)–(10) with $\Phi_j(\boldsymbol{\omega}) = \bar{L}_j$ if $j \in J^I$ and $\Phi_j(\boldsymbol{\omega}) = \bar{\Phi}_j(\boldsymbol{\omega})$ if $j \in J^M$.

Equivalence. To establish equivalence, we construct preferences for the the representative household that yield the labor supply function $\Phi_j(.) = \bar{\Phi}_j(.)$. Specifically, consider the following preferences:

$$U_c = \begin{cases} \left[\sum_{j\in J_c^M} C_j\left(L_j\right)^{\frac{1}{1-\beta}-1}\right]^{1-\beta} & \text{if} \quad \sum_{j\in J_c^M} L_j = \bar{L}_c \text{ and } L_i = \bar{L}_i \ \forall i \in J^I \\ -\infty & \text{otherwise} \end{cases}$$

If $\beta = 0$, it is straightforward to see that the solution of the utility maximization problem yields equation 85. If $\beta \neq 0$, the solution of the maximization problem is the same as equation 86. Since we have assumed a production structure and preferences for goods identical to those in the New Trade Theory Economy, the assumptions on technology and consumption aggregator imposed in the previous section imply that the functions X_{ij} and Ψ_{ij} deliver the equivalence.

D.1.4 Spatial Assignment Models

Environment. Suppose that countries are populated by a continuum of individuals, $\iota \in I_c$, that are heterogeneous in terms of preferences and efficiency across markets (i.e, sector-region pairs). We assume individual ι has market specific preferences, $a_j(\iota)$, and market specific efficiency, $e_j(\iota)$. In particular, if employed in market j, we assume that individual ι has homothetic preferences given by

$$U_j(\iota) = a_j(\iota) + \nu_j \cdot U_j\left(\left\{c_{ij}\right\}_j\right),$$

with a budget constraint given by

$$\sum_i p_{ij}c_{ij} = \rho_j w_j \cdot e_j(\iota).$$

We further assume that individuals take independent draws of $(a_j(\iota), e_j(\iota))$ from a common distribution:

$$\left\{a_j(\iota), e_j(\iota)\right\}_j \sim F_c(\boldsymbol{a}, \boldsymbol{e}).$$

On the production side, we maintain the same structure of the Generalized Spatial Economy. That is, there is a representative competitive firm in each market with the production technology in (4).

Equilibrium. We start by characterizing spending shares across markets. Conditional on choosing j, individuals choose spending shares that minimize total cost:

$$P_j\left(\left\{p_{ij}\right\}_j\right) \equiv \sum_i p_{ij}c_{ij} \text{ s.t. } U_j\left(\left\{c_{ij}\right\}_j\right) = 1, \tag{87}$$

with associated spending shares given by

$$x_{ij} \in X_{ij}\left(\left\{p_{ij}\right\}_i\right). \tag{88}$$

The solution of this problem implies that, for individual ι, the utility of being employed in j is $U_j(\iota) = a_j(\iota) + \rho_j \nu_j \omega_j \cdot e_j(\iota)$. Thus, the set of individuals choosing j is

$$I_j\left(\{\rho_i \nu_i \omega_i\}_i\right) \equiv \{(\boldsymbol{a}, \boldsymbol{e}) : \ a_j + e_j \cdot \rho_j \nu_j \omega_j \geq a_i + e_i \cdot \rho_i \nu_i \omega_i \ \forall i\},$$

with the associated labor supply given by

$$L_j = \bar{\Phi}_j\left(\{\rho_i \nu_i \omega_i\}_i\right) \equiv \int_{I_j\left(\{\rho_i \nu_i \omega_i\}_i\right)} e_j \ dF_c(\boldsymbol{a}, \boldsymbol{e}). \tag{89}$$

Notice that the function $\bar{\Phi}_j(.)$ is homogeneous of degree zero with $\frac{\partial \bar{\Phi}_j}{\partial \omega_j} \geq 0$ and $\frac{\partial \bar{\Phi}_j}{\partial \omega_i} \leq 0$.[30] In addition, the profit maximization problem of producers implies that

$$p_{ij} = \tau_{ij} p_{ii} \tag{90}$$

and

$$Q_i \equiv \frac{p_{ii}}{P_i} = \frac{\omega_i}{\Psi_i\left(\{L_j\}_j\right) \zeta_i}. \tag{91}$$

Finally, labor market clearing condition implies that

$$w_i L_i = \sum_j x_{ij} \cdot \rho_j w_j L_j$$

which, by defining $e_i = L_i \Psi_i\left(\{L_j\}_j\right) \zeta_i$, yields

$$p_{ii} e_i = \sum_j x_{ij} \cdot \rho_j p_{jj} e_j. \tag{92}$$

The equilibrium of this economy is characterized by $\{p_{ii}, \omega_i, L_i, Q_i\}$ solving equations (87)–(92). Thus, the equilibrium can be written as $\{p_{ii}, \omega_i, L_i, Q_i\}$ solving (6)–(10) with $(X_{ij}(.), \Psi_j(.))$ and $\Phi_j(.) = \bar{\Phi}_j(.)$.

Equivalence. To establish the equivalence, we propose preferences for the representative household of the Generalized Spatial Economy that yield the labor supply function $\Phi_j(.) = \bar{\Phi}_j(.)$. Specifically, consider the following preferences:

$$U_c = \begin{cases} \sum_j \nu_j C_j + \bar{U}\left(\{L_j\}\right) & \textbf{if} \quad \{L_i\} \in \mathcal{L}_c \\ -\infty & \textbf{otherwise} \end{cases}$$

where $\mathcal{L}_c \equiv \left\{\{L_i\}_i : \ \exists I_i \text{ s.t.} \quad L_i = \int_{I_i} e_j \ dF(a,e)\right\}$ is the set of feasible employment allocations, and

$$\bar{U}\left(\{L_j\}\right) \equiv \max_{\{I_j(a,e)\}} \int \sum_j a_j I_j(a,e) \ dF(a,e) \tag{93}$$

subject to

[30] The homogeneity of $\bar{\Phi}_j(.)$ follows immediately from the definition of I_j. To see that $\frac{\partial \bar{\Phi}_j}{\partial \omega_j} \geq 0$ and $\frac{\partial \bar{\Phi}_j}{\partial \omega_i} \leq 0$, notice that $I_i\left(\tilde{\boldsymbol{\omega}}_c\right) \subset I_i\left(\boldsymbol{\omega}_c\right)$ and $I_j\left(\boldsymbol{\omega}_c\right) \subset I_j\left(\tilde{\boldsymbol{\omega}}_c\right)$ whenever $\tilde{\omega}_j > \omega_j$ and $\tilde{\omega}_i = \omega_i$.

$$L_j = \int I_j(a,e) e_j \; dF(a,e) \; \forall j$$

$$\sum_j I_j(a,e) = 1 \; \forall (a,e),$$

$$I_j(a,e) \geq 0 \; \forall j, \; \forall (a,e).$$

We now show that $\bar{\Phi}_j(.)$ in (89) is the labor supply function implied by the representative agent's problem. We start by substituting the second constraint into the objective function to eliminate $I_i(a,e)$ for an arbitrary i. Then, we consider the problem's Lagrangian:

$$\bar{U}\left(\{L_j\}\right) \equiv \max_{\{I_j(a,e)\geq 0\}} a_i + \int \sum_{j\neq i} \left[\left(a_j + e_j\mu_j - a_i - e_i\mu_i\right) I_j(a,e) - \mu_j L_j + \mu_i L_i \; \right] dF(a,e),$$

where μ_j is the Lagrangian multiplier of the labor supply constraint in market j.

The first-order condition of this problem implies that, for $j \neq i$, $I_j(a,e) = 0$ if $a_i + e_i\mu_i > a_j + e_j\mu_j$. Thus, $I_i(a,e) = 1$ if, and only if, $a_j + e_j\mu_j \leq a_i + e_i\mu_i$. Since i was arbitrarily chosen, we can write

$$\forall i: \; I_i(a,e) = 1 \; \Leftrightarrow \; (a,e) \in I_i\left(\left\{\mu_j\right\}_j\right) \equiv \left\{(\boldsymbol{a},\boldsymbol{e}): \; a_i + e_i\mu_i \geq a_j + e_j\mu_j \; \forall j\right\}.$$

The multiplier vector $\left\{\mu_j\right\}_j$ solves the system of labor supply constraints:

$$L_j = \int_{I_j\left(\left\{\mu_i\right\}_i\right)} e_j \; dF_c(\boldsymbol{e}). \tag{94}$$

Now we are ready to show that $\bar{\Phi}_j(.)$ is the labor supply function implied by the utility maximization problem of the representative household. The representative agent solves the following problem:

$$\max_{\left\{L_j\right\}_j} \sum_j \nu_j\omega_j L_j + \bar{U}\left(\left\{L_j\right\}_j\right)$$

where $\bar{U}(.)$ is given by (93).

The first-order condition of this problem implies that the labor supply function is the solution of the following system:

$$\frac{\partial \bar{U}\left(\left\{\Phi_j\right\}_j\right)}{\partial L_i} = -\rho_i\nu_i\omega_i \quad \forall i.$$

Notice that the Envelope Theorem implies that $\frac{\partial \bar{U}}{\partial L_i} = -\mu_i$ and, therefore, $\mu_i = \rho_i\nu_i\omega_i > 0$. Thus, the labor supply function is

$$\Phi_j\left(\left\{\rho_i\nu_i\omega_i\right\}_i\right) = \int_{I_j\left(\left\{\rho_i\nu_i\omega_i\right\}_i\right)} e_j \; dF_c(\boldsymbol{a},\boldsymbol{e})$$

with

$$I_j\left(\left\{\rho_i\nu_i\omega_i\right\}_i\right) \equiv \left\{(\boldsymbol{a},\boldsymbol{e}): \; a_j + e_j\cdot\rho_j\nu_j\omega_j \geq a_i + e_i\cdot\rho_i\nu_i\omega_i \; \forall i\right\}.$$

This implies that $\Phi_j\left(\left\{\rho_i\nu_i\omega_i\right\}_i\right) = \bar{\Phi}_j\left(\left\{\rho_i\nu_i\omega_i\right\}_i\right)$, and therefore the equivalence is proven.

D.1.5 Spatial Assignment Models with Other Factors in Production

Environment. Consider an economy with a representative household with the preferences in (1)–(2) subject to the budget constraint in (3). We impose additional restrictions on preferences to obtain the equivalence result. First, assume that individuals employed in all sectors of region r have identical preferences $V_{kr}(.) = V_r(.)$. Second, assume that preferences are such that the labor supply function is invertible (up to a scalar). Third, we restrict attention to counterfactual exercises in which ν_j remains constant and, therefore, ν_j can be normalized to one.

In every market i, there is a representative competitive firm that uses labor and other factors in production, with the following Cobb-Douglas production function:

$$Y_i = \tilde{\zeta}_i \tilde{\Psi}_i(\boldsymbol{L}) \cdot (L_i)^{\alpha_i} (T_i)^{1-\alpha_i}. \tag{95}$$

Each region r has an endowment of the other factor, $\bar{T}_r$.We assume that T_i is mobile across sectors within a region, but that it is immobile across regions – like land in spatial models. Similar to Caliendo et al. (2014), there is a national mutual fund that owns the other factor in all regions. We assume that the local government in region r owns a share κ_r of the national fund, and it transfers all dividends to local residents. In particular, we impose that the transfer rate to individuals employed in sector k of region r, ρ_{kr}, is inversely proportional to the share of labor in the total cost of the sector,

$$\rho_{kr} = \rho_r / \alpha_{kr}. \tag{96}$$

Equilibrium. To characterize the equilibrium, it is useful to work with the adjusted wage rate, $\tilde{w}_{k,r} \equiv w_{k,r}/\alpha_{k,r}$. The representative household's cost minimization problem yields spending share and price indices that are given by

$$x_{ij} \in X_{ij}\left(\{p_{ij}\}_i\right) \quad \text{and} \quad P_j = P_j\left(\{p_{ij}\}_i\right). \tag{97}$$

As in Section (2), the utility maximization problem of the representative household yields the labor supply function. Using the transfer rule in (96), the labor supply in j is

$$L_{so} \in \Phi_{so}\left(\{\rho_r \tilde{\omega}_{kr}\}_{kr}\right). \tag{98}$$

Notice that the invertibility assumption implies that, for any arbitrary reference sector-region pair sr, we can write

$$\frac{\rho_o \tilde{\omega}_{ko}}{\rho_r \tilde{\omega}_{sr}} = \Phi^{-1}_{sr,ko}\left(\{L_i\}_i\right).$$

In addition, the profit maximization problem of firms implies that

$$p_{ij} = \tau_{ij} p_{ii}$$

where

$$p_{ii} = \frac{\tilde{w}_i}{\zeta_i \tilde{\Psi}_i(\boldsymbol{L})} \cdot \left(\frac{R_i}{\tilde{w}_i}\right)^{1-\alpha_i}$$

where R_i is the price of other factor faced by the producer in i, and $\zeta_i \equiv \tilde{\zeta}_i (1-\alpha_i)^{(1-\alpha_i)}$.

To obtain the equilibrium level of R_i, consider the market clearing condition for the other factor

in region r, $\bar{T}_r = \sum_k T_{kr}$, which implies that

$$R_r = \frac{1}{\bar{T}_r} \sum_k (1 - \alpha_{kr}) \, \tilde{w}_{kr} L_{kr}$$

We use this expression to eliminate R_i in the expression of p_{ii} for sector s in region r with $i \equiv (s, r)$. After some manipulation, we obtain

$$p_{ii} = \frac{\tilde{w}_i}{\zeta_i \tilde{\Psi}_i (\boldsymbol{L})} \cdot \left(\frac{1}{\bar{T}_r} \sum_k (1 - \alpha_{kr}) \frac{\rho_r \tilde{\omega}_{kr}}{\rho_r \tilde{\omega}_{sr}} L_{kr} \right)^{1-\alpha_i}$$

Thus, the invertibility of the labor supply function yields

$$p_{ii} = \frac{\tilde{w}_i}{\zeta_i \Psi_i (\boldsymbol{L})} \tag{99}$$

with

$$\bar{\Psi}_{sr} (\boldsymbol{L}) \equiv \tilde{\Psi}_{sr} (\boldsymbol{L}) \left(\frac{1}{\bar{T}_r} \sum_k (1 - \alpha_{kr}) \, \Phi^{-1}_{sr,kr} (\{L_i\}_i) \, L_{kr} \right)^{\alpha_i - 1}. \tag{100}$$

To close the equilibrium, consider the labor market clearing condition that can be written in terms of the revenue share accruing to labor in every sector-region pair:

$$\tilde{w}_{ko} L_{ko} = \sum_{s,r} x_{ko,sr} \cdot (\rho_r \tilde{w}_{sr} L_{sr}) \, . \tag{101}$$

Finally, the transfer rate in region r is determined by its share in the divided paid by the mutual fund:

$$\rho_r \sum_s \tilde{w}_{sr} L_{sr} = \kappa_r \sum_{k,o} (1 - \alpha_{ko}) \tilde{w}_{ko} L_{ko} \tag{102}$$

where the left hand side is region r's total transfer payments, and the right hand side is region r's share in the total land revenue in the country.

The equilibrium of this economy is characterized by$\{p_{ii}, \tilde{w}_i, L_i, P_i\}$ that solve equations 101–98. Thus, the equilibrium can be written as $\{p_{ii}, \tilde{w}_i, L_i, P_i\}$solving 7–10 with $(X_{ij}(.), \Phi_j(.))$, $\Psi_j(.) = \bar{\Psi}_j (.)$, conditional on the transfer rule $\{\rho_r\}$ in (102).

Equivalence. To establish the equivalence, we consider the Generalized Spatial Competitive Economy of Section 2, with $\Psi_j(.) = \bar{\Psi}_j (.)$ in (100) and the transfer rule in (102). To the extent that the share of labor in production remains constant, the counterfactual changes in wages, $\hat{w}_i$, is identical to the counterfactual change in the normalized wage, $\hat{\tilde{w}}_i$.

This establishes that the Generalized Spatial Competitive Economy is equivalent to spatial assignment models with other factors of production that are mobile across sectors but not across regions – e.g., land and other natural resources. A similar argument yields the equivalence with models with other factors of production that are mobile across both regions and sectors. The only restriction is that the invertibility step to obtain (100) requires the same transfer rate across markets in the country.

D.2 Extensions

D.2.1 Generalized Spatial Competitive Economy with Multiple Labor Types

Consider the Generalized Spatial Competitive Economy. Assume that production requires multiple factors such that

$$Y_{ij} = \frac{\zeta_i \Psi_i \left(\boldsymbol{L}_c\right)}{\tau_{ij}} F_{ij}\left(\left\{L_{ij,g}\right\}_g\right)$$

where $F_{ij}\left(.\right)$ has constant returns to scale.

We write the equilibrium in terms of factor-content of trade as in Adao et al. (2017). To this end, we define the utility function as

$$\bar{U}_j\left(\left\{L_{ij,g}\right\}\right) \equiv U_j\left(\left\{F_{ij}\left(\left\{L_{ij,g}\right\}\right)\right\}\right)$$

and define the effective price of each factor as

$$p_{ij,g} = \tau_{ij,g} \cdot \frac{w_{i,g}}{\zeta_i \Psi_i \left(\boldsymbol{L}_c\right)}.$$

This problem yields spending shares in terms of factor content:

$$x_{ij,g} = X_{ij,g}\left(\left\{p_{ij,g}\right\}_{i,g}\right) \equiv \frac{\partial \log P_j\left(\left\{p_{lj,g}\right\}_l\right)}{\partial \log p_{ij,g}} \tag{103}$$

where P_j is the minimum cost of achieving one unit of the consumption index in j,

$$P_j = P_j\left(\left\{p_{ij,g}\right\}_i\right) \equiv \min_{\{c_{ij}\}} \sum_i p_{ij,g} L_{ij,g} \quad \text{s.t.} \quad U_j\left(\left\{F_{ij}\left(\left\{L_{ij,g}\right\}\right)\right\}\right) = 1. \tag{104}$$

The rest of the analysis remains unchanged. We just have to compute the Factor Content of Trade across markets, which, in this simple model, is straightforward:

$$x_{ij,g} = \alpha_{i,g}\left(\left\{w_{i,g}\right\}\right) \cdot X_{ij}$$

where $\alpha_{i,g}\left(\left\{w_{i,g}\right\}\right)$ is the share of factor g in the production cost of market i. Therefore, the model is isomorphic to the baseline model, except that a market i is defined as a sector-region-group triple.

D.2.2 Generalized Spatial Competitive Economy with Intermediate Goods in Production

Assume that production uses labor and other intermediate goods according to the following production function:

$$Y_{ij} = \frac{\zeta_i}{\tau_{ij}} \Psi_i \left(\boldsymbol{L}_c\right) F\left(L_{ij}, \tilde{F}\left(\left\{M_{li}\right\}_l\right)\right)$$

where $F(.)$ and $\tilde{F}(.)$ are homogeneous of degree one, and M_{li} is the quantity of good l used in production by i.

This production function covers multiple sector gravity models reviewed in Costinot and Rodriguez-Clare (2013). In this class of models, final preferences and production imply identical spending shares across origins within a sector. However, sector-level spending shares differ across industries and final consumption. In quantitative models, it is typically assumed that sector-level

spending shares are constant, and identical to those in Input-Output Tables – for example, see Caliendo and Parro (2014).

Let $\tilde{P}_i$ be the PPI in region-sector pair i. Cost minimization implies that

$$\tilde{P}_i\left(\{p_{li}\}_l\right) \equiv \min_{M_{li}} \sum p_{li} M_{li} \quad \text{s.t.} \quad \tilde{F}\left(\{M_{li}\}_l\right) = 1$$

and the input shares are given by

$$\lambda_{li} \equiv \frac{\partial \ln \tilde{P}_i\left(\{p_{li}\}_l\right)}{\partial \ln p_{li}},$$

with a matrix of input shares given by $\Lambda = [\lambda_{ij}]_{ij}$.

The cost minimization problem in market i is

$$C_i\left(w_i, \tilde{P}_i\right) \equiv \min_{L,\tilde{F}} w_i L + \tilde{P}_i \tilde{F} \quad \text{s.t.} \quad F\left(L, \tilde{F}\right) = 1,$$

implying that

$$p_{ii} = \frac{C_i\left(w_i, \tilde{P}_i\right)}{\zeta_i \Psi_i\left(\boldsymbol{L}_c\right)}$$

Notice that $C_i\left(.\right)$ homogeneous of degree one, with associated labor spending shares given by

$$\alpha_i\left(w_i, \tilde{P}_i\right) \equiv \frac{\partial \ln C_i\left(w_i, \tilde{P}_i\right)}{\partial \ln w_i}$$

Thus, we can write

$$Q_i \equiv \frac{p_{ii}}{P_i}\left[C_i\left(1, \frac{\tilde{P}_i}{w_i}\right)\right]^{-1} = \tau_{ij,g} \cdot \frac{\omega_i}{\zeta_i \Psi_i\left(\boldsymbol{L}_c\right)} \tag{105}$$

To close the model, consider the market clearing condition for labor in each market. The total revenue of market i from sales in market j is

$$X_{ij} = x_{ij}\rho_j w_j L_j + \lambda_{ij}(1-\alpha_j)\sum_d X_{jd}$$

$$X_{ij} = x_{ij}\rho_j w_j L_j + \lambda_{ij}\frac{1-\alpha_j}{\alpha_j} w_j L_j$$

$$X_{ij} = \left(x_{ij}\rho_j + \lambda_{ij}\frac{1-\alpha_j}{\alpha_j}\right) w_j L_j$$

Thus,

$$w_i L_i = \sum_j \alpha_i \left(x_{ij}\rho_j + \lambda_{ij}\frac{(1-\alpha_j)}{\alpha_j}\right) w_j L_j. \tag{106}$$

Notice that the equilibrium requires knowledge of the cost function, $C_i(.)$, and the associated input shares, $\mathbf{\Lambda}$, and labor share, $\alpha_i\left(.\right)$. With this information, we can proceed as before and consider changes in the system of equilibrium equations (7)-(10).

D.2.3 Generalized Spatial Competitive Economy with Commuting

Let L_{jl} be the mass of labor residing in region j and working on region l, with C_{jl} denoting the associated consumption index of these workers. The representative agent's problem is

$$\max_{\left\{\{c_{ijl}\}_i, L_{jl}\right\}_{jl}} U\left(\{\nu_{jl}C_{jl}\}_{j,l}, \{L_{jl}\}_{j,l}\right)$$

subject to

$$C_{jl} = U_j\left(\{c_{ijl}\}\right)$$
$$\sum_i p_{ij}c_{ijl} = w_l L_{jl}$$

As in the baseline model, the first-stage problem implies that

$$P_{jl}\left(\{p_{ij}\}_i\right) \equiv \min_{c_{cji}} \sum_j p_{ij}c_{ijl} \quad \text{s.t.} \quad U_j\left(\{c_{ijl}\}\right) = 1 \tag{107}$$

with the spending shares given by

$$x_{ij} \in X\left(\{p_{oj}\}_o\right) \tag{108}$$

Thus,

$$\sum_i p_{ij}c_{ijl} = P_j C_{jl} \quad \Rightarrow \quad C_{jl} = \frac{w_l}{P_j}L_{jl} = \omega_{jl}L_{jl}$$

where $\omega_{jl} = w_l/P_j$ is the real wage of working in market l and residing in market j. Thus,

$$\max_{\{L_{jl}\}_{jl}} U\left(\{(\nu_{jl}\omega_{jl})\cdot L_{jl}\}_{j,l}, \{L_{jl}\}_{j,l}\right)$$

which yields the labor supply mapping,

$$L_{ji} \in \Phi_{ji}\left(\{\nu_{dl}\omega_{dl}\}_{d,l}\right). \tag{109}$$

The total labor supply in region i is

$$L_i = \sum_j L_{ji} \quad \text{with } L_{ji} \in \Phi_{ji}\left(\{\nu_{dl}\omega_{dl}\}_{d,l}\right).$$

As in the baseline model, we consider the profit maximization problem of firms in market j. This yields the following local competitiveness equation,

$$Q_{ji} \equiv \frac{P_{ii}}{P_j} = \frac{\omega_{ji}}{\zeta_i \Psi_i(L_c)}. \tag{110}$$

Notice that Q_{ji} is a bilateral measure of competitiveness that adjusts the baseline measure by the real exchange rate differences between markets, $Q_{ji} \equiv \frac{P_{ii}}{P_i}\frac{P_i}{P_j}$.

To close the model, we consider the labor market clearing condition: total labor payments to labor in market i equals total revenue of market i from selling to all other markets in the world economy. That is,

$$\sum_i w_i L_{ij} = \sum_j x_{ij} \cdot \left(\sum_l w_l L_{jl} \right)$$

which is equivalent to

$$\sum_i p_{ii} e_{ij} = \sum_j x_{ij} \cdot \left(\sum_l p_{ll} e_{jl} \right) \tag{111}$$

where $e_{ij} = \zeta_i \Psi_i(L_c) L_{ij}$.

The competitive equilibrium in this economy corresponds to $\{p_{ii}, w_i, P_j, L_{ij}\}$ such that conditions (107)–(111) hold. In this case, we need to extend the notion of labor supply to capture commuting flows across markets. In other words, counterfactual predictions require knowledge of the extended labor supply mapping, $L_{ji} \in \Phi_{ji}\left(\{\nu_{dl}\omega_{dl}\}_{d,l}\right)$.

D.2.4 Spatial Generalized Competitive Economy with Selection in Production

TBA