

Trade, Agglomeration Effects, and Labor Markets: Theory and Evidence

Rodrigo Adao
Chicago Booth

Costas Arkolakis
Yale

Federico Esposito
Tufts

July 2017 [preliminary]

Abstract

We develop a unified theoretical and empirical framework to study the impact of trade shocks on local labor markets. We start by explicitly characterizing, in a class of trade and geography models, a structural relationship between trade and labor market outcomes at the regional level. We show that both labor supply and agglomeration elasticities are central in determining this relationship. To identify these key elasticities, we propose a new empirical methodology that uses as an instrument the impact of exogenous trade shocks on changes of the endogenous variables predicted by our general equilibrium model. This methodology yields the most efficient estimator of the structural elasticities, i.e. a Model-implied Optimal IV (MOIV). We then apply our methodology to evaluate the aggregate impact of trade shocks affecting regional labor markets in the U.S.

1 Introduction

The analysis of the labor market consequences of the recent integration of the world economy has risen to the forefront of the debate regarding trade policy design. Recent empirical evidence documents that exposure to international trade shocks is associated with changes in employment and wages across regional labor markets in both developing and developed countries – see e.g. [Topalova \(2010\)](#), [Kovak \(2013\)](#) and [Autor et al. \(2013\)](#). The research design of the existing difference-in-difference empirical approaches addresses many challenges in the identification of the impact of trade shocks on labor markets. However it presents two important shortcomings.¹ First, this approach alone is insufficient to recover the aggregate impact of trade shocks on labor market outcomes, which is absorbed by the time fixed effect. Second, even if the regional exposure measure is exogenous, the empirical results have a limited structural interpretation, making it harder to use existing findings for trade policy evaluation. To address these issues, a growing body of literature has proposed structural frameworks of trade and labor markets.²

In this paper, we develop a unifying theoretical and empirical framework to study the impact of trade shocks on employment and wages. Our approach moves beyond existing structural papers in two important ways. First, in a general environment, we explicitly characterize the regional relationship between trade exposure and labor market outcomes. We show that both the labor supply elasticities and the agglomeration forces are central for its general equilibrium predictions regarding both aggregate and differential effects of international trade shocks on regional labor markets outcomes. Second, we turn to the problem of estimating the structural elasticities of the model governing the relationship between trade and labor markets. We propose a novel empirical methodology that identifies the structural elasticities by using as instrument the impact of *exogenous* trade shocks on the *endogenous* variables predicted by our general equilibrium model. In this theoretical environment, we show that a model-implied IV yields the most efficient estimator of the structural elasticities, i.e. it is the Model-implied Optimal IV (MOIV). We then apply our methodology to evaluate the aggregate impacts of trade shocks affecting regional labor markets in the U.S.

In the first tier of our analysis, we start with a simple model of labor supply and agglomeration, where the elasticities governing these two effects are constant and invariant across regions, and embed it into a standard constant elasticity of substitution (CES) Armington

¹See [Muendler \(2017\)](#) for a review of the evidence on the impact of trade on labor markets and the details of the difference-in-difference empirical approach.

²See, for example, [Burstein et al. \(2015\)](#), [Lee \(2015\)](#), [Galle et al. \(2015\)](#), [Adão \(2015\)](#) and [Caliendo et al. \(2015\)](#).

model. In this environment, we show that the local labor market outcomes are characterized solely by two equations: a labor supply equation that links labor supply to real wages, and a “local relative competitiveness” equation that links trade to real wages and employment. Both these equations have attained prominent position in recent trade theory: the former in evaluating labor market outcomes — for a review, see [Goldberg \(2015\)](#), while the latter in geography models of entry and agglomeration since the seminal work of [Krugman \(1991\)](#) – for a review, see [Redding and Rossi-Hansberg \(2016\)](#). While the literature has mostly focused on each of these channels separately, we show that the link between trade and labor market outcomes depends on their combined strength.

In addition, we show in this simple model that, conditional on the labor supply and agglomeration elasticities, the change in the regional demand-adjusted domestic trade share is a sufficient statistic for the effect of international trade shocks on regional employment and real wages. This property, which is an extension of the result of [Arkolakis et al. \(2012\)](#) for endogenous labor supply, is a key prediction of the model that allows us to discipline the empirical relationship between trade shocks and labor market outcomes. Importantly, it allows us to segment the structure of the model into a “local labor market module”, that combines the local labor supply and competitiveness equations as a function of trade, and a “trade module”, which determines, given labor, the trade flows across regions and countries. This result is key for the identification of the effects of trade shocks, as it implies that these shocks are connected to the local labor module only through the regional trade shares. Therefore our model offers a natural exclusion restriction that can be used in the identification of the parameters.

We next establish that this result is pervasive in a class of trade and geography models featuring endogenous labor supply and agglomeration forces. To do so, we develop a multiple-sector framework with a flexible elasticity structure of bilateral trade flows, which largely generalize existing constant elasticity gravity models. Our environment allows sectoral labor supply in a region to depend on the entire vector of real wages of regions and sectors in the country.³ In addition, we model agglomeration forces without committing to any particular micro-foundation, allowing it to depend on employment on every sector and region of the country.⁴ We show that the sufficient statistics result retains, albeit now the labor market outcomes depends on the assumptions for the labor supply and agglomeration functions and the vector of all trade flows. Importantly, regional trade shares still constitute the only

³Our environment accommodates heterogeneous individuals that can adjust their labor supply in terms of status of labor force participation, hours of work, sector of employment, and region of residence.

⁴Agglomeration effects could be capturing standard forces such as endogenous firm entry, as in the [Krugman \(1980\)](#) and [Melitz \(2003\)](#) model, or productivity spillovers and congestion forces, as in [Krugman \(1991\)](#) and [Allen and Arkolakis \(2014\)](#).

channel of transmission of trade shocks from the “trade module”, i.e. the part of the economy that determines trade flows through labor market clearing and trade balance, to the “local labor market module”.

Having established these theoretical predictions, we turn to the problem of documenting the causal impact of changes in trade flows on employment and real wages, which identify the agglomeration and the labor supply elasticities in our model. The challenge in estimating this causal impact arises from the fact that, in general equilibrium, the trade flows and labor market outcomes are correlated with unobserved local shocks to productivity and labor supply. To circumvent this problem, we exploit the sufficient statistics structure of the model. Specifically, we use the “exact hat-algebra” technique proposed by [Dekle et al. \(2008\)](#) to compute the *predicted* impact of observable trade cost shocks on the endogenous variables in general equilibrium. This *predicted* change in the endogenous variables is then used to construct moment conditions for the estimation of the effect of *actual* changes in the endogenous variables on changes in labor market outcomes across regions.

We show that such Model-Implied IV is *Optimal*, in the sense that it minimizes the asymptotic variance of the estimator, and we further show that it can be easily implemented through a Two-Stage GMM methodology.⁵ The intuition for the optimality result is that, in our general equilibrium model, the predicted change in the endogenous variables corresponds to the total effect of the exogenous trade shock on the endogenous regional outcomes. That is, through the lens of the model, it includes all the general equilibrium channels through which the shock affects regional labor markets, and thus uses all the available information to identify the unknown structural parameters, leading to the most precise estimates.

Our empirical methodology has several advantages. First, the model-implied instrument relies on standard forces in trade models: the region’s direct and indirect exposure to trade cost shocks through the world trade network. It provides a disciplined form of computing regional exposure to aggregate shocks affecting all parts of the country simultaneously. Second, the moment conditions clearly delineate the main exogeneity assumption required for identification: regional labor supply and productivity shocks are mean-independent from the observable trade cost shock and the initial trade network. This exogeneity condition is similar to the exogeneity restrictions that empirical papers in international trade typically make (see [Topalova \(2010\)](#), [Kovak \(2013\)](#), [Ebenstein et al. \(2013\)](#) and [Pierce and Schott \(2016\)](#)).⁶

⁵See [Chamberlain \(1987\)](#) for a seminal paper on optimal instrumental variables.

⁶Notice that our exogeneity restriction is potentially weaker than the one made in other empirical works, such as [Autor et al. \(2013\)](#), which require the entire vector of imports to be mean independent from unobserved local shocks. In addition, our methodology requires weaker assumptions compared to other structural approaches, such as Maximum Likelihood and Simulated Method of Moments. While these approaches rely on assumptions about the entire distribution of the unobserved shocks, our methodology only requires orthogonality restrictions between unobserved local shocks and one observed source of trade cost shocks.

Under this assumption, our methodology leads to consistent estimates of the structural parameters even if our general equilibrium model is mis-specified. Third, the optimality of our methodology is intrinsically connected to the ability of our general equilibrium model to generate predictions that are correlated with actual changes in trade and labor market outcomes across regions. In this sense, the first-stage of this methodology is equivalent to a test of the predictive power of the responses implied by our general equilibrium following the observed trade cost shocks. Lastly, our empirical strategy leads to a simple GMM estimator that can be used for the estimation of structural parameters in a wide class of general equilibrium models.

In the last part of the paper, we apply our methodology to evaluate the impact of trade openness on real wages and employment. We use US regional data on employment, real wages and state trade shares for the years 1997-2012. Using different measures of actual changes in international trade costs, our structural estimation indicates an aggregate labor supply elasticity in line with the literature estimates reviewed by [Chetty \(2012\)](#). The impact of trade outcomes on labor market outcomes yields an agglomeration elasticity which is slightly lower than the unitary agglomeration elasticity implied by [Krugman \(1980\)](#). This number is somewhat larger than recent estimates – see e.g. [Ahlfeldt et al. \(2015\)](#), [Kline and Moretti \(2014\)](#) and [Peters \(2017\)](#). We fail to find significant responses in terms of migration. This result is quite 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 make three contributions to the reduced form literature studying the impact of trade on local labor markets, such as [Topalova \(2010\)](#), [Kovak \(2013\)](#) and [Autor et al. \(2013\)](#). First, we provide a novel direct evidence on the impact of trade openness on local real wages and employment across US states. Second, since our measure of local exposure to international trade shock is the effect predicted by our general equilibrium model, it captures direct and indirect exposure to shocks through the international trade network. This is a significant departure from the existing literature relying on variations of Bartik-like measures of local exposure to international trade shocks. Lastly, our general equilibrium framework allows us to compute *both* the local and aggregate impacts of any counterfactual change in international trade costs.

Our paper is also related to the literature quantifying the labor market consequences of international trade shocks using calibrated gravity models of international trade – e.g., [Galle et al. \(2015\)](#), [Caliendo et al. \(2015\)](#), [Lee \(2015\)](#) and [Burststein et al. \(2016\)](#). Similar to these papers, our general equilibrium model enables us to compute the aggregate impact on labor market outcomes of counterfactual changes in trade costs. However, we depart

from this literature by proposing a framework to empirically assess the implications of our model regarding the impact of trade openness on labor market outcomes. We then use this framework to discipline the magnitude of the labor supply responses and agglomeration forces used in the counterfactual predictions of gravity models. We see our approach as a novel tool to evaluate the empirical content of quantitative models of international trade.

There is a number of recent papers that employ model-consistent instruments to identify the relevant parameters – e.g. [Monte et al. \(2015\)](#), [Donaldson and Hornbeck \(2016\)](#), [Faber and Gaubert \(2016\)](#), [Allen and Arkolakis \(2016\)](#) and [Bartelme et al. \(2017\)](#). The feature that makes our approach distinct is the natural segmentation of the labor market module from the trade module. This allows us to provide precise conditions for the identification of the key structural parameters in the context of a straightforward GMM estimator. In particular, we prove that our estimator is consistent and that is the one that minimizes the asymptotic variance among all possible instrument, i.e. it is an optimal IV (see [Chamberlain \(1987\)](#)). While the study of optimal instruments is new in general equilibrium, some work has been done in partial equilibrium industrial organization contexts (see e.g. [Berry et al. \(1995\)](#) and [Reynaert and Verboven \(2014\)](#)). The difference with these latter approaches is that our analysis can yield a precise analytical characterization of the instrument, while the industrial organization literature has solely relied on simulations.

This paper is structured as follows. Section 2 describes a simple Armington framework with elastic labor supply and agglomeration force, that conveys the intuition for our main results, which are shown in a more general model in Section 3. Section 4 discusses our novel empirical methodology, which we then implement in Section 5 using US state-level data. Armed with the theoretical model and the estimates, in Section 6 we conduct a number of counterfactual exercises. Section 7 concludes.

2 A Simple Model of Agglomeration Effects and Endogenous Labor Supply

In this section we present a simple framework that conveys the intuition for our main results, which are shown in a more general model in Section 3. Specifically, we consider an Armington model with endogenous labor supply and agglomeration forces in production. We show that these features are central for the quantitative predictions of the model regarding the effect of international trade on local labor markets. More generally, we establish a structural relationship between international trade flows and local labor markets outcomes that is

controlled by the assumptions about regional labor supply and agglomeration effects in the regional production function.

2.1 Environment

We begin by discussing the main points of our analysis in a simple model of trade and labor supply. In this very stylized setup, we highlight the key role of the elasticity of labor supply and the degree of agglomeration in determining the labor market outcomes as a function of trade cost shocks.

Preferences. We consider regional economies indexed by i . Each region has a representative household with quasi-linear preferences over consumption and labor:

$$U_i(C_i, L_i) = C_i - \tilde{v}_i \frac{L_i^{1+1/\phi}}{1 + 1/\phi}, \quad (1)$$

where L_i is the labor supply, and C_i is a Constant Elasticity of Substitution (CES) Armington aggregator over the differentiated goods produced by each region:

$$C_i \equiv \left(\sum_j (C_{ji})^{\frac{\varepsilon}{1+\varepsilon}} \right)^{\frac{1+\varepsilon}{\varepsilon}}. \quad (2)$$

with $\varepsilon > 0$.

This separable preference structure yields a two-stage problem for the representative household. In the first-stage problem, the representative household minimizes the cost of the consumption goods conditional on a level of aggregate consumption. This implies that the spending share of region i on the good produced on region j is

$$x_{ij} = \frac{(P_{ij})^{-\varepsilon}}{\sum_{ij} (P_{ij})^{-\varepsilon}}, \quad (3)$$

with the associated CES price index defined as the weighted sum of the local prices,

$$P_i = \left[\sum_j P_{ji}^{-\varepsilon} \right]^{-\frac{1}{\varepsilon}}. \quad (4)$$

In the second-stage problem, the representative household chooses the aggregate levels of labor supply and consumption subject to the regional budget constraint, i.e. $w_i L_i = P_i C_i$.⁷

⁷The budget constraint imposes trade balance: labor is the only source of regional income. Our main

This problem yields the regional labor supply equation:

$$\log L_i = \phi \log \omega_i + \nu_i \quad (5)$$

where $\omega_i \equiv w_i/P_i$ is the real wage of region i , and $\nu_i \equiv -\phi \log \tilde{\nu}_i$ is an unobserved labor supply shifter.

Expression (5) is central for our analysis: it implies that changes in the regional real wage trigger changes in the regional labor supply. In this simple model, the magnitude of this response is regulated by the parameter ϕ , which we assume to be positive, $\phi > 0$. This assumption is consistent with empirical evidence establishing that workers increase their labor supply in response to higher wages, implying a positive elasticity of labor supply – for reviews, see [Keane and Rogerson \(2012\)](#) and [Chetty et al. \(2013\)](#).⁸

Production. We assume that labor is the only factor of production. Shipment of goods involves bilateral iceberg costs denoted by $\tau_{ij} \geq 1$. We assume that the price of good produced in region i and sold in region j is given by

$$P_{ij} = \tau_{ij} \frac{w_i}{\tilde{\xi}_i (L_i)^\psi}. \quad (6)$$

In this expression the term $\tilde{\xi}_j$ is a productivity shifter that captures unobserved local productivity shocks and the term $(L_j)^\psi$ captures scale effects in regional production. The parameter ψ controls the magnitude of these scale effects: it is the elasticity of regional production costs with respect to regional employment. The log-linear agglomeration externality is the simplest choice to model these effects, but it illustrates local scale effects implied by a variety of assumptions in the production structure – for instance, spatial externalities in [Krugman \(1991\)](#) or Marshallian production externalities in [Ethier \(1982\)](#) – that would yield a positive scale coefficient, $\psi > 0$.⁹ Yet, assuming that land or housing is a factor of production, as in [Donaldson and Hornbeck \(2016\)](#), may yield a negative coefficient. The special case of $\psi = 1$ corresponds to the model in [Krugman \(1980\)](#), where the combination of increasing returns and monopolistic competition generates agglomeration forces through firm entry. Since our focus is on the aggregate implications of the interaction between agglomeration

insights hold even in the presence of regional income transfers if these do not affect regional labor supply through income effects. Alternatively, as in [Dekle et al. \(2008\)](#), it is possible to account for exogenous trade imbalances by introducing lump-sum transfers between regions: $E_i = w_i L_i + \kappa_i$ with $\sum_i \kappa_i = 0$.

⁸In the Appendix, we provide alternative micro foundations for the dependence of local employment on local real wages. We show that, under specific functional form assumptions, the same expression arises in an economy populated by heterogeneous individuals in terms of work dis-utility that can adjust both the number of hours worked and labor force participation status.

⁹For a review of the literature on local agglomeration forces, see [Redding and Rossi-Hansberg \(2016\)](#).

forces and elastic labor supply, we follow a recent literature that incorporates the scale effects without committing to any particular micro-foundation (e.g., [Allen and Arkolakis \(2014\)](#)).

We now consider the ratio between the price of domestic goods and all other available goods, P_{ii}/P_i , as a measure of the relative competitiveness of local producers. In this simple model, the production structure in (6) implies that

$$\frac{P_{ii}}{P_i} = \frac{\omega_i}{\tilde{\xi}_i (L_i)^\psi}$$

This equation connects the demand for local production labor to the real wage, regulated by the agglomeration elasticity. It reserves a special role for trade captured by the price term: as the price index for domestic goods consumed domestically increases versus the overall price index, the demand for domestic labor declines and the relationship between local labor and local real wage is adjusted. To complete the picture, we need to substitute out the import demand which, using the gravity equation (3), yields

$$\ln \omega_i = \psi \ln L_i + \ln \tilde{x}_{ii} + \xi_i \tag{7}$$

where $\tilde{x}_{ii} \equiv x_{ii}^{-1/\varepsilon}$, is defined as the demand adjusted trade openness and in this simple model is inversely related to the domestic trade share.

We call equation (7) the “local relative competitiveness” condition. It links real wages and employment to the relative production costs of domestic goods, which depend on the trade links that a location has with other locations through trade. Because of the particular constant elasticity structure of this example, all these links are summarized through the demand-adjusted domestic trade share term \tilde{x}_{ii} . In Section 3, we consider an environment without gravity, in which case the entire vector of trade spending shares of the region determine its demand-adjusted trade openness.

2.2 Local Labor Market Module

The local labor supply and relative competitiveness conditions, equations (5) and (7), constitute the base of our analysis and we refer to them as the “*Local Labor Market Module*”. In this simple example, a powerful intuition arises for the propagation of trade shocks to local labor markets: knowledge of the two key parameters (ϕ, ψ) and the overall demand-adjusted domestic trade share is sufficient to measure the effect of trade on the local labor market module. We summarize this result in the following proposition.

Proposition 1. *Given (ϕ, ψ) , the change in the regional demand-adjusted domestic trade*

share, $\tilde{x}_{ii} \equiv (x_{ii})^{-1/\varepsilon}$, is a sufficient statistics for the effect of international trade shocks on regional employment and real wages, ω_i and L_i .

Proposition 1 is directly related to the result in [Arkolakis et al. \(2012\)](#) that, in constant elasticity gravity models, the domestic trade share is a sufficient statistic for the impact of trade shocks on the real wage. In this simple constant elasticity example, we show that complex trade linkages across countries can be parsimoniously summarized by the simple demand-adjusted trade share, while in the next section we show, in a much richer framework, that these links can be represented by the vector of *all* bilateral trade flows. Most importantly, this result illustrates how foreign shocks, either to trade, productivity or labor supply, propagate to the local market through trade alone.

Applying the implicit function theorem on equations (5) and (7), we have that

$$\frac{d \log \hat{\omega}_i}{d \log \hat{x}_{ii}} = \frac{1}{1 - \phi\psi} \quad \text{and} \quad \frac{d \log \hat{L}_i}{d \log \hat{x}_{ii}} = \frac{\phi}{1 - \phi\psi}. \quad (8)$$

Equation (8) indicates the importance of the strength of scale effects and labor supply responses for the quantitative predictions of the model. The assumptions about the magnitude of these forces are central for the model-implied relationship between changes in regional trade outcomes and regional labor market outcomes. To see this more clearly, assume that either labor supply is exogenous, $\phi = 0$, or that scale effects do not exist, $\psi = 0$. Under these assumptions, equation (8) reduces to the effect of the domestic trade share on real wages in [Arkolakis et al. \(2010\)](#): the trade elasticity controls the magnitude of changes in the domestic trade share on real wage. However, the values for ψ and ϕ affect this prediction of the model. In fact, by changing ϕ and ψ , it is possible to obtain any relationship between the demand-adjusted domestic trade share and local labor market outcomes. Intuitively, the impact of demand-adjusted domestic trade share on the real wage is associated with changes in the labor supply. In the presence of local scale effects, the response in local labor supply translates into responses in local productivity, that further affect real wages.

2.3 World Trade Module

We now turn to the full characterization of the general equilibrium and most importantly the determination of trade flows. To close the model, we impose that the only source of income is labor, which implies that $Y_i = w_i L_i$. Hence, labor market clearing implies the following condition: for all i ,

$$w_i L_i = \sum_j x_{ij} \cdot (w_j L_j). \quad (9)$$

where the gravity equation in (3) implies that

$$x_{ij} = \left(\tau_{ij} \cdot \frac{w_i}{P_j} \right)^{-\varepsilon} \cdot \left(\tilde{\xi}_i L_i^\psi \right)^\varepsilon \quad (10)$$

and

$$P_j^{-\varepsilon} = \sum_i (\tau_{ij} \cdot w_i)^{-\varepsilon} \cdot \left(\tilde{\xi}_i L_i^\psi \right)^\varepsilon. \quad (11)$$

Equations (9) and (11) are the standard equilibrium equations of gravity models of trade and economic geography and as such we define this system as the “Trade Module”.

The equilibrium is defined as the vector $\{w_j, L_j, P_j, \{x_{ij}\}_i\}$ such that equations (5), (7), (9) and (11) hold for every region j . In Appendix 9.5 we follow the strategy in [Alvarez and Lucas \(2007\)](#) to show that bounds on the labor supply and agglomeration elasticities lead to a unique world equilibrium.¹⁰

A crucial feature of our model is that the Labor Market Module directly connects trade and labor market outcomes at the regional-level, conditional on the parameters ϕ and ψ that regulate the strength of labor supply responses and scale effects. However, the only effect of trade costs τ_{ij} on local labor markets is through the demand-adjusted trade openness, \tilde{x}_{ii} . Therefore, any shock to trade barriers is excluded from the Labor Market Module. We will exploit this feature of the model to obtain moment conditions for the estimation of the structural elasticities ϕ and ψ .

3 General Model

This section outlines our main insights in a multiple-sector framework with connected regional labor markets. We introduce macro-level restrictions that link regional trade outcomes to regional labor market outcomes within a country. We demonstrate, in further generality, that this relationship is shaped by the assumptions regarding the strength of labor supply responses and scale effects. For this reason, we argue that it is important to directly estimate such relationship in the data.

¹⁰In particular, we require the labor supply elasticity to be bounded above by the demand elasticity. Intuitively, this requires that the feedback effects of shocks that increase production through increases in labor supply and the agglomeration forces are balanced by strong declines in labor demand. In addition, we require agglomeration forces to be bounded, as in economic geography models (see [Fujita et al. \(1999\)](#) and [Allen and Arkolakis \(2014\)](#)).

3.1 Environment

Consider a world economy composed of countries, c , that are collections of regional labor markets, $r \in R_c$, with multiple sectors indexed by k . Each country is populated by individuals that can choose to supply their labor in any sector and region. We assume that individuals may choose to engage in home production, in which case they supply their labor to the home sector ($k = 0$). Throughout this section, we denote vectors of variables with bold symbols.

Preferences for consumption goods. We assume that individuals have identical homothetic preferences over consumption goods across sectors,

$$C_r = U(C_{r,1}, \dots, C_{r,K})$$

where $C_{r,k}$ is the consumption aggregator of goods in sector k . Denote the prices of goods in sector k , region r , for goods originating from i as $P_{ir,k}$.

Our first assumption imposes restrictions on the bilateral trade flows implied by the sector-level consumption aggregator.

A1. *In sector k , region r 's spending shares on goods produced in all world regions, $\mathbf{x}_{r,k} \equiv (x_{ir,k})_i$, has the following form,*

$$\mathbf{x}_{r,k} = \boldsymbol{\chi}_k(\mathbf{P}_{r,k}) \tag{12}$$

where $\boldsymbol{\chi}_k(\cdot)$ is an invertible function (up to a scalar) of the vector of prices $\mathbf{P}_{r,k} \equiv (P_{ir,k})_i$. Moreover, the unit cost of consumption in sector k and region r is

$$P_{r,k} \equiv f_k(\mathbf{P}_{r,k}). \tag{13}$$

Assumption A1 is central for relating prices to trade shares at the regional level. The restrictions in A1 effectively impose that the sector-level consumption aggregator is homothetic, since sectoral spending shares are independent of the overall sectoral expenditure. Note that, by construction, $f_k(\cdot)$ is homogeneous of degree one. More importantly, A1 also imposes invertibility of sectoral spending shares, which implies that relative local prices can be written solely as a function of local spending shares:

$$\left\{ \frac{P_{jr,k}}{P_{rr,k}} \right\}_j = \tilde{\boldsymbol{\chi}}_k(\mathbf{x}_{r,k}). \tag{14}$$

As discussed in [Berry et al. \(2013\)](#), a sufficient condition for a demand function to be

invertible over its support is that it satisfies the *connected substitutes* property. This property is implied by the gross substitutes property traditionally used to guarantee equilibrium uniqueness in general equilibrium models (Arrow and Hahn (1971)). However, it is more general since it allows for zeros in the matrix of cross-price elasticities as long as there are connected substitutability paths between all goods. In our environment, Assumption A1 allows for a general sector-level substitution pattern across goods supplied by producers in different world regions. Since it involves aggregation at the country-sector level, it is satisfied under certain joint restrictions on preferences and production. For example, it is easy to verify that, under Armington preferences, perfect competition and external increasing returns to scale, any utility function satisfying connected substitutes yields spending shares that satisfy A1. In addition, it is easy to show, following closely the logic of Adao et al. (2017), that a Ricardian multi-country model with a general homothetic utility function over goods and constant returns to scale is isomorphic to the Armington model with spending shares satisfying A1.¹¹

Labor Supply. Our second assumption imposes restrictions on the labor supply in any sector-region of a country.

A2. *In country c , the labor supply vector, $\mathbf{L}_c \equiv \{L_{r,k}\}_{r,k}$ is a function of the effective real wage vector, $\tilde{\omega}_c \equiv \{\nu_{c,k}\omega_{r,k}\}_{r,k}$:*

$$\mathbf{L}_c = \tilde{\Phi}_c(\tilde{\omega}_c) \quad (15)$$

where $\tilde{\Phi}_c(\cdot)$ is an *invertible* function, and $\nu_{r,k}$ is a sector-region labor supply shifter..

There are two central restrictions in Assumption A2. First, it implies that the vector of effective real wages in the entire country is the only variable vector in the model affecting the level of labor supply in any sector-region pair. This assumption means that sector-region labor supply shifters are log-additive in the endogenous real wage. Importantly, notice that this labor supply function allows for non-employment across regions, since it explicitly includes the labor supply to the home sector. Second, A2 imposes that labor supply is invertible. This restriction plays an important role in our empirical strategy, since it allows the recovery of the sector-region labor supply shifter from observable labor market outcomes,

$$\log \nu_{r,k} = \log \omega_{r,k} - \log \Phi_{r,k}(\mathbf{L}_c). \quad (16)$$

Assumption A2 is satisfied by the main constant elasticity trade and geography models existing in the literature, since labor allocations in those models are a function of real wages

¹¹See Lemma 1 in Adao et al. (2017) and Scarf et al. (2003) for details.

across all locations and invertibility is always guaranteed (see [Allen et al. \(2014\)](#)). In particular, it covers the two main extremes in the literature: perfect worker mobility, and no worker mobility, as well as intermediate cases with restricted mobility.¹² Moreover, we show in Appendix 8 that A2 is satisfied by the labor supply function implied by an environment where heterogeneous individuals in terms of sector-region preferences choose the number of hours worked, their sector of employment and region of residency. Our empirical application in Section 5 is based on a parametric version of this micro-founded framework.¹³

Production. As in Section (2), we assume that labor is the only factor of production, and that shipment of goods involves bilateral iceberg costs, τ_{ij} . Our last assumption imposes restrictions on the price-index of good produced in region i and sold in region j .

A3. *The price-index of goods produced in region r and sold in region i has the following form:*

$$P_{ri,k} = \tau_{ri,k} \cdot \frac{w_{r,k}}{\xi_{r,k} \Psi_{r,k}(\mathbf{L}_c)} \quad (17)$$

where $\xi_{r,k}$ is a sector-region productivity shifter, and $\Psi_{r,k}(\mathbf{L}_c)$ is the sector-region function of agglomeration forces.

Similarly to the simple one sector model shown in the previous section, we posit that the price of a good in sector k , produced in region r and sold in i , depends on labor costs $w_{r,k}$, iceberg trade costs $\tau_{ri,k}$, an unobserved productivity shifter $\xi_{r,k}$ and on scale effects in regional production, captured by $\Psi_{r,k}(\mathbf{L}_c)$. The key difference relative to the one sector model is that agglomeration forces in region r and sector k are allowed to depend on the whole vector of labor supply in all regions and sectors. Therefore, we can account for very rich patterns of agglomeration across regions and sectors, that we will discipline in the data.¹⁴

¹²Note that in the case of perfect labor mobility any differences in real wages are ruled out in the sectors and regions where allocations are positive. In the other extreme case of perfectly immobile workers, labor supply does not respond to changes in real wage, so that A2 is trivially satisfied. Intermediate cases of restricted labor mobility are discussed in detail in [Allen and Arkolakis \(2013\)](#); [Redding \(2012\)](#); [Redding and Rossi-Hansberg \(2016\)](#); [Allen et al. \(2014\)](#).

¹³Modeling agents heterogeneity has a long tradition in economics starting from [Roy \(1951\)](#). Recently, it has found use in modeling labor market outcomes and sectoral employment in a number of papers, such as [Caliendo et al. \(2015\)](#), [Galle et al. \(2015\)](#), [Lee \(2015\)](#), [Burstein et al. \(2016\)](#) and [Adão \(2015\)](#).

¹⁴Notice here that richer production patterns can be incorporated in A3 without affecting the main result. For example, a price index that is represented by a homogeneous of degree 1 aggregator in all wages can be incorporated, allowing for entry costs as in [Arkolakis \(2010\)](#). In addition, richer input-output patterns of production would require the price index to be represented by an aggregator not only of wages but also of the price index in different sector-regions, as in [Eaton and Kortum \(2002\)](#); [Caliendo and Parro \(2014\)](#). The key condition on the aggregator in both these examples is that it can be ultimately inverted to be written as a function of all available wages alone. If that is possible, repeating the procedure in Section 3.2 will yield the same estimating equations. We choose to present a simpler formulation of A.3 in order to convey the critical steps for the procedure in a more intuitive way and to keep the notation simple.

3.2 Local Labor Market Module

We now derive the equations relating trade outcomes and labor market outcomes across regions of a country. In each sector, we follow the simple model of Section 2 and consider the ratio between the prices of domestic goods and all other available goods, $P_{rr,k}/P_{r,k}$, as our measure of the relative competitiveness of local producers. This ratio captures the cost of domestic goods compared to the cost of all locally available goods and, for this reason, it is linked to consumption pattern embedded in regional trade flows. In particular, A1 yields a connection between relative prices and spending shares:

$$\frac{P_{rr,k}}{P_{r,k}} = \left[f_k \left(\left\{ \frac{P_{jr,k}}{P_{rr,k}} \right\}_j \right) \right]^{-1} = [f_k(\tilde{\boldsymbol{\chi}}_k(\mathbf{x}_{r,k}))]^{-1} \quad (18)$$

where the first equality follows from the homogeneity of $f_{r,k}(\cdot)$, and the second equality form the inverse demand function in (14).

We can use this equation to obtain a structural relationship between trade and labor market outcomes. The price index (17) in A3 implies that

$$\frac{P_{rr,k}}{P_{r,k}} \frac{P_{r,k}}{P_r} = \frac{\omega_{r,k}}{\xi_{r,k} \Psi_{r,k}(\mathbf{L}_c)}$$

Thus,

$$\log \omega_{r,k} - \log \tilde{\boldsymbol{\alpha}}_{r,k} = \log \Psi_{r,k}(\mathbf{L}_c) + \log \xi_{r,k}$$

where the demand-adjusted trade openness $\tilde{\boldsymbol{\alpha}}_{rr,k}$ is defined as

$$\tilde{\boldsymbol{\alpha}}_{r,k} \equiv \frac{P_{r,k}}{P_r} \frac{P_{rr,k}}{P_{r,k}} = \frac{P_{r,k}}{P_r} [f_{r,k}(\{\tilde{\chi}_{jr,k}(\mathbf{x}_{r,k})\})]^{-1}.$$

Finally, we can use A2 to write the Local Labor Market Module in this general environment as the following two relationships:

$$\log \omega_c = \log \Phi_{r,k}(\mathbf{L}_c) + \log \nu_c \quad (19)$$

$$\log \omega_c - \log \tilde{\boldsymbol{\alpha}}_c = \log \Psi_c(\mathbf{L}_c) + \log \xi_c. \quad (20)$$

We can now state the main result of this section.

Theorem 1. *Suppose A1-A3 hold. Given (Φ, Ψ) , the demand-adjusted trade shares, $\tilde{\boldsymbol{\alpha}}_c$, are a sufficient statistic for the effect of international trade on local labor markets, ω_c and \mathbf{L}_c .*

Theorem 1 is a generalization of the sufficient statistic result in [Arkolakis et al. \(2012\)](#)

regarding the effects of trade on real wages and employment at the regional level. In particular, in a setting characterized by complex interactions between sectoral labor supply and agglomeration forces, we show that, under macro restrictions A1-A3, it is sufficient to observe the change in the vector of demand-adjusted trade shares to infer the change in real wages and employments across regions of a country. In fact, the strength of the labor supply and agglomeration interactions shape the quantitative predictions of the model. To see this, we use the implicit function theorem and equations (19)-(20) to explicitly write the change in local labor market outcomes as a function of the change in the vector of demand-adjusted trade shares. Under regularity conditions in the elasticity structure of $\Psi_c(\cdot)$ and $\Phi_c(\cdot)$, we can write

$$d \log \boldsymbol{\omega}_c = (I - \bar{\Psi}_c \bar{\Phi}_c)^{-1} d \log \tilde{\boldsymbol{x}}_c \quad (21)$$

$$d \log \mathbf{L}_c = \bar{\Phi}_c (I - \bar{\Psi}_c \bar{\Phi}_c)^{-1} d \log \tilde{\boldsymbol{x}}_c \quad (22)$$

where $\bar{\Psi}_c \equiv \left[\frac{\partial \log \Psi_{r,k}(\mathbf{L}_c)}{\partial \log L_{r',k'}} \right]_{rk,r'k'}$ and $\bar{\Phi}_c \equiv \left[\frac{\partial \log \tilde{\Phi}_{r,k}(\boldsymbol{\omega}_c)}{\partial \log \omega_{r',k'}} \right]_{rk,r'k'}$ are the elasticity matrices of the functions $\Psi_c(\cdot)$ and $\Phi_c(\cdot)$, respectively.¹⁵

These equations illustrate that our model is a simple generalization of constant elasticity and labor mobility models with agglomeration effects as in [Allen and Arkolakis \(2014\)](#), [Redding \(2012\)](#) or sectoral worker mobility as in [Caliendo et al. \(2015\)](#) and [Galle et al. \(2015\)](#). More generally, our model allows for rich substitution patterns across regions and sectors.

3.3 Trade Module

Similarly to the one sector model, we impose a labor market clearing condition for each sector and region. Thus, for all i ,

$$w_{r,k} L_{r,k} = \sum_j x_{rj,k} \cdot \left(\sum_k w_{j,k} L_{j,k} \right) \quad (23)$$

where x_{ji} is given by equations (12) and (17),

$$x_{rj,k} \equiv \chi_{rj,k} \left(\left\{ \tau_{ji,k} \cdot \frac{w_{j,k} \xi_{i,k} \Psi_{r,k}(L_{r,k})}{w_{r,k} \xi_{r,k} \Psi_{j,k}(L_{j,k})} \right\}_j \right) \quad (24)$$

¹⁵For details on the regularity conditions, see Lemma 2 in [Allen et al. \(2015\)](#).

and the price index is given by (13) and (17),

$$P_{r,k} \equiv f_k \left(\left\{ \tau_{jr,k} \cdot \frac{w_{j,k}}{\xi_{j,k} \Psi_{j,k}(\mathbf{L}_c)} \right\}_j \right). \quad (25)$$

The world equilibrium is characterized by a vector $\{w_{i,k}, L_{i,k}, \mathbf{x}_{i,k}, P_{i,k}, P_i\}$ that satisfies equations (19), (20), (23), (24), and (25).

To recap the main results of the general model, notice that compared to the parametric example of the previous version, the Labor Market Module directly connects trade and labor market outcomes at the regional-level albeit conditional on the general labor supply and agglomeration functions $\Phi_c(\cdot), \Psi_c(\cdot)$. Importantly, the exclusion of the trade costs from the Labor Market Module still retains, and this feature will allow us to use the Trade Module to design a novel IV estimation strategy in the next section.

4 Empirical Strategy

Our main theoretical result establishes a structural relationship between regional trade outcomes and regional labor market outcomes in the class of models satisfying Assumptions A1–A3. This relationship depends crucially on the elasticity structure of the regional labor supply function, Φ , and the regional agglomeration function, Ψ . In this section, we propose a novel strategy to estimate these central elasticities using the general equilibrium properties of our model. Throughout our analysis, we assume that the functions controlling the bilateral trade flows, $\chi_k(\cdot)$, and regional price indices, f and f_k , are known, since the estimation of such functions has been the focus of an extensive literature in international trade.¹⁶

4.1 Parametric Econometric Model

We assume that the world economy is generated by our model in several periods, indexed by t . Our methodology relies on changes in the world equilibrium between periods. We denote as y^t the value of variable y in period t and as $\hat{y}^t = y^t/y^{t_0}$ the change in variable y between a base period t_0 and period t . We assume that we observe changes in bilateral

¹⁶In single-sector gravity models, these functions only depend 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 et al. \(2011\)](#) 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 et al. \(2017\)](#) 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 $\chi_k(\cdot)$, $f_k(\cdot)$ and $f(\cdot)$.

trade flows, $\{\hat{\mathbf{x}}_{i,k}^t\}$, and labor market outcomes, $\{\hat{\omega}_{i,k}^t, \hat{L}_{i,k}^t, \hat{P}_{i,k}^t\}$, but the regional shocks, $\{\hat{\nu}_{i,k}^t, \hat{\xi}_{i,k}^t, \hat{\tau}_{i,k}^t\}$, are unobserved.

We start by imposing parametric restrictions on the functions governing labor supply and agglomeration forces across regions and sectors.

Condition 1. Assume that $\Psi_c(\cdot) = \Psi_c(\cdot|\boldsymbol{\psi})$ and $\Phi_c(\cdot) = \Phi_c(\cdot|\boldsymbol{\phi})$ are log-linear in the vector of unknown parameters $\Theta \equiv (\boldsymbol{\psi}, \boldsymbol{\phi}) \in \mathbb{R}^s$.

Condition 1 imposes that the functions Ψ and Φ are log-linear in a set of parameters $(\boldsymbol{\psi}, \boldsymbol{\phi})$. The log-linearity restriction significantly simplifies the conditions for identification and optimality of our methodology.¹⁷ Given the parameter vector Θ , we use the Local Labor Market and the Trade modules to show that changes in the observed endogenous variables depend on both the changes in the vector of unobserved shocks and the vector of endogenous variables in the base period.

Lemma 1. Define the vector of initial variables, $W_c^{t_0} \equiv (\mathbf{Y}_c^{t_0}, \mathbf{L}_c^{t_0}, \mathbf{x}_c^{t_0}, \mathbf{e}_c^{t_0})$ with $\mathbf{W}^{t_0} \equiv \{W_c^{t_0}\}_c$. Suppose Condition 1 holds in the world equilibrium of the model of Section 3.

The Local Labor Market Module implies that

$$F_{r,k}(\hat{\mathbf{x}}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | W_c^{t_0}, \Theta) = \hat{\epsilon}_{r,k}^t \quad \text{for all } (r, k) \quad (26)$$

where $F(\cdot | W_c^{t_0}, \Theta)$ is a known linear function in Θ , and

$$\hat{\epsilon}_{r,k}^t \equiv (\log \hat{\nu}_{r,k}^t, \log \hat{\xi}_{r,k}^t). \quad (27)$$

The Trade Module implies that

$$A_{r,k}\left(\left\{\hat{\mathbf{x}}_c^t, \hat{\omega}_c^t, \hat{\mathbf{L}}_c^t\right\}, \{\hat{\tau}_c, \hat{\epsilon}_c\} | \mathbf{W}^{t_0}, \Theta\right) = \mathbf{0} \quad \text{for all } (r, k). \quad (28)$$

where $A_{r,k}(\cdot)$ is a known non-linear function.

We prove Lemma (1) in Appendix 9.1. Together, equations (26)–(28) play a central role in the construction of our estimation methodology. Given the parameter vector Θ and the vector of initial conditions $W_c^{t_0}$, the function $F(\cdot | W_c^{t_0}, \Theta)$ relates changes in observed endogenous variables to the unobserved shocks at the regional-level. We exploit this relationship

¹⁷Our results can be extended to allow for general differentiable functions on the parameter vector. In this case however, the conditions for identification are less intuitive and not testable. This result is directly related to the discussion in [Newey and McFadden \(1994\)](#) about identification with moment conditions that are nonlinear in the unknown parameter vector.

to obtain moment conditions for the estimation of Θ . In proposing such moment conditions, we must take into account the fact that, in general equilibrium, regional trade and labor market outcomes are correlated with the unobserved local shocks included in $\hat{\epsilon}_{r,k}^t$. We use the general equilibrium structure of our model to circumvent this challenge. Specifically, we construct model-implied moment conditions using observable shifters of bilateral trade costs, since these are excluded from the Local Labor Market Module in (26), but simultaneously affect the world equilibrium through the Trade Module in (28).

4.2 Construction of Model-implied IV

We now describe the three steps of our methodology.

Step 1: Observable Trade Cost Shifter. The first step of our methodology is to construct a moment condition using the unobserved shocks in the Local Labor Market Module. To this end, we assume that a shifter of bilateral trade cost shocks, $\hat{\mathbf{z}}^t \equiv \{\hat{z}_{ij,k}^t\}$, is observable. We impose a log-linear relationship between the iceberg trade cost $\log \hat{\tau}_{ij,k}^t$ on the observable shifter $\log \hat{z}_{ij,k}^t$:

$$\log \tau_{ij,k}^t = \rho \log \hat{z}_{ij,k}^t + \hat{\eta}_{ij,k}^t \quad \text{where} \quad E[\hat{\eta}_{ij,k}^t | \hat{z}_{ij,k}^t] = 0. \quad (29)$$

The importance of equation (29) lies on the fact that the observable variable $\hat{\mathbf{z}}^t$ affects bilateral trade costs $\boldsymbol{\tau}^t$, which implies that $\hat{\mathbf{z}}^t$ also affects regional markets through the Trade Module in equation (28). However, due to the structure of our model, $\hat{\mathbf{z}}^t$ is excluded from the Local Labor Market Module in equation (26). Thus, we can use $\hat{\mathbf{z}}^t$ to construct moment conditions for the estimation of Θ as long as the following exogeneity condition holds.

Condition 2. Assume that $E[\hat{\epsilon}_{r,k}^t | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0}] = 0$ for every sector k and region r , where $\hat{\epsilon}_{r,k}^t$ is defined in equation (27).

Condition 2 states that, conditional on the vector of endogenous variables in the initial world equilibrium, the trade cost shifter $\hat{\mathbf{z}}$ is mean-independent from local shocks to productivity and labor supply. By the law of iterated expectations, this assumption immediately implies that, for any function $H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$,

$$E[\hat{\epsilon}_{r,k}^t \cdot H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})] = 0. \quad (30)$$

It is important to note that the exogeneity condition that we require for identification is similar to the exogeneity restrictions that empirical papers in international trade typically make (see Topalova (2010), Kovak (2013), Ebenstein et al. (2013) and Pierce and Schott (2016)). This exogeneity restriction is potentially weaker than the one made in other em-

pirical works, such as [Autor et al. \(2013\)](#), which require the entire vector of imports to be mean independent from unobserved local shocks.

Step 2: Recover Unobserved Shocks. The use of the moment condition (30) for the estimation of Θ requires a measure of the unobserved local shocks in $\hat{\epsilon}_{r,k}^t$. In our model, Lemma 1 implies that the Local Labor Market Module in (26) immediately delivers such a measure. Thus, for any function $H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$,

$$E [e_{r,k}^t(\Theta) \cdot H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})] = 0. \quad (31)$$

where

$$e_{r,k}^t(\Theta) \equiv F_{r,k} \left(\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | \mathbf{W}^{t_0}, \Theta \right).$$

As discussed earlier, assumptions A1-A2 immediately imply invertibility. The invertibility of the structural residuals with specific assumptions is obtained also in other frameworks, such as [Berry et al. \(1995\)](#), [Berry et al. \(2013\)](#), [Bartelme \(2015\)](#), [Allen et al. \(2014\)](#), and [Monte et al. \(2015\)](#).

Step 3: Model-implied IV. The final step of our methodology is to propose a function $H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$ capturing the exposure of sector-region pairs to the observable variable $\hat{\mathbf{z}}^t$ conditional the vector of initial world equilibrium. In order to capture all the channels through which $\hat{\mathbf{z}}^t$ affects regional economies, we consider the change in endogenous regional outcomes predicted by our general equilibrium model following the shock $\hat{\mathbf{z}}^t$. Formally, we propose the following function.

Definition 1. *The predicted change in the endogenous variables implied by $\hat{\mathbf{z}}^t$ is $Z_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta) \equiv \left\{ \hat{x}_{r,k}^p, \hat{\omega}_{r,k}^p, \hat{\mathbf{L}}_c^p \right\}$ such that*

$$F_{r,k} \left(Z_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta) | \mathbf{W}^{t_0}, \Theta \right) = \mathbf{1}$$

$$A_{r,k} \left(\left\{ Z_{r',k'}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta) \right\}_{r',k'}, \left\{ \hat{\mathbf{z}}^t, \mathbf{1} \right\} | \mathbf{W}^{t_0}, \Theta \right) = \mathbf{0}.$$

Conditional on the parameter vector Θ and the functions $A_{r,k}(\cdot)$ and $F_{r,k}(\cdot)$, the function $Z_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta)$ is the solution of a non-linear system of equations that depends only on initial trade equilibrium \mathbf{W}^{t_0} and on the observable trade cost shifter $\hat{\mathbf{z}}^t$. To ease the notation, we will henceforth denote the function $Z_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta)$ simply by $Z_{r,k}$. This function is directly related to the "exact hat-algebra" in [Dekle et al. \(2007\)](#) that yields counterfactual changes in endogenous variables using information about the initial equilibrium and some observed changes in trade costs. Here, we use this idea to construct instrumental variables for the estimation of the unknown parameters of our model. Specifically, we apply $Z_{r,k}$ to expression (31) to obtain the following moment condition:

$$E [e_{r,k}^t(\Theta) \cdot \nabla_{\Theta} F_{r,k}(Z_{r,k} | \mathbf{W}^{t_0})] = 0. \quad (32)$$

where $\nabla_{\Theta} F_{r,k}$ does not depend on Θ due to the linearity of $F_{r,k}$ in Lemma (1).

In our general equilibrium model, the function $Z_{r,k}$ corresponds to the total effect of $\hat{\mathbf{z}}^t$ on the endogenous regional outcomes. It includes all the general equilibrium channels through which $\hat{\mathbf{z}}^t$ affects regional labor markets. Although there are other sources of structural shocks that affect regional economies, these are excluded from $Z_{r,k}$ by construction.

In Appendix 9.2, we show that our general equilibrium model implies that the changes in endogenous variables have two components: one stemming from $\hat{\mathbf{z}}^t$ and another stemming from all other shocks. Under Condition 2, these two components are orthogonal. The following lemma presents this result.

Lemma 2. *Suppose Conditions 1 and 2 hold. The first-order approximation of the general equilibrium equations in (26)–(28) implies that*

$$\begin{bmatrix} \hat{x}_{r,k}^t & \hat{\omega}_{r,k}^t & \hat{\mathbf{L}}_c^t \end{bmatrix} = \rho Z_{r,k} + \hat{\mu}_{r,k}^t \quad (33)$$

where

$$E [\hat{\mu}_{r,k}^t | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0}] = 0.$$

This result plays a central role in shaping the properties of the estimator based on the moment condition (32). Intuitively, equation (33) is the first-stage of our empirical methodology: the instrumental variable, $Z_{r,k}$, induces exogenous variation in the endogenous variables, $\{\hat{x}_{r,k}^p, \hat{\omega}_{r,k}^p, \hat{\mathbf{L}}_c^p\}$, that can be used in the estimation of the structural parameters. Moreover, through the lens of our general equilibrium model, the instrumental variable and the endogenous variable have a log-linear relationship. As discussed below, the log-linearity is key for the optimality of the moment condition (32).

4.3 Model-implied Optimal IV

We now define the general class of GMM estimators associated with the moment condition in (32) implied by a given exposure function, $H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$ and discuss the optimality of our proposed model-implied instrumental variables approach for that class of GMM estimators.

Definition 2. *For a generic function $H_{r,k} \equiv H_{r,k}(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$, the GMM estimator associated with moment condition (31) minimizes the following quadratic function:*

$$\hat{\Theta}_H \equiv \min_{\Theta} \left[\sum_{r,k,t} e_{r,k}^t(\tilde{\Theta}) \cdot H_{r,k} \right]' W \left[\sum_{r,k,t} e_{r,k}^t(\tilde{\Theta}) \cdot H_{r,k} \right]$$

where W is the optimal GMM weighting matrix.

We now consider the asymptotic properties of the class of GMM estimators in Definition (2). We show in Appendix 9.3 that, under standard regularity, this class of estimator is consistent, $\hat{\Theta}_H \xrightarrow{p} \Theta$, and asymptotically normal with variance given by

$$\text{Var}(\hat{\Theta}_H) = (E[H_{r,k}G_{r,k}])^{-1} (E[H_{r,k}\Omega_{r,k}H_{r,k}']) (E[H_{r,k}G_{r,k}])^{-1'} \quad (34)$$

where

$$G_{r,k} \equiv E[\nabla_{\Theta} e_{r,k}(\Theta) | \hat{\mathbf{z}}^t, W^{t_0}] \quad \text{and} \quad \Omega_{r,k} \equiv E[e_{r,k}(\Theta)e_{r,k}(\Theta)' | \hat{\mathbf{z}}^t, W^{t_0}].$$

Within the class of estimators in Definition (2), we propose the use of the estimator constructed from the moment condition in (32). Although all the estimators in Definition (2) are consistent, they vary in terms of asymptotic variance — that is, the estimators differ in terms of precision. Thus, in choosing the exposure function $H_{r,k}$, we follow the approach in Chamberlain (1987) and select the one minimizing the asymptotic variance, $\text{Var}(\hat{\Theta}_H)$. Applying the result in Chamberlain (1987) to our environment, we show in Appendix 9.4 that such an estimator is given by

$$H_{r,k}^* \equiv E[\nabla_{\Theta} e_{r,k}^t(\Theta) | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0}] (\Omega_{r,k})^{-1}, \quad (35)$$

which implies that

$$\text{Var}(\hat{\Theta}_{H^*}) = E[G_{r,k}' \Omega_{r,k}^{-1} G_{r,k}] \quad (36)$$

We then combine the linearity of $F_{r,k}(\cdot | \Theta)$ on Θ and expression in (33) to establish that

$$E[\nabla_{\Theta} e_{r,k}^t(\Theta) | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0}] = E\left[\nabla_{\Theta} F_{r,k}\left(\hat{\mathbf{x}}_{r,k}^t, \hat{\boldsymbol{\omega}}_{r,k}^t, \hat{\mathbf{L}}_c^t | W_c^{t_0}\right) | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0}\right] = \rho \nabla_{\Theta} F_{r,k}(Z_{r,k} | W_c^{t_0}).$$

The estimator based on $H_{r,k}^*$ is the most efficient in the class of estimators in Definition (2). That is, the asymptotic variance of this estimator, $\text{Var}(\hat{\Theta}_{H^*})$, is always weakly lower than the asymptotic variance implied by any arbitrary exposure function, $\text{Var}(\hat{\Theta}_H)$. This result is presented in the following theorem.

Theorem 2. *Suppose the world economy is generated by the model in Section 3, satisfying Conditions 1 and 2. The function $H_{r,k}^*(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0}|\Theta)$ in (35) minimizes the asymptotic variance in equation (34).*

Theorem (35) suggests that the optimal IV, in the sense that it minimizes the asymptotic variance of the GMM estimator, is $Z_{r,k}$, the vector of changes in endogenous variables predicted by the general equilibrium model once we shock the system with an exogenous \hat{z} . To implement our optimal GMM estimator, we use the sample analog of moment condition in (32) to define the Model-implied Optimal IV (MOIV).

Definition 3. *The Model-implied Optimal IV estimator, $\hat{\Theta}$, is constructed in two stages.*

Stage 1. *For an initial guess for the parameter vector Θ_0 ,*

$$\hat{\Theta}_1 = \arg \min_{\tilde{\Theta}} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \Theta_0) \right) W \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \Theta_0) \right)' \quad (37)$$

where W is a symmetric positive definite matrix of moment weights.

Stage 2. *Using the first-stage estimates $\hat{\Theta}_1$,*

$$\hat{\Theta} = \arg \min_{\tilde{\Theta}} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \hat{\Theta}_1) \right) (\hat{W})^{-1} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \hat{\Theta}_1) \right)', \quad (38)$$

where $\hat{W} \equiv \sum_{r,k,t} e_{r,k}^t(\hat{\Theta}_1) e_{r,k}^t(\hat{\Theta}_1)'$.

The implementation of MOIV entails two stages. In the first stage, we use a guess of the structural parameters to compute the model-implied IV, $H_{r,k}^*$. Since this variable is a function of $(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0})$, it implies a valid moment condition for the consistent estimation of Θ . Once we obtain this first-stage estimator, we can compute the model-implied instrument again to emulate the main moment condition in (32). In Appendix 9.3, we show that, under standard regularity conditions, this MOIV consistently recovers the *unique* Θ satisfying the moment condition in (32).

5 Empirical Application

To estimate the structural parameters and implement our GMM methodology, we impose parametric restrictions that are consistent with the assumptions in Section 3. As shown in Appendix 8.3, we assume that the agents have a separable utility function in consumption goods and labor. The consumption aggregator is Cobb-Douglas function of sector-level CES

consumption indices of goods produced in different regions. To derive the aggregate labor supply functions, we assume that individuals have heterogeneous preferences for sectors and regions, which are drawn from the Generalized Extreme Value distribution proposed in [McFadden \(1980\)](#),

$$\{a_{r,k}(l)\}_{r,k} \sim \exp \left[- \left(\sum_r a_r^t \cdot \left(\sum_{k=0}^K \nu_{r,k}^t \cdot (a_{r,k})^{\phi_e} \right)^{\frac{\phi_m}{\phi_e}} \right)^{\frac{1}{\phi_m}} \right], \quad (39)$$

where $1 < \phi_m \leq \phi_e$. We allow the agents to choose whether to join the workforce or to work in the home sector, which is indexed by $k = 0$.¹⁸

With this functional form, the parameter ϕ_m represents the elasticity of migration, which regulates the likelihood of choosing to work in a region after an increase in that region's real wage. Instead, the parameter ϕ_e regulates the elasticity of sectoral employment, i.e. the probability of choosing to work in a sector after an increase in the sectoral real wage. Note that the formulation in equation (39) generalizes the distribution assumed in [Caliendo et al. \(2015\)](#), which impose the migration and employment elasticities to be the same. Also notice that the elasticity of employment ϕ_e corresponds to the parameter controlling the between-sector mobility of workers in [Burstein et al. \(2016\)](#) and [Galle et al. \(2015\)](#).

Finally, we assume that agglomeration forces operate at the state-sector level. Specifically, we impose that

$$P_{ij,k}^t = \tau_{ij,k}^t \cdot w_{i,k}^t (L_{i,k}^t)^{-\psi} \cdot \xi_{i,k}^t. \quad (40)$$

Equations for the estimation. Under the parametric assumptions above, we show in the Appendix 8.3 that our model yields a set of simple equations relating trade and labor market outcomes across regions and sectors. We use these equations to estimate the parameter controlling the local agglomeration forces, ψ , and the parameters controlling local labor supply responses, (ϕ_e, ϕ_m) .

To estimate the employment elasticity ϕ_e , we regress the following specification:

$$\Delta \log (l_{r,k}^t / l_{r,0}^t) = \phi_e \Delta \log \omega_{r,k}^t + \nu_{r,k}^t, \quad (41)$$

where $l_{r,k}^t / l_{r,0}^t$ is the employment share in sector k relative to unemployment in region r . To estimate the migration elasticity ϕ_m , we run

$$\Delta \log (n_r^t) = - \left(\frac{\phi_m}{\phi_e} \right) \Delta \log (l_{r,0}^t) + d_t + a_r^t, \quad (42)$$

¹⁸For simplicity, we normalize the real wage for non-employed workers to one. Also, it is straightforward to add the hours margin as well. We implement this extended specification as robustness test, and we do not find significant responses of hours worked in the data.

where n_r^t is the share of working age population in region r . Finally, to obtain the agglomeration elasticity we regress

$$\Delta \log \omega_r^{k,t} - \Delta \log \tilde{x}_{r,k}^t = \psi \Delta \log L_{r,k}^t + \Delta \xi_{r,k}^t, \quad (43)$$

where the demand-adjusted trade share is given by $\hat{x}_{r,k}^t = (\hat{x}_{rr,k}^t)^{-\frac{1}{\varepsilon}} \frac{\hat{P}_{r,k}^t}{\hat{P}_r^t}$. To recap, in order to estimate the structural parameters we need to observe the trade elasticity, ε , as well as the empirical counterparts of trade shares, real wages, employment across sectors and regions. We describe the sources of these data in the following section.

5.1 Data

In this section we describe in detail the datasets we use and how we construct the relevant variables. We use data for the years 1997, 2002, 2007 and 2012. We can divide our data sources into four groups.

Labor Market Outcomes. To construct measures for the nominal wage and labor supply for each US state, we exploit the richness of the Current Population Surveys - Merged Outgoing Rotation Groups dataset. The CPS is the monthly household survey conducted by the Bureau of Labor Statistics to measure labor force participation and employment, in which about 50,000 households per month are queried. We use the reported weekly earnings and hours of these households to construct the average nominal wage and the average number of hours worked for each state, by weighting the individual values with the official weights reported by the CPS.¹⁹

Price Data. We construct a state-level measure of the price index P_i using data from the Cost of Living Index, published by the Council for Community and Economic Research (C2ER). This index is based on a survey that is conducted every quarter and that records the prices for each urban area within the US.²⁰ It is a well-known source of data on living cost differentials among U.S. cities (see [Moretti \(2013\)](#)). Since the price index is at the urban area level, we aggregate prices at the state level by taking a population-weighted average of the prices.

As a robustness, we also construct state-level price indices using the Nielsen Homescan Dataset, for the years 2004, 2007 and 2012. This dataset provides detailed information on purchases, trips of purchases, household and product characteristics. We follow closely the

¹⁹We follow the cleaning procedure of [Autor et al. \(2008\)](#) to adjust for top censoring, outliers, and time consistency of variables. We sample all individuals aged between 16 and 64 years.

²⁰The price index is available also for 5 categories of goods (Grocery, Housing, Utilities, Transportation, Health Care).

procedure proposed by [Handbury and Weinstein \(2014\)](#) to construct price indices that take into account for product, buyer and retailer heterogeneity (see Appendix 10 for details).

Trade data. We use data on bilateral trade flows for 35 countries from the WIOD.²¹ We consider in the analysis all the 37 ISIC sectors included in the WIOD, and we aggregate them to two sectors, Manufacturing and Non-Manufacturing. We merge this matrix of international trade flows with a matrix of bilateral flows between the 50 US states.²² To construct this *regional* matrix of trade we use between-states shipments data from the Commodity Flow Survey. Finally, we use Census data at the port of entry/exit level to obtain trade flows between each of the US states and each of the 35 countries in the WIOD. Merging these three matrices of trade flows leads to a matrix of bilateral trade flows between the 50 US states and 35 countries. In Appendix 10 we explain in detail the construction of this matrix. We also take into account for trade imbalances, which we assume to be exogenous.

Gravity Calibration. To implement the estimation strategy of Section 3, it is necessary to calibrate the parameters controlling the gravity structure of our model. The CES assumption implies a constant trade elasticity ε , and we follow [Adao et al. \(2017\)](#) by setting $\varepsilon = 6$.

5.2 Observed Trade Costs

A key step to construct the moment conditions in (30) is to obtain observable shifters of international trade costs that satisfy the exogeneity restriction in Condition 2.

Our baseline results use the gravity structure of our model to compute changes in bilateral trade cost shocks. The CES assumption implies that we can write the gravity equation as:

$$(\hat{\tau}_{ji,k} \hat{\tau}_{ij,k})^{-\varepsilon} = \frac{\hat{x}_{ji,k}}{\hat{x}_{jj,k}} \cdot \frac{\hat{x}_{ij,k}}{\hat{x}_{ii,k}}. \quad (44)$$

To recover the change in bilateral trade costs, we follow the same approach as [Head and Ries \(2001\)](#) and impose symmetry in trade cost shocks, $\hat{\tau}_{ji} = \hat{\tau}_{ij}$. Thus, for every bilateral pair of world regions, our baseline measure of the change in trade cost is

$$\hat{z}_{ji}^t = \left(\frac{\hat{x}_{ji,k}}{\hat{x}_{jj,k}} \cdot \frac{\hat{x}_{ij,k}}{\hat{x}_{ii,k}} \right)^{-\frac{1}{2\varepsilon}}. \quad (45)$$

Conditional on the measure of trade costs, we compute the predicted changes in endogenous variables, $Z_{r,k} \equiv Z_{r,k}(\hat{z}^t, \mathbf{W}^{t_0} | \Theta)$, shown in definition 1. These predicted changes is then

²¹See [Dietzenbacher et al. \(2013\)](#) for details about the WIOD database. See Appendix for a full list of the countries in the sample.

²²We aggregate DC with Maryland.

used to construct the moment conditions in (32) and to implement our Two-Stages GMM methodology.

5.3 Montecarlo Analysis

In Section 4 we have shown that the MOIV delivers the most efficient estimator, and thus the lowest variance. In order to assess the gains from our methodology, we perform a simple Montecarlo analysis.

Starting from the observed initial conditions in 2012, we simulate 300 economies by taking i.i.d. draws of the unobserved local shocks $(\hat{\xi}_{i,k}, \hat{\nu}_{i,k})$ from a log-normal distribution, as well as draws of the observable component of trade costs, $\hat{z}_{ij,k}$, and of the unobservable component, $\hat{\eta}_{ij,k}$, such that $\ln \hat{\tau}_{ij,k} = \ln \hat{z}_{ij,k} + \ln \hat{\eta}_{ij,k}$.²³ Then we feed these shocks into the model and create “fake” data on changes in endogenous outcomes for US states in several periods. Using the fake data, we construct different IVs and estimate the structural parameters for each economy. Finally, for each IV and parameter, we compute the average and the standard deviation of the estimates.

Using this data, we estimate the structural parameters using the observable shocks to the bilateral trade cost of US states and China, $\hat{z}_{Ci,k}$ and $\hat{z}_{iC,k}$. We consider three different instruments. As the first instrument, we use directly the observable cost shifter of trading with China, $\hat{z}_{Ci,k}$ and $\hat{z}_{iC,k}$. The second instrument is the observable shifter interacted with initial sectoral employment shares in the state, $l_{r,k}^{t_0} \cdot \hat{z}_{Ci,k}$ and $l_{r,k}^{t_0} \cdot \hat{z}_{iC,k}$. Lastly, we use the same observable trade cost shocks to implement the MOIV as described in Section 4. For all instruments, we estimate equations (41)-(43). The results are summarized in the Table 1.

Note that, since the exogeneity condition holds by construction (because trade costs are orthogonal to the unobserved local shocks), all IVs are consistent, and the average estimates are close to the true values in all cases. In addition, it is evident how the MOIV delivers estimates with a much smaller standard deviation compared to the other instruments. In the simulations, the standard deviation of the estimates obtained with our methodology is 2%–40% of that obtained with the alternative instruments.

5.4 Results

We finally implement the empirical strategy outlined in Section 4 to estimate the structural parameters in equations (41)–(43). To this end, we compute the predicted change in the endogenous variables, $H_{r,k}^*(\hat{z}^t, \mathbf{W}^{t_0} | \Theta)$, using \hat{z}^t obtained from equation (45). In the estimation, we only use the cost of bilaterally trading with China in the periods of 1997–2002,

²³We set the average reduction of trade costs from and to China at -15%.

Table 1: Montecarlo Simulations

IV 1: \hat{z}_{Ci} and \hat{z}_{iC}			
Parameter	True value	Mean	St. Dev.
ψ	0.5	0.509	0.085
ϕ_e	1.4	1.643	1.382

IV 2: \hat{z}_{Ci} and \hat{z}_{iC} interacted with $l_{r,k}^{to}$			
Parameter	True value	Mean	St. Dev.
ψ	0.5	0.511	0.211
ϕ_e	1.4	1.030	2.119

IV 3: MOIV using \hat{z}_{Ci} and \hat{z}_{iC}			
Parameter	True value	Mean	St. Dev.
ψ	0.5	0.500	0.005
ϕ_e	1.4	1.498	0.053

2002–2007, and 2007–2012. We also include in all specifications census division, sector and period dummies, and initial labor market conditions interacted with period dummies. These controls capture state-specific time trends in labor market outcomes associated with the pre-period sector employment composition, share of college graduates in labor force, and female labor force participation.²⁴

5.4.1 Structural Parameters

Table 2 shows the estimates of the structural parameters, (ψ, ϕ_e, ϕ_m) , along with the standard errors clustered at the state-level. Our point estimate of the agglomeration elasticity is $\psi = 0.185$, which multiplied by the trade elasticity roughly corresponds to the unitary agglomeration elasticity in [Krugman \(1980\)](#). The estimates of agglomeration forces in the literature present large variation. [Ahlfeldt et al. \(2015\)](#), using detailed data for the city of Berlin, find substantial and statistically significant agglomeration forces, with an estimated elasticity of productivity with respect to the surrounding concentration of workplace employment of 0.08. [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.

The other two columns report our estimates of the structural parameters governing regional labor supply. Our point estimate for the extensive margin elasticity is $\phi_e = 1.5$, implying that a 1% increase in the sector relative wage triggers a 1.5% increase in the sector

²⁴This set of controls is similar to that used [Autor et al. \(2013\)](#).

Table 2: Estimation of Structural Parameters

	Agglomeration	Labor Supply Elasticity	
	Elasticity	Between sectors	Between regions
	ψ	ϕ_e	$-\phi_m/\phi_e$
	0.185*	1.562**	-1.27
	(0.100)	(0.766)	(1.539)
F-stat	11.29	7.268	2.661

Note. Sample of state-sector pairs (2 sectors x 50 states x 3 time periods 1997-2002, 2002-2007, 2007-2012). All regressions include sector and census division dummies along with period dummies interacted with the following initial characteristics: share of college graduates in labor force, female labor force participation, 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

relative employment share. In terms of employment shares, if real wages increase in sector k by 1%, the employment share in sector k increases by $\phi_e l_{r,k}(1 - l_{r,k})$. Using the national employment composition, this implies an extensive margin elasticity of 0.17% in manufacturing and 0.36% non-manufacturing. Our estimates are close to the estimates of the extensive margin elasticity in the literature reviewed by [Chetty \(2012\)](#). Finally, we estimate a non-significant migration elasticity. For this parameter, our model has a very weak predictive power, leading to very imprecise estimates. 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 \(2016\)](#)).

5.4.2 Model Fit

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. In this section, we use the structural estimates reported in Table 2.

We consider the effect of changes in the cost of trading with China. Figure 5.4.2 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 strong relationship allows the estimation of the structural parameters in Table 2.

Notice however that the magnitude of this slope is large: i.e. the model predicts correctly the direction of the relationship but under-predicts the size of the impact. As shown in Panel A of Table 3, the slope varies between 5 for the response in trade shares and 25 for the response in employment. This could be indicative that the trade shocks have a much higher labor market impact than that predicted by our model or that the China shock is correlated

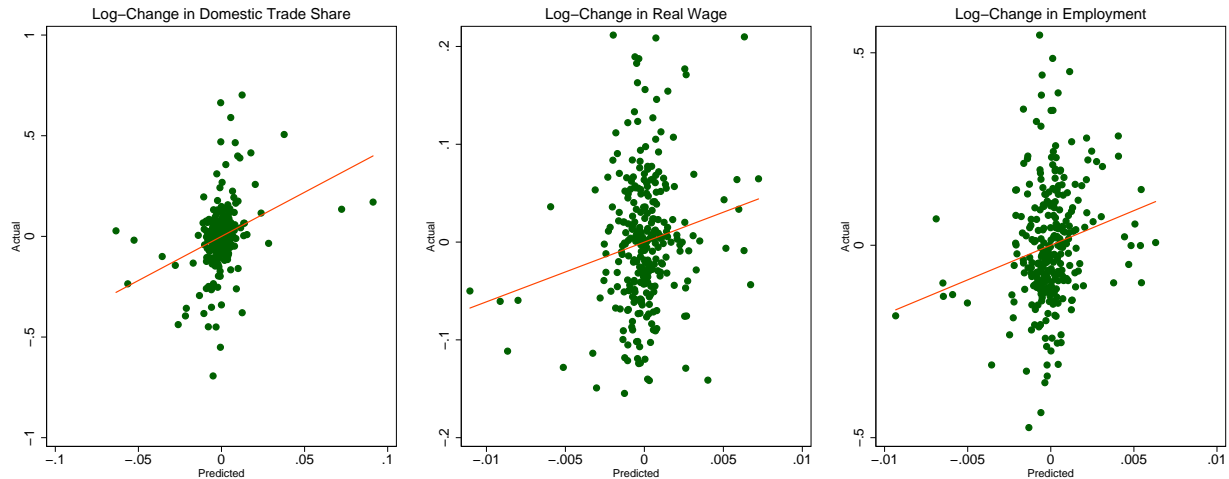


Figure 1: Model Fit: China Shock

Table 3: Model Fit

	Log-change in		
	Domestic Trade Share	Real Wage	Total Employment
	(1)	(2)	(3)
Panel A: Chinese shock			
Predicted change	4.52*** (1.447)	6.12*** (2.271)	24.67*** (7.343)
R2	0.24	0.55	0.60
Panel B: All shocks			
Predicted change	1.27*** (0.093)	0.58*** (0.225)	0.87* (0.456)
R2	0.74	0.55	0.63

Note. Sample of state-sector pairs (2 sectors x 50 states x 3 time periods 1997-2002, 2002-2007, 2007-2012). All regressions include sector and census division dummies along with period dummies interacted with the following initial characteristics: share of college graduates in labor force, female labor force participation, 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

with other shocks. To investigate the first hypothesis, we compare the correlation of actual changes in the endogenous variables with the predicted changes computed with all the trade cost shocks affecting the world economy. Panel B shows that the slope with all shocks is very close to one, which suggests that the model yields changes in labor market outcomes whose magnitude are consistent with those in the data.

6 Counterfactual analysis

We consider the counterfactual of reducing the trade costs from China to US states by the amount estimated in the data between 1997-2007. We keep all the other trade costs and shocks fixed and consider the calibrated economy in 1997. In this exercise, we use the estimated parameters reported in Table 2. Figures 2 and 3 illustrate the impact on real

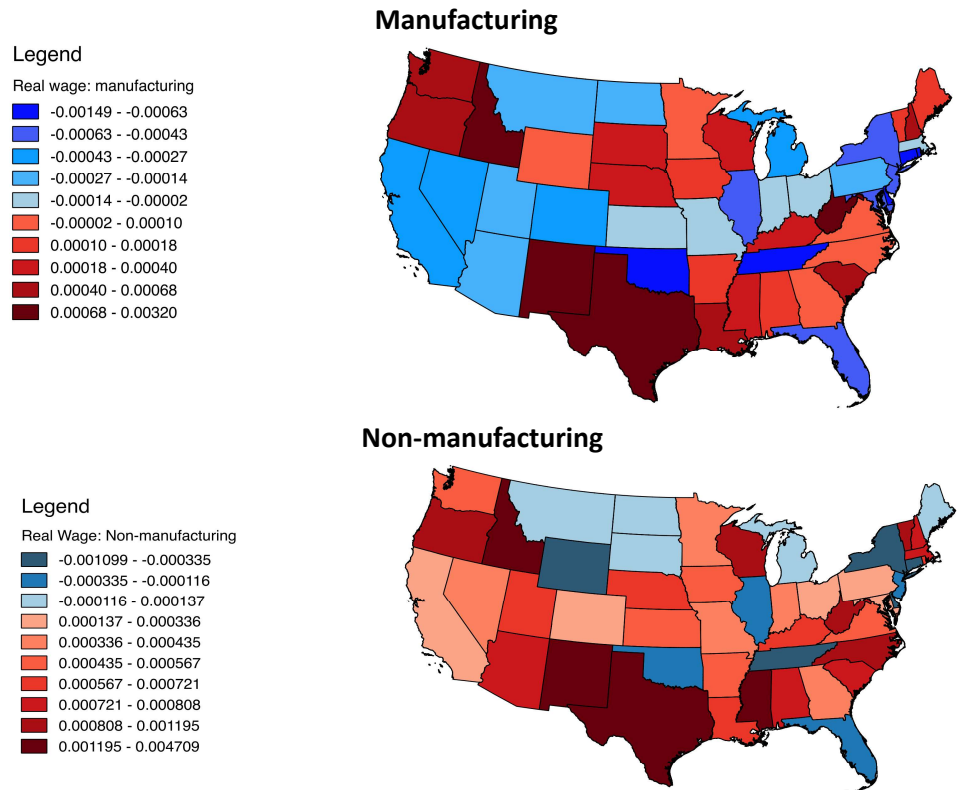
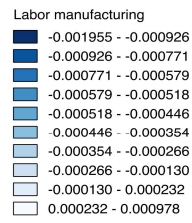
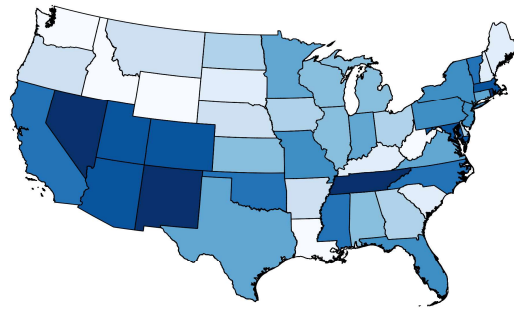


Figure 2: Counterfactual Predictions: China Shock on Real Wage

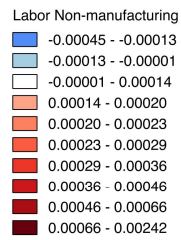
Legend



Manufacturing



Legend



Non-manufacturing

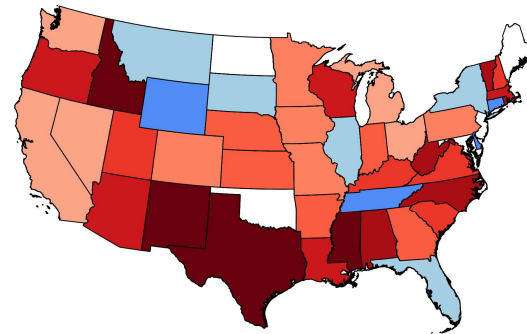


Figure 3: Counterfactual Predictions: China Shock on Employment

wages and employment, respectively. The effects are very heterogeneous owing to the different exposure of the regions to China, directly or indirectly, and to their different degrees of specialization on manufacturing and non-manufacturing. Interestingly workers in manufacturing lose both in terms of real wages and employment, as a result of the exposure to China. The resulting decrease in the price of tradables that leads to increase in the real wage of manufacturing workers does not compensate for the loss in competitiveness of the manufacturing sector in most of the states. The drop in employment is strong across states. It worth notice the patterns of spatial correlation that result from the strong special links of nearby states.

As expected, non-manufacturing workers generally gain and non-manufacturing employment increases for most of the states. But the drop of the real wages and employment in the manufacturing has spillover effects on the non-manufacturing sectors because of demand spillovers. As a result some states experience drop in non-manufacturing real wages and employment.

Finally, notice that the magnitudes of these changes are small. This indicates that the model predicts small aggregate effects of trade shocks. Even when we consider the projection between actual and predicted changes in the endogenous variables reported in Table 3, the aggregate impacts are moderate: in response to the shock manufacturing employment falls by 1.2% and non-manufacturing employment increases by .5%. In general, the model predicts small changes in the overall employment level in the US.

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 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

- 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., 2013. 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., 2016. The welfare effects of transportation infrastructure improvements.
- Allen, T., Arkolakis, C., Li, X., 2015. On the existence and uniqueness of trade equilibria. Tech. rep., mimeo, Northwestern and Yale Universities.
- Allen, T., Arkolakis, C., Takahashi, Y., 2014. Universal gravity. Tech. rep., Mimeo Northwestern and Yale University.
- Alvarez, F., Lucas, R. E., 2007. General equilibrium analysis of the Eaton-Kortum model of international trade. *Journal of Monetary Economics* 54 (6), 1726–1768.
- Arkolakis, C., 2010. Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy* 118 (6), 1151–1199.
- Arkolakis, C., Costinot, A., Rodríguez-Clare, A., 2010. Gains from trade under monopolistic competition: A simple example with translog expenditure functions and Pareto distributions of firm-level productivity. Manuscript, MIT, Penn State, and Yale University.
- Arkolakis, C., Costinot, A., Rodríguez-Clare, A., 2012. New trade models, same old gains? *American Economic Review* 102 (1), 94–130.

- Arrow, K. J. K. J., Hahn, F. H., 1971. General competitive analysis. Tech. rep.
- 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.
- Bartelme, D., 2015. Trade costs and economic geography: evidence from the us. Work. Pap., Univ. Calif., Berkeley.
- Bartelme, D., Costinot, A., Donaldson, D., Rodriguez-Clare, A., 2017. Sector-level economies of scale: Estimation using trade data.
- Berry, S., Gandhi, A., Haile, P., 2013. Connected substitutes and invertibility of demand. *Econometrica* 81 (5), 2087–2111.
- Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. *Econometrica* 63 (4), 841–890.
- Burstein, A., Morales, E., Vogel, J., 2015. Accounting for changes in between-group inequality. 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.
- Caliendo, L., Dvorkin, M., Parro, F., 2015. The impact of trade on labor market dynamics. 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.
- Chamberlain, G., 1987. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics* 34 (3), 305–334.
- 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.

- 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.
- Dekle, R., Eaton, J., Kortum, S., 2007. Unbalanced trade. Tech. rep., National Bureau of Economic Research.
- Dekle, R., Eaton, J., Kortum, S., 2008. Global rebalancing with gravity: Measuring the burden of adjustment. *IMF Staff Papers* 55 (3), 511–540.
- Dietzenbacher, E., Los, B., Stehrer, R., Timmer, M., de Vries, G., 2013. The construction of world input–output tables in the wiod project. *Economic Systems Research* 25 (1), 71–98.
- Dix Carneiro, R., Kovak, B. K., 2016. Trade liberalization and regional dynamics.
- Donaldson, D., Hornbeck, R., 2016. Railroads and american economic growth: A market access approach. *The Quarterly Journal of Economics*, qjw002.
- Dvorkin, M., et al., 2013. Sectoral shocks, reallocation and unemployment in a model of competitive labor markets. In: 2013 Meeting Papers. No. 1229. Society for Economic Dynamics.
- Eaton, J., Kortum, S., 2002. Technology, geography and trade. *Econometrica* 70 (5), 1741–1779.
- Ebenstein, A., Harrison, A., McMillan, M., Phillips, S., 2013. Why are american workers getting poorer? estimating the impact of trade and offshoring using the cps.
- Ethier, W. J., 1982. Decreasing costs in international trade and frank graham's argument for protection. *Econometrica: Journal of the Econometric Society*, 1243–1268.
- Faber, B., Gaubert, C., 2016. Tourism and economic development: Evidence from mexico's coastline. Tech. rep., National Bureau of Economic Research.
- Fujita, M., Krugman, P., Venables, A. J., 1999. *The Spatial Economy: Cities, Regions, and International Trade*. MIT Press, Boston, Massachusetts.
- Galle, S., Rodriguez-Clare, A., Yi, M., 2015. Slicing the pie: Quantifying the aggregate and distributional effects of trade. Unpublished manuscript, Univ. Calif., Berkeley.
- Goldberg, P. K., 2015. Introduction. In: *Trade and Inequality*.
- 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.
- Keane, M., Rogerson, R., 2012. Micro and macro labor supply elasticities: A reassessment of conventional wisdom. *Journal of Economic Literature* 50 (2), 464–476.
- 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.
- 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.
- Lee, E., 2015. Trade, inequality, and the endogenous sorting of heterogeneous workers. Tech. rep., mimeo.
- McFadden, D., 1980. Econometric models for probabilistic choice among products. *Journal of Business*, S13–S29.
- Melitz, M. J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71 (6), 1695–1725.
- 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.
- Newey, W. K., McFadden, D., 1994. Large sample estimation and hypothesis testing. *Handbook of econometrics* 4, 2111–2245.

- 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.
- Roy, A. D., 1951. Some thoughts on the distribution of earnings. *Oxford economic papers* 3 (2), 135–146.
- Scarf, H. E., Wilson, C. A., Jin, S., 2003. Uniqueness of equilibrium in the multi-country ricardo model. COWLES FOUNDATION DISCUSSION PAPER.
- 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 Micro-foundations of the model

8.1 Deriving the Price Index

In this section we describe how to derive the expression for the price index in a variety of market structures.

8.1.1 Monopolistic competition

We follow Krugman (1980) and assume that there is a large mass of potential entrants that produce a differentiated good and operate in monopolistic competition. The production cost of q units is

$$c_i(q) = w_i \cdot \left(\frac{q}{z_i \cdot A_i \cdot f_1(L_i)} + F_i \right) \quad \text{and} \quad F_i = \mu_i \cdot f_2(L_i)$$

where w_i is the wage in market i , F_i is an entry cost, and the function $f_1(\cdot)$ and $f_2(\cdot)$ capture local agglomeration forces. The price index is given by

$$P_{ij}^{1-\sigma} = M_i \left(\frac{\sigma}{\sigma-1} \frac{\tau_{ij} w_i}{A_i \cdot f_1(L_i)} \right)^{1-\sigma}$$

We combine the free entry condition with the labor market clearing condition to obtain the number of firms in region i . The aggregate labor spending in region i is

$$w_i L_i = M_i \cdot \left(\frac{w_i}{A_i \cdot f_1(L_i)} \cdot \sum_j \tau_{ij} \left(\frac{\sigma}{\sigma-1} \frac{1}{A_i \cdot f_1(L_i)} \tau_{ij} w_i \right)^{-\sigma} \frac{E_j}{(P_j)^{1-\sigma}} \right) + M_i \cdot w_i F_i$$

The free entry condition in market i is

$$\frac{1}{\sigma} \sum_j \left(\frac{\sigma}{\sigma-1} \frac{1}{A_i} \frac{\tau_{ij} w_i}{P_j} \right)^{1-\sigma} E_j = w_i F_i$$

Combining these expressions, we get

$$M_i = \frac{1}{\sigma \mu_i} \cdot \frac{L_i}{f_2(L_i)}.$$

Thus, the price index is

$$\begin{aligned} P_{ij} &= \left(\frac{1}{\sigma \mu_i} \cdot \frac{L_i}{f_2(L_i)} \right)^{\frac{1}{1-\sigma}} \frac{\sigma}{\sigma-1} \frac{\tau_{ij} w_i}{A_i \cdot f_1(L_i)} = \\ &= (\tau_{ij} w_i) \left(\frac{L_i}{f_2(L_i)} \right)^{\frac{1}{1-\sigma}} \frac{1}{f_1(L_i)} \xi_i \\ &= (\tau_{ij} w_i) \cdot [\Psi(L_i)]^{\frac{1}{1-\sigma}} \xi_i \end{aligned}$$

where

$$\xi_i \equiv \frac{\sigma}{\sigma-1} \frac{(\mu_i)^{\frac{1}{\sigma-1}}}{A_i}$$

and

$$\Psi(L_i) \equiv \left(\frac{L_i}{f_2(L_i)} \right) \left(\frac{1}{f_1(L_i)} \right)^{1-\sigma}$$

8.1.2 Perfect competition

The price index shown in the main text can be micro-founded also with a perfect competition model, as in [Eaton and Kortum \(2002\)](#), with productivity spillovers. Assume that firms are heterogeneous in their productivity z , and the production cost of q units is:

$$c_i(q, z) = w_i \frac{q}{z A_i f_1(L_i)}$$

where w_i is the wage in market i and the function $f_1(\cdot)$ captures local agglomeration forces. Assume that the productivities are drawn from a Frechet distribution:

$$F_i(z) = e^{-T_i z^{-\theta}}$$

where $T_i > 0$ and $\theta > \sigma - 1$. It is straightforward to show that the price index is:

$$\begin{aligned} P_{ij}^{1-\sigma} &= \int p^{1-\sigma} dp = \\ &= M_i \int \left(\frac{w_i \tau_{ij}}{A_i z f_1(L_i)} \right)^{1-\sigma} df_i(z) = \\ &= M_i \left(\frac{w_i \tau_{ij}}{A_i f_1(L_i)} \right)^{1-\sigma} \int (z)^{\sigma-1} df_i(z) = \\ &= M_i \left(\frac{w_i \tau_{ij}}{A_i f_1(L_i)} \right)^{1-\sigma} \left[\Gamma \left(\frac{\theta + 1 - \sigma}{\theta} \right) \right]^{1/(1-\sigma)} \end{aligned}$$

and thus:

$$P_{ij} = \tau_{ij} w_i \Psi(L_i) \xi_i$$

where

$$\xi_i \equiv \frac{1}{A_i} \Gamma \left(\frac{\theta + 1 - \sigma}{\theta} \right)$$

and

$$\Psi(L_i) = \frac{1}{f_1(L_i)}$$

8.2 Deriving the Labor Supply Function

Suppose that a country c is a collection of regional labor markets $r = 1, \dots, R_c$. Each region has $K + 1$ sectors, including the home sector ($k = 0$). Countries are populated by a mass \bar{L}_c of heterogeneous individuals that decide their region of residence and their sector of employment. If employed in sector k of region r , we assume that individual ι 's utility is

$$U_{r,k}(\iota) = a_{r,k}(\iota) \cdot \tilde{v}_{r,k} \cdot \bar{U}(\zeta_{r,k} C, h), \quad (46)$$

where U is twice differentiable, strictly concave, increasing in consumption c , and decreasing in worked hours h . As in the one sector model, we impose a restriction on the utility function such that the substitution effect always dominates the income effect in the labor supply decision. We also assume that

$$C = U(\{C_{r,k}\}_k)$$

where $U(\cdot)$ is homogeneous of degree one, and $C_{r,k}$ is the sector-level consumption aggregator.

In expression 46, the term $\tilde{\nu}_{r,k}$ is the component of preferences common to all individuals, and the term $a_{r,k}(\iota)$ is individual ι 's idiosyncratic preference shifter, which is assumed to be independently drawn from a full-support distribution,

$$\{a_{r,k}(\iota)\}_{r,k} \sim F_c(\cdot).$$

If employed in sector k of region r , individual ι receives for each hour worked a wage per hour worked of $w_{r,k}$. Thus, the preference structure in (46) implies that individual ι 's payoff is

$$U_{r,k}(\iota) \equiv a_{r,k}(\iota) \cdot \nu_{r,k} \cdot \omega_{r,k}, \quad (47)$$

where $\omega_{r,k} \equiv w_{r,k}/P_r$ is the real wage in sector k and region r .²⁵

We assume that non-employed households obtain consumption in terms of home production, in which case they are employed in the home sector, $k = 0$. We set the real wage earned with unemployment to one, $\omega_r^0 = 1$.²⁶

The total labor supply in sector k of region r is

$$L_{r,k} = h_{r,k} \cdot N_{r,k} \quad (48)$$

where, in sector k of region r , $h_{r,k}$ is the average number of worked hours per employee and $N_{r,k}$ is the total number of employees.

In this environment, the average number of worked hours is a function of the real wage:

$$h_{r,k} = \Phi^h(\zeta_{r,k}\omega_{r,k}) \quad (49)$$

where the regularity conditions on $U(\cdot)$ imply that $\Phi^h(\cdot)$ is differentiable and increasing.²⁷

²⁵Since $\frac{\partial U}{\partial c} > 0$, the function $g(x) \equiv \max_h \cdot U(x \cdot h, h) = U(x\bar{h}(x), \bar{h}(x))$ is strictly increasing in x . Thus, for any utility function $u(c, h)$, the transformed utility $\tilde{u}(c, h) \equiv g^{-1}(u(c, h))$ yields the same choices and implies that

$$U_{r,k}(\iota) \equiv a_{r,k}(\iota) \cdot \nu_{r,k} \cdot \max_h \cdot \tilde{u}(\zeta_{r,k}\omega_{r,k}h, h) = a_{r,k}(\iota) \cdot (\tilde{\nu}_{r,k}\zeta_{r,k}) \cdot \omega_{r,k}.$$

²⁶Our normalization is equivalent to the assumption that home production has a linear technology that yields utility equivalent to one unit of the aggregate consumption good. Alternatively, [Dvorkin et al. \(2013\)](#) and [Caliendo et al. \(2015\)](#) assume that there is an exogenous unemployment benefit $b_{r,0} > 0$.

²⁷Note that $\frac{\partial \log h_{r,k}}{\partial \log \omega_{r,k}}$ is the Marshallian elasticity of labor supply. Formally, the Marshallian elasticity is strictly positive if, and only if,

$$\frac{\partial \bar{U}}{\partial C} + \frac{\partial^2 \bar{U}}{\partial C^2} \cdot C + \frac{\partial^2 \bar{U}}{\partial L \partial C} \cdot L > 0.$$

Also, we have that the total number of employees equals:

$$N_{r,k} = \bar{L}_c \cdot \int_{A_{r,k}} dF(a) \equiv \Phi_{r,k}^e(\{\nu_{r',k'}\omega_{r',k'}\})$$

$$\text{where } A_{r,k} \equiv \{a_{r,k} : \nu_{r,k}\omega_{r,k}a_{r,k} \geq \nu_{r',k'}\omega_{r',k'}a_{r',k'} \quad \forall r', k'\}.$$

Notice that the function $\Phi_{r,k}^e(\cdot)$ is homogeneous of degree zero with $\frac{\partial \Phi_{r,k}^e}{\partial \omega_{r,k}} > 0$ and $\frac{\partial \Phi_{r,k}^e}{\partial \omega_{r',k'}} < 0$.²⁸

This general labor supply structure covers the single-sector model with segmented labor markets presented in Section 2. The benchmark model corresponds to the case in which individuals cannot move across regions, and there is only one sector in each region, $K = 1$. In this case, individuals choose either to engage in home production or to sell working hours in the labor market.

Using the structure above, we can write the total labor supply as:

$$L_{r,k} = \Phi^h(\zeta_{r,k}\omega_{r,k}) \Phi_{r,k}^e(\{\nu_{r',k'}\omega_{r',k'}\}) \equiv \Phi_{r,k}(\boldsymbol{\omega}_c) \quad (50)$$

where $\boldsymbol{\omega}_c \equiv \{\omega_{r,k}\}_{r,k}$ is the vector of wages. The multiple sectors structure implies that the labor supply elasticities are:

$$\frac{\partial \Phi_{r,k}}{\partial \omega_{r,k}} > 0 \quad \text{and} \quad \frac{\partial \Phi_{r,k}}{\partial \omega_{r',k'}} < 0.$$

8.3 Parametric restrictions for empirical analysis

If employed in sector k of region r , we assume that individual l 's utility is

$$U_{r,k}(l) = a_{r,k}(l) \cdot \tilde{\nu}_{r,k} \cdot \bar{U}(\zeta_{r,k}C, h), \quad (51)$$

where \bar{U} is given by:

$$\bar{U}(\zeta_{r,k}C, h) = \left[(1 + \phi_h)\zeta_{r,k}C - \phi_h h^{1+\frac{1}{\phi_h}} \right]^{\frac{1}{1+\phi_h}}. \quad (52)$$

where C is a Cobb-Douglas aggregator of consumption levels $C_{r,k}$, and where $C_{r,k}$ is, in turn, a CES consumption aggregator of goods in sector k and region r . This functional form implies that the indirect utility derived from living in region r and working in sector k is simply the real wage $\omega_{r,k}$ multiplied by the preference shifters:

²⁸The homogeneity of $\Phi_{r,k}^e(\cdot)$ follows immediately from the definition of $A_{r,k}$. Since $F(\cdot)$ has full support, $\frac{\partial \Phi_{r,k}^e}{\partial \omega_{r,k}} > 0$ and $\frac{\partial \Phi_{r,k}^e}{\partial \omega_{r',k'}} < 0$ are implied by $A_{r',k'}(\tilde{\boldsymbol{\omega}}_c) \subset A_{r',k'}(\boldsymbol{\omega}_c)$ and $A_{r,k}(\boldsymbol{\omega}_c) \subset A_{r,k}(\tilde{\boldsymbol{\omega}}_c)$ whenever $\tilde{\omega}_{r,k} > \omega_{r,k}$ and $\tilde{\omega}_{r',k'} = \omega_{r',k'}$.

$$U_{r,k}(\iota) = a_{r,k}(\iota) \cdot (\nu_{r,k}\omega_{r,k}).$$

To characterize the allocation of individuals to regions and sectors, we follow [McFadden \(1980\)](#) and assume that $\{a_{r,k}(\iota)\}_{r,k}$ is independently drawn across individuals from the following GEV distribution:

$$\{a_{r,k}(\iota)\}_{r,k} \sim \exp \left[- \left(\sum_r A_r \cdot \left(\sum_{k=0}^K (a_{r,k})^{\phi_e} \right)^{\frac{\phi_m}{\phi_e}} \right)^{\frac{1}{\phi_m}} \right] \quad (53)$$

It is easy to verify that the aggregate labor supply in sector k of region r is given by

$$L_r^k = h_r^k \cdot l_r^k \cdot n_r \cdot \bar{L}_c \quad (54)$$

where h_r^k is the number of hours worked by individuals employed in sector k of region r ,

$$h_r^k = (\zeta_{r,k}\omega_{r,k})^{\phi_h}, \quad (55)$$

the share of residents of region r employed in sector k is:

$$l_r^k = \frac{(\nu_r^k \omega_{r,k})^\kappa}{1 + \sum_{s=1}^K (\nu_r^s \omega_{r,s})^\kappa}, \quad (56)$$

and the share of national population residing in region r is:

$$n_r = \frac{A_r \cdot (l_r^0)^{-\frac{\eta}{\kappa}}}{\sum_j A_j \cdot (l_j^0)^{-\frac{\eta}{\kappa}}}. \quad (57)$$

9 Analytical Proofs for the Empirical Methodology

9.1 Proof of Lemma 1

We write the change between two world equilibria in the model of Section 3. Consider first the Local Labor Market Module in (15)–(20). Using the Invertibility Assumption, the labor supply equation in (15) yields

$$\nu_{r,k}\omega_{r,k} = \tilde{\Phi}_{r,k}(\mathbf{L}_C^{t_0}|\phi),$$

which implies that

$$\log \hat{\nu}_{r,k}^t = \log \hat{\omega}_{r,k}^t - \log \frac{\Phi_{r,k}(\mathbf{L}_C^{t_0} \hat{\mathbf{L}}_C^t|\phi)}{\tilde{\Phi}_{r,k}(\mathbf{L}_C^{t_0}|\phi)} \quad (58)$$

where $\hat{\omega}_{j,k}^t = \hat{w}_{j,k}^t / \hat{P}_j^t$.

Immediately, equation (20) implies that

$$\log \hat{\xi}_{r,k}^t = \log \hat{\omega}_{r,k}^t - \log \hat{x}_{r,k}^t - \log \frac{\Psi_{r,k}(\mathbf{L}_C^{t_0} \hat{\mathbf{L}}_C^t | \boldsymbol{\psi})}{\Psi_{r,k}(\mathbf{L}_C^{t_0} | \boldsymbol{\psi})} \quad (59)$$

where, by definition,

$$\hat{x}_{j,k}^t = \frac{\hat{P}_{j,k}^t f_{j,k}(\{\tilde{\chi}_{ij,k}(\mathbf{x}_{j,k}^{t_0})\}_i)}{\hat{P}_j^t f_{j,k}(\{\tilde{\chi}_{ij,k}(\mathbf{x}_{j,k}^{t_0} \hat{\mathbf{x}}_{j,k}^t)\}_i)}.$$

Now consider the Trade Module in equations (23)–(24). These equations imply that

$$Y_{i,k}^{t_0} \cdot (\hat{\omega}_{i,k}^t \hat{L}_{i,k}^t \hat{P}_i^t) = \sum_j (x_{ij,k}^{t_0} e_{j,k} E_j) \cdot \hat{x}_{ij,k} \cdot \hat{e}_{j,k} \cdot \hat{E}_j \quad (60)$$

with

$$\hat{E}_j = \sum_{k=1}^K Y_{j,k}^{t_0} \cdot (\hat{\omega}_{j,k}^t \hat{L}_{j,k}^t \hat{P}_j^t) \quad (61)$$

$$\hat{x}_{ij,k}^t = \frac{1}{x_{ij,k}^{t_0}} \chi_{ij,k} \left(\left\{ \tilde{\chi}_{ij,k}(\mathbf{x}_{j,k}^{t_0}) \cdot \hat{\tau}_{ij,k}^t \cdot \frac{\hat{P}_i^t}{\hat{P}_j^t} \cdot \frac{\hat{\omega}_{i,k}^t}{\hat{\omega}_{j,k}^t} \cdot \frac{\hat{\xi}_{j,k}^t}{\hat{\xi}_{i,k}^t} \cdot \frac{\Psi_{j,k}(\mathbf{L}_C^{t_0} \hat{\mathbf{L}}_C^t | \boldsymbol{\psi})}{\Psi_{j,k}(\mathbf{L}_C^{t_0} | \boldsymbol{\psi})} \cdot \frac{\Psi_{i,k}(\mathbf{L}_C^{t_0} | \boldsymbol{\psi})}{\Psi_{i,k}(\mathbf{L}_C^{t_0} \hat{\mathbf{L}}_C^t | \boldsymbol{\psi})} \right\}_i \right) \quad (62)$$

$$\hat{P}_{j,k} = f_k \left(\left\{ \tilde{\chi}_{ij,k}(\mathbf{x}_{j,k}^{t_0}) \cdot \hat{P}_i^t \cdot \frac{\hat{\omega}_{i,k}^t}{\hat{\xi}_{i,k}^t} \cdot \frac{\Psi_{i,k}(\mathbf{L}_C^{t_0} | \boldsymbol{\psi})}{\Psi_{i,k}(\mathbf{L}_C^{t_0} \hat{\mathbf{L}}_C^t | \boldsymbol{\psi})} \right\}_i \right) \quad (63)$$

where, in sector k of region j at period t_0 , $X_{ij,k}^{t_0}$ denotes the imports from region i and $Y_{j,k}^{t_0}$ denotes total output.

Finally, the aggregate preferences across sectors implies that

$$\hat{e}_{j,k} = \chi_j \left\{ \tilde{\chi}_j(\mathbf{e}_j^{t_0}) \cdot \hat{P}_{j,k} \right\} \quad (64)$$

$$\hat{P}_j = f_j \left(\tilde{\chi}_j(\mathbf{e}_j^{t_0}) \cdot \hat{P}_{j,k} \right) \quad (65)$$

Define the vector of parameters, $\Theta \equiv (\boldsymbol{\phi}, \boldsymbol{\psi})$, and the vector of initial conditions, $W_c^{t_0} \equiv (\mathbf{L}_c^{t_0}, \mathbf{x}_c^{t_0}, \mathbf{e}_c^{t_0}, \mathbf{Y}_c^{t_0}, \mathbf{E}_c^{t_0})$. Equations (58)–(59) can be written as

$$F_{r,k}(\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | W_c^{t_0}, \Theta) = \hat{\epsilon}_{r,k}^t,$$

equations (60)–(65) can be written as

$$A_{r,k} \left(\left\{ \hat{\mathbf{x}}_c^t, \hat{\omega}_c^t, \hat{\mathbf{L}}_c^t \right\}_c, \left\{ \hat{\tau}_c^t, \hat{\epsilon}_c^t \right\} | \left\{ W_c^{t_0} \right\}_c, \Theta \right) = 0. \quad \blacksquare$$

9.2 Proof of Lemma 2

Define the vector of endogenous variables, $\hat{Y}_{r,k}^t \equiv \{\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{L}_c^t\}$. Equations (26) and (28) imply

$$F_{r,k}(Y_{r,k}^t | W_c^{t_0}, \Theta) = \hat{\epsilon}_{r,k}^t \quad (66)$$

$$A_{r,k}\left(\left\{\hat{Y}_c^t\right\}, \left\{\hat{\tau}_c^t, \hat{\epsilon}_c^t\right\} \mid \left\{W_c^{t_0}\right\}_c, \Theta\right) = 0. \quad (67)$$

Consider a log-linear approximation of the labor market module above:

$$F_{r,k}^{Z_{r,k}} \cdot d \log Z_{r,k}^t = d \log \epsilon_{r,k}^t + op(1) \quad (68)$$

where $F_{r,k}^Z \equiv \nabla_{Z_{r,k}} F_{r,k}(\hat{Z}_{r,k}^t | W_c^{t_0}, \Theta)$, and $op(1)$ is the first-order approximation error. Similarly, the trade module can be approximated as:

$$\sum_{i,d} A_{r,k}^{Z_{i,d}} \cdot d \log Y_{i,d}^t = \left[- \sum_{i,j,d} A_{r,k}^{\tau_{ij,d}} \cdot d \log \tau_{ij,d}^t \right] + \left[- \sum_{i,d} A_{r,k}^{\epsilon_{i,d'}} \cdot d \log \epsilon_{i,d}^t \right] + op(1) \quad (69)$$

where

$$A_{r,k}^{Z_{i,d}} \equiv \nabla_{Z_{i,d}} A_{r,k}\left(\left\{\hat{Y}_c^t\right\}, \left\{\hat{\tau}_c, \hat{\epsilon}_c\right\} \mid \left\{W_c^{t_0}\right\}_c, \Theta\right)$$

$$A_{r,k}^{\tau_{ij,d}} \equiv \nabla_{\tau_{ij,d}} A_{r,k}\left(\left\{\hat{Y}_c^t\right\}, \left\{\hat{\tau}_c, \hat{\epsilon}_c\right\} \mid \left\{W_c^{t_0}\right\}_c, \Theta\right)$$

$$A_{r,k}^{\epsilon_{i,d}} \equiv \nabla_{\epsilon_{i,d}} A_{r,k}\left(\left\{\hat{Y}_c^t\right\}, \left\{\hat{\tau}_c, \hat{\epsilon}_c\right\} \mid \left\{W_c^{t_0}\right\}_c, \Theta\right).$$

Equations (68) and (69) constitute a linear system with the following form

$$A(\mathbf{W}^{t_0}, \Theta) d\mathbf{Y}^t = \tilde{B}^1(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\tau}^t + \tilde{B}^2(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\epsilon}^t + op(1)$$

where $d\mathbf{Y}^t \equiv [d \log Y_{r,k}^t]_{r,k}$, $d\boldsymbol{\tau}^t \equiv [d \log \tau_{ij,k}^t]_{ij,k}$, and $d\boldsymbol{\epsilon}^t \equiv [d \log \epsilon_{r,k}^t]_{r,k}$. Thus, any solution of this system has the following form:

$$d\mathbf{Y}^t = B^1(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\tau}^t + B^2(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\epsilon}^t + op(1)$$

Substituting the relationship between τ_{ij} and observable shifters z_{ij} in equation (29),

$$d\mathbf{Y}^t \approx \rho B^1(\mathbf{W}^{t_0}, \Theta) dz^t + B^2(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\eta}^t + B_2(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\epsilon}^t$$

By setting $\hat{\tau}^t = \hat{z}^t$ and $\hat{\epsilon}^t = 1$, the same first-order approximation yields that $H_{r,k}^*(\hat{z}^t, \mathbf{W}^{t_0}) \approx B_1(\mathbf{W}^{t_0}, \Theta) dz^t$. Thus, the expression above can be written as

$$d \log Y_{r,k}^t \approx \rho H_{r,k}^*(\hat{z}^t, \mathbf{W}^{t_0}) + \hat{\mu}_{r,k}^t$$

where

$$\hat{\mu}_{r,k}^t \equiv B_{r,k}^1(\mathbf{W}^{t_0}, \Theta) \hat{\boldsymbol{\eta}}^t + B_{r,k}^2(\mathbf{W}^{t_0}, \Theta) d\boldsymbol{\epsilon}^t$$

By the exogeneity of $\hat{\boldsymbol{z}}^t$ in Condition 2,

$$E[\hat{\mu}_{r,k}^t | \hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0}] = B^2(\mathbf{W}^{t_0}, \Theta) E[\hat{\boldsymbol{\eta}}^t | \hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0}] + H_i^r(W^{t_0}, \beta) E[d\boldsymbol{\epsilon}^t | \hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0}] = \mathbf{0} \blacksquare$$

9.3 Asymptotic Properties of the Model-implied Optimal GMM

9.3.1 Identification

Let us define the following function:

$$J(\tilde{\Theta}, \bar{\Theta}) \equiv E \left[e_{r,k}^t(\tilde{\Theta}) \cdot \nabla_{\Theta} F_{r,k}(H_{r,k}^*(\hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0} | \bar{\Theta}) | W_c^{t_0}) \right]' W E \left[e_{r,k}^t(\tilde{\Theta}) \cdot \nabla_{\Theta} F_{r,k}(H_{r,k}^*(\hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0} | \bar{\Theta}) | W_c^{t_0}) \right].$$

Proposition 2. *Suppose Conditions 1 and 2 hold. If $\rho \neq 0$ and*

$$\text{rank } E \left[\nabla_{\tilde{\Theta}} F_{r,k} \left(\hat{\boldsymbol{x}}_{r,k}^t, \hat{\boldsymbol{\omega}}_{r,k}^t, \hat{\boldsymbol{L}}_c^t | W_c^{t_0}, \tilde{\Theta} \right) \cdot \nabla_{\Theta} F_{r,k} \left(H_{r,k}^*(\hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0} | \Theta_0) | W_c^{t_0} \right) \right] = \dim \Theta \quad (70)$$

Then,

$$\Theta = \arg \min_{\tilde{\Theta}} J(\tilde{\Theta}, \tilde{\Theta}) \quad (71)$$

Proof. Θ is the *unique global minimum* of $J(\tilde{\Theta}, \tilde{\Theta})$ if, and only if $J(\Theta, \Theta) < J(\tilde{\Theta}, \tilde{\Theta})$ for all $\tilde{\Theta} \neq \Theta$. To show this, we first establish that Θ is the unique minimum of the following constrained minimization problem:

$$\Theta = \arg \min_{\tilde{\Theta}} J(\tilde{\Theta}, \Theta_0) \quad \text{for a given } \Theta_0 \in \mathbb{R}^s. \quad (72)$$

The quadratic form of $J(\tilde{\Theta}, \Theta_0)$ immediately implies that $J(\tilde{\Theta}, \Theta_0) \geq 0$ for all $\tilde{\Theta}$. Also, Condition 2 implies that, for any $\tilde{\Theta}$, $E[e_{r,k}^t(\tilde{\Theta}) \cdot \nabla_{\Theta} F_{r,k}(H_{r,k}^*(\hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0} | \Theta_0) | W_c^{t_0})] = 0$ and, therefore, $J(\Theta, \Theta_0) = 0$. To show that Θ is the unique minimum of the constrained minimization problem, it is sufficient to show that $J(\tilde{\Theta}, \Theta_0)$ is convex in $\tilde{\Theta}$. To this end, recall that $e_{r,k}^t(\tilde{\Theta}) \equiv F_{r,k} \left(\hat{\boldsymbol{x}}_{r,k}^t, \hat{\boldsymbol{\omega}}_{r,k}^t, \hat{\boldsymbol{L}}_c^t | W_c^{t_0}, \tilde{\Theta} \right)$ and, therefore, the Hessian matrix of $J(\tilde{\Theta}, \Theta_0)$ is

$$D_{\tilde{\Theta}}^2 J(\tilde{\Theta}, \Theta_0) = G \left(\tilde{\Theta}, \Theta_0 \right)' W G \left(\tilde{\Theta}, \Theta_0 \right)$$

where $G \left(\tilde{\Theta}, \Theta_0 \right) \equiv E \left[\nabla_{\tilde{\Theta}} F_{r,k} \left(\hat{\boldsymbol{x}}_{r,k}^t, \hat{\boldsymbol{\omega}}_{r,k}^t, \hat{\boldsymbol{L}}_c^t | W_c^{t_0}, \tilde{\Theta} \right) \cdot \nabla_{\Theta} F_{r,k} \left(H_{r,k}^*(\hat{\boldsymbol{z}}^t, \mathbf{W}^{t_0} | \Theta_0) | W_c^{t_0} \right) \right]$.

Thus, to show that $D_{\tilde{\Theta}}^2 J(\tilde{\Theta}, \tilde{\Theta})$ is positive definite, it is sufficient to show that $G \left(\tilde{\Theta}, \Theta_0 \right)$ has full rank. By Lemma (1), $F_{r,k}$ is linear in $\tilde{\Theta}$, which implies that $G \left(\tilde{\Theta}, \Theta_0 \right)$ does not

depend on $\tilde{\Theta}$. By the rank condition in (70), $G(\tilde{\Theta}, \Theta_0)$ has full rank.

Second, we show that Θ is the unique minimum of the unconstrained minimization problem. Take any $\tilde{\Theta}$ such that $\tilde{\Theta} \neq \Theta$. Since Θ is the solution of the constrained problem in (72), $J(\Theta, \tilde{\Theta}) < J(\tilde{\Theta}, \tilde{\Theta})$. Moreover, $J(\Theta, \Theta) = J(\Theta, \tilde{\Theta}) = 0$, which implies that $J(\Theta, \Theta) < J(\tilde{\Theta}, \tilde{\Theta})$. ■

9.3.2 Consistency

Throughout this section, we assume that the rank condition in (70) is satisfied and, therefore, Θ is the unique global minimum of (71). To show the consistency of the Model-implied Optimal GMM in Definition (3), we only need to impose the sufficient conditions of Theorems 2.6 and 2.7 in [Newey and McFadden \(1994\)](#). In our environment, Condition 1 immediately implies that $h_{r,k}^t(\tilde{\Theta}, \bar{\Theta})$ is linear in $\tilde{\Theta}$. Thus, standard regularity conditions imply the consistency of $(\hat{\phi}, \hat{\psi})$.

Proposition 3. *Suppose the world economy is generated by the model in Section 2, satisfying Conditions 1 and 2. Suppose also that Θ is the unique global minimum of the problem in (71). Assume that Θ is in the interior of a convex set B , and (iii) $E[h_{r,k}^t(\tilde{\Theta}, \bar{\Theta})] < \infty$ for all $(\tilde{\Theta}, \bar{\Theta}) \in B \times B$. Then, $\hat{\Theta} \xrightarrow{p} \Theta$.*

9.3.3 Asymptotic Normality

We now derive the asymptotic distribution of the Model-implied Optimal GMM. The estimator can be written as

$$\begin{aligned}\hat{\Theta} &= \arg \min_{\tilde{\Theta}} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \hat{\Theta}) \right) \hat{W} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \hat{\Theta}) \right)' \\ \hat{\Theta} &= \arg \min_{\tilde{\Theta}} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \Theta_0) \right) \hat{W} \left(\sum_{r,k,t} h_{r,k}^t(\tilde{\Theta}, \Theta_0) \right)'\end{aligned}$$

where \hat{W} is positive definite matrix with $\hat{W} \xrightarrow{p} W$, and Θ_0 is an arbitrary parameter vector.

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:

$$(\hat{\Theta}, \hat{\Theta}) \equiv \arg \min_{\tilde{\Theta}, \bar{\Theta}} \left(\sum_{r,k,t} H_{r,k}^t(\tilde{\Theta}, \bar{\Theta}) \right) \bar{W} \left(\sum_{r,k,t} H_{r,k}^t(\tilde{\Theta}, \bar{\Theta}) \right)' \quad (73)$$

where $H_i^t(\tilde{\Theta}, \bar{\Theta}) \equiv \begin{bmatrix} h_{r,k}^t(\tilde{\Theta}, \bar{\Theta}) & h_{r,k}^t(\bar{\Theta}, \Theta_0) \end{bmatrix}$.

The estimator in expression (73) is asymptotically normal under the standard regularity conditions in Theorem 4.3 of [Newey and McFadden \(1994\)](#). In our environment, Condition 1 immediately implies that $H_i^t(\tilde{\Theta}, \bar{\Theta})$ is continuously differentiable on $(\tilde{\Theta}, \bar{\Theta})$.

Lemma 3. *Suppose the world economy is generated by the model in Section 2, satisfying Conditions 1 and 2. Suppose also that (i) the conditions in Proposition (3) hold, (ii) Θ is in the interior B , with $E[h_{r,k}^t(\Theta, \Theta)] = 0$ and $E[|h_{r,k}^t(\Theta, \Theta)|^2] < \infty$, (iii) $E[\sup \nabla_{\tilde{\Theta}} |h_{r,k}^t(\tilde{\Theta}, \tilde{\Theta})|] < \infty$, (iv) $\tilde{G}W\tilde{G}$ non-singular with $\tilde{G} \equiv [\nabla_{(\tilde{\Theta}, \bar{\Theta})} H_{r,k}^t(\tilde{\Theta}, \bar{\Theta})]$. Then,*

$$\sqrt{N} \left((\hat{\Theta}, \hat{\Theta} - (\Theta, \Theta)) \right) \xrightarrow{d} N \left(0, (\tilde{G}'W\tilde{G})^{-1} (\tilde{G}'W\tilde{\Lambda}W\tilde{G}) (\tilde{G}'W\tilde{G})^{-1} \right)$$

with $\tilde{\Lambda} \equiv E \left[H_i^t(\tilde{\Theta}, \bar{\Theta})' H_i^t(\tilde{\Theta}, \bar{\Theta}) \right]$, and $N = T \cdot K \cdot I$.

We use this lemma to establish the asymptotic distribution of the second-stage estimator $\hat{\Theta}$. Define $h_{r,k}^t \equiv h_{r,k}^t(\Theta, \Theta)$ and $\bar{h}_{r,k}^t \equiv h_{r,k}^t(\Theta, \Theta_0)$. By definition, \tilde{G} and $\tilde{\Lambda}$ are given by

$$\tilde{\Lambda} = E \begin{bmatrix} h_{r,k}^{t'} h_{r,k}^t & h_{r,k}^{t'} \bar{h}_{r,k}^t \\ \bar{h}_{r,k}^{t'} h_{r,k}^t & \bar{h}_{r,k}^{t'} \bar{h}_{r,k}^t \end{bmatrix} \quad \text{and} \quad \tilde{G} = \begin{bmatrix} G & G_1 \\ 0 & G \end{bmatrix}$$

where

$$G \equiv E \left[\nabla_{\tilde{\Theta}} F_{r,k} \left(\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | W_c^{t_0}, \tilde{\Theta} \right) \cdot \nabla_{\Theta} F_{r,k} \left(H_{r,k}^*(\hat{z}^t, \mathbf{W}^{t_0} | \Theta_0) | W_c^{t_0} \right) \right].$$

$$G_1 \equiv E \left[e_{r,k}^t(\Theta) \nabla_{\Theta_0} \nabla_{\Theta} F_{r,k} \left(H_{r,k}^*(\hat{z}^t, \mathbf{W}^{t_0} | \Theta_0) | W_c^{t_0} \right) \right]$$

By Condition 2, any function of $(\hat{z}^t, \mathbf{W}^{t_0})$ is orthogonal to $e_{r,k}^t(\Theta)$, which implies that $G_1 = 0$. Thus, Lemma (3) immediately implies the following result.

Proposition 4. *Suppose the world economy is generated by the model in Section 2, satisfying Conditions 1 and 2. Suppose also that (i) the conditions in Proposition (3) hold, (ii) Θ is in the interior B , with $E[h_{r,k}^t(\Theta, \Theta)] = 0$ and $E[|h_{r,k}^t(\Theta, \Theta)|^2] < \infty$, (iii) $E[\sup \nabla_{\tilde{\Theta}} |h_{r,k}^t(\tilde{\Theta}, \tilde{\Theta})|] < \infty$, (iv) GWG non-singular. Then,*

$$\sqrt{N} \left(\hat{\Theta} - \Theta \right) \xrightarrow{d} N \left(0, (G'WG)^{-1} (G'W\Lambda WG) (G'WG)^{-1} \right)$$

where $\Lambda \equiv E [h_{r,k}^{t'} h_{r,k}^t]$, and $N = T \cdot K \cdot I$. If $W = \Lambda^{-1}$, then

$$\sqrt{N} (\hat{\Theta} - \Theta) \xrightarrow{d} N \left(0, (G' \Lambda^{-1} G)^{-1} \right).$$

9.4 Proof of Theorem 2

To prove the result, we follow Chamberlain (1987) to construct the moment conditions that minimizes the asymptotic variance of the GMM estimator. Our treatment of the problem follows closely Newey and McFadden (1994).

Consider the class of GMM estimators obtained from the moment condition in

$$E [e_{r,k}^t(\Theta) \cdot H_{r,k}(\hat{z}^t, W^{t_0})] = 0 \quad \text{for any function } H_{r,k}(\cdot).$$

Using the optimal weighting matrix in the formula in Proposition (4), we obtain the asymptotic variance of the GMM estimator for any function $H_{r,k}(\cdot)$:

$$V(H) = (E [H_{r,k}(\hat{z}^t, W^{t_0}) G_{r,k}(\hat{z}^t, W^{t_0})])^{-1} \left(E [H_{r,k}(\hat{z}^t, W^{t_0}) e_{r,k}(\Theta) e_{r,k}(\Theta)' H_{r,k}(\hat{z}^t, W^{t_0})'] \right) (E [H_{r,k}(\hat{z}^t, W^{t_0}) G_{r,k}(\hat{z}^t, W^{t_0})])^{-1'} \quad (74)$$

where

$$G_{r,k}(\hat{z}^t, W^{t_0}) \equiv E [\nabla_{\Theta} e_{r,k}(\Theta) | \hat{z}^t, W^{t_0}]$$

We first establish the following the lemma.

Lemma 4. *Consider the function:*

$$\begin{aligned} H_{r,k}^*(\hat{z}^t, W^{t_0}) &\equiv G_{r,k}(\hat{z}^t, W^{t_0})' (\Omega(\hat{z}^t, W^{t_0}))^{-1} \\ \Omega(\hat{z}^t, W^{t_0}) &= E [e_{r,k}(\Theta) e_{r,k}(\Theta)' | \hat{z}^t, W^{t_0}]. \end{aligned}$$

Then, $V(H) - V(H^*)$ is positive semi-definite for any function $H(\cdot)$.

Proof of Lemma 4. The asymptotic variance of H^* is

$$V(H^*) = E \left[G_{r,k}(\hat{z}^t, W^{t_0})' (\Omega(\hat{z}^t, W^{t_0}))^{-1} G_{r,k}(\hat{z}^t, W^{t_0}) \right].$$

Thus,

$$V(H) - V(H^*) = (E [H_{r,k}^t G_{r,k}^t])^{-1} \left(E [(H_{r,k}^t e_{r,k}^t) (H_{r,k}^t e_{r,k}^t)'] \right) (E [H_{r,k}^t G_{r,k}^t])^{-1'} - (E [G_{r,k}^t \Omega^{-1} G_{r,k}^t])^{-1}$$

$$= (E [H_{r,k}^t G_{r,k}^t])^{-1} \left(E \left[(H_{r,k}^t e_{r,k}^t) (H_{r,k}^t e_{r,k}^t)' \right] - E [H_{r,k}^t G_{r,k}^t] (E [G_{r,k}^t \Omega^{-1} G_{r,k}^t])^{-1} E [H_{r,k}^t G_{r,k}^t]' \right) (E [H_{r,k}^t G_{r,k}^t])^{-1}.$$

Let us define

$$U_{r,k}^t \equiv H_{r,k}^t e_{r,k}^t - E \left[(H_{r,k}^t e_{r,k}^t) (G_{r,k}^{t'} \Omega^{-1} e_{r,k}^t)' \right] (E [G_{r,k}^t \Omega^{-1} G_{r,k}^t])^{-1} G_{r,k}^{t'} \Omega^{-1} e_{r,k}^t.$$

This implies that

$$E [U_{r,k}^t U_{r,k}^{t'}] = E \left[(H_{r,k}^t e_{r,k}^t) (H_{r,k}^t e_{r,k}^t)' \right] - E [H_{r,k}^t G_{r,k}^t] (E [G_{r,k}^t \Omega^{-1} G_{r,k}^t])^{-1} E [H_{r,k}^t G_{r,k}^t]'. \quad \blacksquare$$

Therefore,

$$V(H) - V(H^*) = (E [H_{r,k}^t G_{r,k}^t])^{-1} (E [U_{r,k}^t U_{r,k}^{t'}]) (E [H_{r,k}^t G_{r,k}^t])^{-1'}.$$

Hence, $V(H) - V(H^*)$ is positive semi-definite because $E [U_{r,k}^t U_{r,k}^{t'}]$ is positive semi-definite. \blacksquare

From Lemma (4), the optimal IV is

$$G_{r,k}(\hat{\mathbf{z}}^t, W^{t_0}) \equiv E [\nabla_{\Theta} e_{r,k}(\Theta) | \hat{\mathbf{z}}^t, W^{t_0}] = E \left[\nabla_{\Theta} F_{r,k} \left(\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | W_c^{t_0}, \Theta \right) | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0} \right].$$

From Lemma (1), $F_{r,k}(\cdot | \Theta)$ is linear in Θ . So,

$$G_{r,k}(\hat{\mathbf{z}}^t, W^{t_0}) = \nabla_{\Theta} F_{r,k} \left(E \left[\hat{x}_{r,k}^t, \hat{\omega}_{r,k}^t, \hat{\mathbf{L}}_c^t | \hat{\mathbf{z}}^t, \mathbf{W}^{t_0} \right] | W_c^{t_0}, \Theta \right).$$

Using Lemma (2),

$$G(\hat{\mathbf{z}}^t, W^{t_0}) = \rho \nabla_{\Theta} F_{r,k} (H_{r,k}^*(\hat{\mathbf{z}}^t, \mathbf{W}^{t_0} | \Theta) | W_c^{t_0}, \Theta).$$

From expression (74), it is straightforward to see that the asymptotic variance is invariant to constant linear transformations of $H_{r,k}^*(\cdot)$. \blacksquare

9.5 Additional results

Lemma 5. *Let $x \in \mathbb{R}_+^N$ be the solution of the system $H_n(x) = 0$ for $n = 1, \dots, N$. Suppose that (i) $H_i(tx) = 0 \Leftrightarrow H_i(x) = 0$, and (ii) H_i is strictly increasing in x_n for all $n \neq i$. Then, x is unique up to a scalar.*

Proof. We proceed by contradiction. Take x and x' such that $x' \neq Kx$ for all $K \in \mathbb{R}$ and $H_i(x) = H_i(x') = 0$ for all i . Define $\kappa \equiv \arg \min_n \{x_n/x'_n\}$, and consider the normalized

vector $x'' \equiv \kappa x'$. By the definition of κ , $x''_n \leq x_n$ for all n with at least one strict inequality and at least one equality. By condition (ii), this implies that $H_j(x'') > H_j(x) = 0$ for j with $x''_j = x_j$. But condition (i) implies that $H_j(x'') = 0$. ■

Proposition 5. *Suppose that $0 \leq \phi_i(\omega) < \varepsilon$ and $0 \leq \psi_i(L) \leq 1$ for all i . Then, the system of equations (3), (5), (9) and (11) determines a unique solution for $\{\omega_j, L_j, x_{jj}, P_j\}$.*

Proof. This is equivalent to establishing that $\{w_j, P_j\}$ is the unique solution of the following system:

$$F_i(w, P) = G_i(w, P) = 0 \quad \text{for all } i$$

where

$$F_i(w, P) \equiv (P_i)^{-\varepsilon} - \sum_j K_{ij} \cdot \tilde{\Psi} \left(\frac{w_j}{P_j} \right) \cdot (w_j)^{-\varepsilon} \quad (75)$$

$$G_i(w, P) \equiv - (w_i)^{1+\varepsilon} \left[\left(\tilde{\Psi} \left(\frac{w_i}{P_i} \right) \right)^{-1} \Phi \left(\frac{w_i}{P_i} \right) \right] + \sum_j B_{ij} \cdot (P_j)^\varepsilon \cdot \left(w_j \Phi \left(\frac{w_j}{P_j} \right) \right) \quad (76)$$

with $\tilde{\Psi}(\omega) \equiv \Psi(\Phi(\omega))$, $K_{ji} \equiv (\tau_{ji})^{-\varepsilon} \xi_j$, and $B_{ij} \equiv \xi_i (\tau_{ij})^{-\varepsilon}$. To establish the proposition, we show that condition (i) and (ii) of Lemma 5 are satisfied by the system in (75)–(76). It is easy to check that condition (i) holds since $F_i(tw, tP) = t^{-\varepsilon} F_i(w, P)$ and $G_i(tw, tP) = t^{1+\varepsilon} G_i(w, P)$. Thus, it is sufficient to show that

$$\frac{\partial F_i}{\partial P_j} P_j > 0 \quad \text{for } i \neq j, \quad \text{and} \quad \frac{\partial F_i}{\partial w_j} w_j > 0 \quad \text{for all } j$$

$$\frac{\partial G_i}{\partial P_j} P_j > 0 \quad \text{for all } j, \quad \text{and} \quad \frac{\partial G_i}{\partial w_j} w_j > 0 \quad \text{for all } j \neq i.$$

Case 1: $0 < \psi < 1$. Using the fact that $\frac{d \log \tilde{\Psi}(\omega)}{d \log \omega} = \psi \phi$,

$$\frac{\partial F_i}{\partial P_j} P_j = \psi \phi \cdot K_{ji} \cdot \tilde{\Psi} \left(\frac{w_i}{P_i} \right) \cdot (w_j)^{-\varepsilon} \quad \text{for } i \neq j$$

$$\frac{\partial F_i}{\partial w_j} w_j = (\varepsilon - \psi \phi) \cdot K_{ji} \cdot \tilde{\Psi} \left(\frac{w_i}{P_i} \right)^\varepsilon \cdot (w_j)^{-\varepsilon} \quad \text{for all } j$$

$$\frac{\partial G_i}{\partial P_j} P_j = B_{ij} (P_j)^\varepsilon \cdot \left(w_j \Phi \left(\frac{w_j}{P_j} \right) \right) \cdot (\varepsilon - \phi) \quad \text{for } i \neq j$$

$$\frac{\partial G_i}{\partial P_i} P_i = \left(w_i \Phi \left(\frac{w_i}{P_i} \right) \right) (P_i^\varepsilon x_{ii}^{-1}) [(1 - \psi) \phi + x_{ii} (\varepsilon - \phi)]$$

$$\frac{\partial G_i}{\partial w_j} w_j = B_{ij} (P_j)^\varepsilon \cdot \left(w_j \Phi \left(\frac{w_j}{P_j} \right) \right) (1 + \phi) \quad \text{for } i \neq j$$

It is straightforward to check that $\frac{\partial G_i}{\partial w_j} w_j > 0$ and $\frac{\partial F_i}{\partial P_j} P_j > 0$. Finally, $\psi < 1$ and $\phi \leq \varepsilon$ implies that $\varepsilon - \phi\psi \geq 0$ and, therefore, $\frac{\partial F_i}{\partial w_j} w_j > 0$.

Case 2: $\psi = 1$.

$$(P_i)^{-\varepsilon} = \sum_j (\tau_{ji})^{-\varepsilon} \cdot w_j^{-\varepsilon}$$

$$G_i(w, P) \equiv w_i \cdot \Phi \left(w_i \left(\sum_k (\tau_{ki})^{-\varepsilon} \cdot w_k^{-\varepsilon} \right)^{\frac{1}{\varepsilon}} \right) - \sum_j \frac{(\tau_{ij})^{-\varepsilon} w_i^{-\varepsilon}}{\sum_k (\tau_{kj})^{-\varepsilon} \cdot w_k^{-\varepsilon}} \cdot \left(w_j \Phi \left(w_j \left(\sum_k (\tau_{kj})^{-\varepsilon} \cdot w_k^{-\varepsilon} \right)^{\frac{1}{\varepsilon}} \right) \right)$$

Thus,

$$\frac{\partial G_i}{\partial w_d} w_d = -w_i L_i \cdot x_{di} \phi + \phi \sum_j x_{ij} \cdot x_{dj} \cdot w_j L_j - x_{id} w_d L_d - \varepsilon \sum_j x_{ij} x_{dj} \cdot w_j L_j$$

$$\frac{\partial G_i}{\partial w_d} w_d = -w_i L_i \cdot x_{di} \phi - x_{id} w_d L_d - (\varepsilon - \phi) \sum_j x_{ij} x_{dj} \cdot w_j L_j$$

Thus, the condition is satisfied if $\varepsilon \geq \phi \geq 0$. ■

10 Data Appendix

10.1 Data Construction

10.1.1 World Trade Matrix

We construct a matrix of bilateral sector-level trade flows among 50 US states and 35 foreign countries for 1997, 2002, 2007 and 2012. To this end, we merge information on bilateral trade flows of 36 countries extracted from the World Input-Output Database (WIOD) 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 consider the 8 sectors shown in Table 4 constructed from the aggregation of the tradeable industries in the WIOD and CFS. In order to avoid zeros in the trade matrix, we merge DC and Maryland into a single state, and consider the subset of countries in Table 5. We proceed in four steps.

First, we construct foreign trade flows of US states for each sector, year and foreign country. Let $(Z_{dj}^{kt}, Z_{jd}^{kt})$ 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 $(Z_{dj}^{kt}, Z_{jd}^{kt})$ 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 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 WIOD database. To this end, we use US domestic sales in the WIOD 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 WIOD 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:

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

10.1.2 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 8 sectors (Table 4).²⁹ We consider the sample all individuals aged between 16 and 64 years in the survey, and we follow the cleaning procedure of [Autor et al. \(2008\)](#) 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 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.

²⁹To compute labor market outcomes, we build a crosswalk table between the sectors in Table 4 and the NAICS-based industry classification in the CPS-MORG. The crosswalk table is available upon request.

Industry Description	CFS	US Foreign Trade	WIOD	
	SCTG	HS (2-digit)	release 2013	relese 2016
Agriculture, Hunting, Forestry and Fishing	01-05	01-14	01	01-03
Food, Beverages and Tobacco	06-09	15-24	03	05
Mining, Coke, Refined Petroleum	10-19	25-27	02, 08	04, 10
Chemical Products, Plastic, Rubber	20-24	28-40	09-10	11-13
Wood, Pulp, Paper, Printing, Textiles, Leather	25-30	41-67	04-07	06-09
Non-Metallic Mineral, Basic Metals, Machinery	31-34	68-84	11-13	14-15, 19
Electrical and Optimal Equipment, Transport Equipment	35-38	85-92	14-15	17-18, 20-21
Other	39-43, 99, 00	93-99	16-35	22-56

Table 4: Industry Description

10.1.3 Construction of Price Index

To construct the state-sector price index, we use the Nielsen HomeScan dataset, for the years 2004, 2007 and 2012. This dataset provides detailed information on purchases, trips of purchases, household characteristics and product characteristics. We first adjust the raw price data to control for product, buyer and retailer heterogeneity (see Handbury and Weinstein (2015) for a detailed discussion). The adjustment is accomplished by running the following regression by each year:

$$p_{usrh,t} = \alpha_{u,t} + \alpha_{s,t} + \alpha_{r,t} + Z_{h,t}\beta_t + \epsilon_{usrh,t}$$

where $p_{usrh,t}$ is the log of price of UPC u , state s , store r and household h in year t ; $\alpha_{u,t}$, $\alpha_{s,t}$ and $\alpha_{r,t}$ are UPC, state and store fixed effects in year t ; $Z_{h,t}$ contains a vector of household characteristics variables and β_t is a vector of corresponding coefficients. Therefore $Z_h\beta_t$

Sample of Countries			
AUS	Australia	ITA	Italy
AUT	Austria	JPN	Japan
BLX	Belgium-Luxembourg	MEX	Mexico
BAL	Estonia-Latvia-Lithuania	NLD	Netherlands
BRA	Brazil	POL	Poland
BGR	Bulgaria	PRT	Portugal
CAN	Canada	KOR	Republic of Korea
CHN	China	ROU	Romania
CZE	Czech Republic	RUS	Russia
DNK	Denmark	SVK	Slovak Republic
FIN	Finland	SVN	Slovenia
FRA	France	ESP	Spain
DEU	Germany	SWE	Sweden
GRC	Greece	TUR	Taiwan
HUN	Hungary	TWN	Turkey
IDN	India	GBR	United Kingdom
IND	Indonesia	RoW	Rest of World
IRL	Ireland		

Table 5: Sample of Countries

controls for buyer heterogeneity since it carries the influence of household characteristics on the price level. The household characteristics variables include a set demographic dummies of household size, gender, age, marital status and race of the head of household; and also household income which is correlated with buyer search intensity. Thus, the adjusted price, denoted as $\tilde{P}_{usrh,t}$, equals

$$\tilde{P}_{usrh,t} \equiv \exp(p_{usrh,t} - \hat{\alpha}_{r,t} - Z_{h,t}\hat{\beta}_t)$$

The second step is to construct the Exact Price Index (EPI), following Broda and Weinstein (2010). It takes into account of product, buyer and retailer heterogeneity by using the adjusted priced discussed above, and also make correction for product availability differences and substitution biases. Therefore, the EPI of state s in year t is the price level, relative to the national average, that a consumer would face if all products varieties were available.

Before showing the formal expression for the EPI, we start with a set of definitions. First, Define $g \in \{1, \dots, G\}$ be the product group; $b \in \{1, \dots, B\}$ be the product brand-module; and $u \in \{1, \dots, U\}$ be the product UPC. These terms of categories are defined in the same way as Nielsen Scan data. Then, let B_g and U_g denote the sets of all brand-modules and UPCs in the product group g ; and U_b denote the set of all UPCs in the brand-module b .

Next, we define the subsets of UPCs, and brand-modules that are available and purchased in state s and year t . Denote $U_{bs,t}$, $U_{gs,t}$ and $B_{gs,t}$ as the subset of all UPCs in brand-module b that are purchased in state s in year t , the subset of all UPCs in group g that are purchased in state s in year t and the subset of all brand-modules in product group g that are purchased in state s in year t . Denote $U_{bs,t}$, $U_{gs,t}$ and $B_{gs,t}$ as the subset of all UPCs in brand-module b that are purchased in state s in year t , the subset of all UPCs in group g that are purchased in state s in year t and the subset of all brand-modules in product group g that are purchased in state s in year t . Let $v_{us,t}$ denote the value of UPC u purchased in state s in year t . With the adjusted price $\tilde{P}_{usrh,t}$, we define $\tilde{v}_{us,t} \equiv \sum_{h \in H_c} \sum_{r \in R_c} \tilde{P}_{usrh,t} q_{ucr,t}$, where $q_{ucr,t}$ is the quantity of UPC u purchased by household h in state s , store r in year t . With the adjusted price $\tilde{P}_{usrh,t}$, we define $\tilde{v}_{us,t} \equiv \sum_{h \in H_c} \sum_{r \in R_c} \tilde{P}_{usrh,t} q_{ucr,t}$, where $q_{ucr,t}$ is the quantity of UPC u purchased by household h in state s , store r in year t .

The Exact Price Index is:

$$EPI_{s,t} = \prod_{g \in G} [CEPI_{gs,t} VA_{gs,t}]^{w_{gs}}$$

where $CEPI_{gs,t} \equiv \prod_{u \in U_{gs,t}} \left(\frac{\tilde{v}_{us,t}/q_{us,t}}{\sum_s \tilde{v}_{us,t}/\sum_s q_{us,t}} \right)^{w_{us}}$ is the conventional exact price index for product group g , which is a sales-weighted average of each UPC price available in state s and

year t ; $VA_{gs,t} \equiv (s_{gs,t})^{\frac{1}{1-\sigma_g^a}} \prod_{b \in B_{gc,t}} (s_{bs,t})^{\frac{w_{bc}}{1-\sigma_g^w}}$ is the variety adjustment term to adjust for the different availability of good in different states, where the first term corrects the importance of missing products and, the second term adjusts variety availability for each brand-module in a state. $s_{gs,t}$ and $s_{bs,t}$ are defined as share of national brand-module b expenditures that is spent on the set of UPCs that are sold in state s in year t , and the share of national product-group expenditures g that is spent on the set of brand-modules that are sold in in state s in year t .³⁰ w_{us} , w_{bs} and w_{gs} are log-ideal UPC, brand-model and group CES weights defined by Sato (1976) and Vartia (1976). σ_g^a and σ_g^w are elasticities across brand-modules within a product group, and across UPCs within a brand-module, respectively.

10.2 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 $e_i^{k,t_0} \equiv E_i^{k,t_0} / \sum_k E_i^{k,t_0}$ 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 , $\hat{l}_r^{k,t}$:

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

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

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

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

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

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

³⁰Since the Nielsen scan data are just a sample of the full set of UPCs available in each state, we estimate the two shares as the asymptotes of the shares accumulation curves in each state and each year (see Handbury and Weinstein (2015) for detailed discussion).

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

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

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,

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

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

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

4) Create domestic trade shares

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

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

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

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

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.