Aggregate Implications of Firm Heterogeneity:
A Nonparametric Analysis of Monopolistic Competition Trade Models*

Rodrigo Adão
Chicago Booth & NBER
rodrigo.adao@chicagobooth.edu

Costas Arkolakis
Yale & NBER
costas.arkolakis@yale.edu

Sharat Ganapati
Georgetown
sg1390@georgetown.edu

November 2, 2020

Abstract

We measure the role of firm heterogeneity in counterfactual predictions of monopolistic competition trade models without parametric restrictions on the distribution of firm fundamentals. We show that two bilateral elasticity functions are sufficient to nonparametrically compute the counterfactual aggregate impact of trade shocks, and recover changes in economic fundamentals from observed data. These functions are identified from two semiparametric gravity equations governing the impact of bilateral trade costs on the extensive and intensive margins of firm-level exports. Applying our methodology, we estimate elasticity functions that imply an impact of trade costs on trade flows that falls when more firms serve a market because of smaller extensive margin responses. Compared to a baseline where elasticities are constant, firm heterogeneity amplifies both the gains from trade in countries with more exporter firms, and the welfare gains of European market integration in 2003-2012.

Keywords: International Trade, Nonparametric counterfactuals, Semiparametric estimation, Firm Heterogeneity

*We thank Xiaohong Chen, Phil Haile, Giovanni Compiani, Chang-Tai Hsieh, Samuel Kortum, and Steve Redding for valuable comments. We thank Takashi Onada, Mengly Xu, and Linghui Wu for research assistance. We are grateful to Peter Morrow for sharing his data. The support of the National Science Foundation under grant 2018234 is gratefully acknowledged. Part of this research was conducted while Ganapati was a Fellow at Dartmouth College. All errors are our own.

1 Introduction

The international trade field has been transformed by the study of firm heterogeneity. The observed correlation between firm characteristics and export participation is the cornerstone of the workhorse monopolistic competition model, linking decisions of heterogeneous firms to international trade and welfare – see Melitz (2003) and, for a review, Melitz and Redding (2014). Typically, to specify firm heterogeneity, the literature relies on restrictive parametric assumptions about the distribution of firm fundamentals. Such restrictions are useful in practice: approximations help with estimation and functional forms facilitate extrapolating aggregate counterfactual predictions from observed firm-level outcomes. However, an important question is whether these parametric distributional assumptions are essential for incorporating firm heterogeneity into empirical and counterfactual analyses in international trade. Furthermore, to what extent do these assumptions determine the implications of firm heterogeneity for the economy's aggregate response to international trade shocks?[1]

This paper aims to measure the aggregate importance of firm heterogeneity by developing a foundation for a widely used class of monopolistic competition trade models of firm heterogeneity, without any parametric restrictions on the distribution of firm fundamentals. We show that two elasticity functions are sufficient to nonparametrically compute the counterfactual aggregate impact of trade shocks, and recover changes in economic fundamentals from observed data. These functions are identified from two semiparametric gravity equations governing the impact of bilateral trade costs on the extensive and intensive margins of firm-level exports. Estimating these equations only requires bilateral data on a trade cost shifter, the share of a country's firms exporting to a destination, and their average exports. Conditional on these elasticity functions, more information on cross-section variation in firm-level outcomes does not affect the model's aggregate counterfactual predictions. We exploit this characterization to provide estimates of these two sufficient elasticity functions and quantify the importance of firm heterogeneity in measuring the aggregate implications of international market integration.

Our starting point is an augmented version of the multi-country workhorse monopolistic competition model with constant elasticity of substitution (CES) preferences. Firms are heterogeneous with respect to productivity, demand, and variable and fixed trade costs across destinations. An extensive literature has imposed different parametric distributional

[1] A series of recent papers have shown that parametric distributional assumptions are important to determine both the adjustment margins driving the model's counterfactual predictions and the set of moments used for identification – e.g., Chaney (2008); Arkolakis, Costinot and Rodríguez-Clare (2012); Melitz and Redding (2014, 2015); Bas, Mayer and Thoenig (2017); Fernandes, Klenow, Denisse Periola, Meleshchuk and Rodriguez-Clare (2019).

assumptions on some of these dimensions of firm heterogeneity to generate the observed patterns of cross-firm variation in productivity, sales, and entry in different markets – e.g., Chaney (2008), Helpman, Melitz and Rubinstein (2008), Eaton, Kortum and Kramarz (2011), Redding (2011), Head, Mayer and Thoenig (2014). In contrast, we do not restrict the distribution of these firm fundamentals. We thus only acknowledge that such heterogeneity exists and evaluate how it affects the response of aggregate outcomes to trade shocks.

We show that the distribution of all these sources of firm heterogeneity can be folded into two functions that yield semiparametric gravity equations for the extensive margin (i.e., exporter firm share) and intensive margin (i.e., average firm exports) of firm sales in each market. The extensive margin gravity equation links a flexible function of the exporter firm share to exogenous bilateral trade costs, and endogenous origin and destination fixed-effects (i.e., wages and price indices). The intensive margin gravity equation connects average firm exports to a flexible function of the exporter firm share, as well as bilateral trade costs and origin and destination fixed-effects.[2] The functions in the two gravity equations control the elasticities of the extensive and intensive margins of firm exports to changes in bilateral trade costs. The elasticity of each margin is an arbitrary function of the initial exporter firm share. These functions also determine the aggregate trade elasticity, which combines the elasticities of the two margins and, thus, may vary arbitrarily with the initial exporter firm share.

With this characterization, we establish that the two gravity elasticity functions of firm exports summarize the role that firm heterogeneity plays in the model's counterfactual predictions following shocks in trade costs and productivity. Formally, three objects are sufficient for conducting counterfactual exercises: (i) bilateral data on trade flows and export firm shares, (ii) the elasticity of substitution across varieties, and (iii) the gravity elasticity functions of the extensive and intensive margins of firm exports. Thus, the distribution of firm fundamentals only affects the model's aggregate counterfactual predictions through the shape and level of the two gravity elasticity functions.

We provide theoretical results that synthesize the debate on the aggregate implications of firm heterogeneity. For small shocks, the aggregate trade elasticity is sufficient to characterize the response of aggregate outcomes, in line with Arkolakis et al. (2012). For large shocks, the trade elasticity changes along the adjustment path to the new equilibrium due to changes in exporter firm shares, implying that firm heterogeneity plays an important role for the model's predictions, in line with Melitz and Redding (2015) and Feenstra (2018). We show, however, that this role is summarized by the gravity elasticities of the firm export margins and their

[2] We derive these semiparametric gravity equations using the same type of inversion argument in Berry and Haile (2014) and Adão (2015). The function of the exporter firm share in the intensive margin equation captures composition effects in exports generated by endogenous firm entry: average firm exports depend on the average sales cost of the endogenous pool of exporter firms.

dependence on the firm exporter share. When both elasticity functions are constant, the aggregate trade elasticity is also constant and the model's predictions are isomorphic to those of the constant-elasticity gravity trade models in Arkolakis et al. (2012).

We then derive expressions to nonparametrically recover changes in economic fundamentals from observed data. Given the two gravity elasticity functions, the inversion of changes in bilateral (variable and fixed) trade costs only requires data on changes in trade flows, firm exporter shares, and wages. This result does not hinge on the type of symmetry assumption imposed by Head and Ries (2001). The inversion of domestic productivity shifters also requires observing changes in the country's price index, as in Eaton, Kortum, Neiman and Romalis (2016).

We conclude our theoretical analysis by deriving nonparametric expressions for the impact of trade shocks on welfare that extend sufficient statistics in the literature – e.g., Melitz (2003), Arkolakis et al. (2012) and Melitz and Redding (2015). In particular, we derive an extension of the formula in Arkolakis et al. (2012) for small shocks in which welfare gains combine the initial trade elasticity, and changes in the domestic spending share and number of entrants. By integrating this formula for large shocks, we show that welfare gains must account for the correlation between changes in the trade elasticity and changes in domestic spending share and firm entry, which depends on the two elasticity functions of firm exports.

Our theoretical results indicate that different parametric assumptions in the literature only affect the model's aggregate predictions through what they imply for the shape and level of the gravity elasticity functions of the firm export margins. Moreover, such parametric assumptions play the key role of allowing the identification of the export margin elasticities from micro cross-sectional variation in observed firm outcomes.[3] Without parametric distributional assumptions, cross-firm variation in sales and exports does not identify the two functions necessary for counterfactual analysis. Instead, we argue that, in a more general environment without parametric restrictions in the distribution of firm fundamental, one should directly estimate the gravity elasticity functions of the firm export margins.

We do so by extending standard tools for gravity estimation to recover the two sufficient elasticity functions using the model-implied semiparametric gravity equations of firm exports. We exploit cross-country variation in exporter firm share and average firm exports induced by observed shifters of bilateral trade costs (conditional on origin and destination fixed-effects). This requires two main assumptions. First, different country groups must have common trade

[3]For instance, the elasticities of all margins do not vary with the exporter firm share if firm productivity has the Pareto distribution in Chaney (2008). The elasticities are decreasing in the exporter firm share if productivity has the Truncated Pareto distribution in Melitz and Redding (2015) or the Log-normal distribution in Head et al. (2014). These parametric assumptions allow the identification of the elasticity functions and the extrapolation of counterfactual predictions from cross-firm variation in sales and exports.

elasticity functions. Second, observed cost shifters must satisfy the same set of assumptions necessary for the consistent estimation of constant-elasticity gravity models – for a review, see Head and Mayer (2014). Intuitively, our methodology extends that of standard constant-elasticity gravity estimation by specifying the elasticity of each export margin to be a flexible function of the observed exporter firm share.

Our estimates yield three main insights about the impact of trade costs on the different export margins.[4] First, the extensive margin of firm entry is more responsive when few firms serve a market (in line with Kehoe and Ruhl (2013) and Kehoe, Rossbach and Ruhl (2015)). We estimate an extensive margin elasticity of six for low levels of the exporter firm share, but this elasticity is only two if more than 10% of the country's firms export to a destination. Second, we estimate an intensive margin elasticity that is positive (as in Fernandes et al. (2019)) but not sensitive to the number of exporter firms. As in Melitz (2003), our intensive margin elasticity entails strong composition effects: an increase of 1% in the number of exporters is associated with a reduction of 0.2% in average firm exports due to the higher cost of marginal exporters. Lastly, the combination of these two margins yields an aggregate trade elasticity that declines with the firm exporter share, varying between seven and four.

We conclude by studying the welfare consequences of two types of trade shocks in 2012: (i) moving from the current equilibrium to autarky by setting trade costs to infinity, and (ii) reverting trade costs within the European single market to its level of 2003. We measure the importance of firm heterogeneity for the shock's impact by comparing results implied by our baseline estimates of the gravity elasticity functions and those implied by the benchmark constant-elasticity gravity model of bilateral trade flows.[5]

In our first exercise, the welfare impact of moving to autarky obtained with our baseline estimates differs on absolute value by an average of 7% from those implied by the constant-elasticity gravity benchmark. Firm heterogeneity has a substantial impact on some countries in which our baseline gains differ by more than 25% from the gains in the benchmark model. We find that, because of the non-linear trade elasticities, our baseline gains are systematically larger than those in the benchmark for countries with a higher fraction of firms exporting.

Our second exercise quantifies the welfare impact of changes in trade costs within the European single market from 2003 to 2012. Using our theoretical inversion result, we use changes in observed outcomes to recover actual changes in fixed and variable costs of selling in different markets. We find that the impact of such shocks was highly heterogeneous across

[4] We use a sample of exporter-importer pairs in 2012 covering 88% of world trade. Following the gravity estimation literature, our trade cost shifter is bilateral distance.

[5] As discussed above, firm heterogeneity does not matter in this benchmark since it is isomorphic to the gravity trade models in Arkolakis et al. (2012). So, counterfactual predictions only require the aggregate constant trade elasticity and the initial trade flow matrix.

countries. While the average welfare gain was only 0.5% in Western Europe, it was 3.4% in East Europe (Estonia and Hungary both saw gains of over 5%). Isolating the impact of different shocks, we show that changes in fixed costs had only a small welfare impact. Most of the gains came from reductions in variable costs, with the larger gains in East Europe driven by stronger reductions in their export costs.[6] Compared to the constant-elasticity gravity benchmark, the average welfare impact of European market integration is 8% higher with our baseline estimates. This is because trade shocks affect several country pairs whose exporter firm shares yield a trade elasticity above the benchmark constant trade elasticity.

Our paper is related to an extensive literature using variations of the framework in Melitz (2003) together with parametric distributional assumptions to conduct empirical and counterfactual analyses – for a review, see Melitz and Redding (2014). We complement this literature by showing how to introduce firm heterogeneity in such analyses while dispensing parametric restrictions on the distribution of firm fundamentals. Our analysis indicates that the role of firm heterogeneity is summarized by two trade elasticity functions that are identified from semiparametric gravity equations for the margins of firm exports.[7] Parametric assumptions made in the literature are central to extrapolate counterfactual predictions from micro moments associated with cross-firm heterogeneity. Our empirical results show that such an approach may have a large quantitative impact on the contribution of firm heterogeneity for the model's predictions as it effectively restricts the two elasticity functions.

Our empirical analysis relies on two semiparametric gravity equations of firm exports. It is thus related to the literature estimating extensions of the log-linear gravity equation of bilateral trade flows – e.g., Novy (2013), Fajgelbaum and Khandelwal (2016), and Lind and Ramondo (2018). We also show that semiparametric gravity equations arise when we extend our baseline environment to allow for (i) tariffs, multiple sectors, multiple factors, and input-output links as in Costinot and Rodriguez-Clare (2013) and Caliendo and Parro (2014), (ii) zero bilateral trade flows as in Helpman et al. (2008), (iii) multi-product firms as in Bernard, Redding and Schott (2011) and Arkolakis, Ganapati and Muendler (2019b), and (iv) non-CES demand functions with a single price aggregator as in Matsuyama and Ushchev (2017) and Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2019a).

We also contribute to a recent literature focusing on nonparametric counterfactual analysis in international trade models (Adao, Costinot and Donaldson, 2017; Bartelme, Lan and Levchenko, 2019). Counterfactual predictions in these settings require knowledge of

[6]The impact of our baseline inverted shocks is larger and more heterogeneous than the impact of either observed changes in tariffs or symmetric trade shocks inverted with a nonparametric extension of the procedure in Head and Ries (2001).

[7]Fernandes et al. (2019) show that counterfactual analysis can be conducted for any given parametric distribution of firm productivity, but its implementation still requires the distribution to be specified.

multivariate functions whose nonparametric estimation is challenging in finite samples – for example, Adao et al. (2017) must estimate a country's demand function for all factors in the world economy. Compared to these papers, we show how to construct nonparametric counterfactuals in a different class of models featuring monopolistic competition and increasing returns to scale. Moreover, it is simple to implement our methodology in finite samples since it only requires estimating two univariate functions of the exporter firm share using semiparametric gravity equations. Further work that offers sufficient statistics for welfare and trade in general environments requires parametric assumptions for the actual implementation of empirical and counterfactual analyses (e.g. Baqaee and Farhi (2019); Kleinman, Liu and Redding (2020); Adao, Arkolakis and Esposito (2020)).

Finally, we offer formal results on how to recover changes in economic fundamentals that can be used in counterfactual analysis. These results constitute a nonparametric extension of the identification procedure in Eaton et al. (2016) in which we recover (potentially asymmetric) changes in fixed and variable trade costs, as well as changes in productivity. This is different from the focus on symmetric bilateral variable trade shifters, as in the approach of Adao et al. (2017), or productivity and amenity shocks (given bilateral trade shifters), as for example in the spatial models in Allen and Arkolakis (2014) and Redding (2016).

Our paper is organized as follows. Section 2 derives the semiparametric gravity equations for the extensive and intensive margins of firm exports. In Section 3, we present our nonparametric counterfactual analysis and inversion of economic fundamentals. Section 4 outlines the methodology to estimate the two main elasticity functions in the model. In Section 5, we report our baseline estimation results. Section 6 conducts counterfactual exercises. Section 7 concludes.

2 Model

We consider an economy in which firms are heterogeneous in terms of productivity, demand, and trade (variable and fixed) costs. The equilibrium of this economy entails two semi-parametric gravity equations for the extensive and intensive margins of firm exports. In general equilibrium, the functions in these two gravity equations along with country-level fundamentals determine trade flows, firm entry, price indices, and wages.

2.1 Environment

Demand. Each country j has a representative household that inelastically supplies $\bar{L}_j$ units of labor. In each country j, the representative household has Constant Elasticity of

Substitution (CES) preferences over the continuum of available varieties from different origins i, $\omega \in \Omega_{ij}$, with elasticity of substitution $\sigma > 1$. This yields j's demand for variety ω from country i:

$$q_{ij}(\omega) = \left(\bar{b}_{ij} b_{ij}(\omega)\right) \left(\frac{p_{ij}(\omega)}{P_j}\right)^{-\sigma} \frac{E_j}{P_j}, \tag{1}$$

where, in market j, E_j is the total spending, $p_{ij}(\omega)$ is the price of variety ω of country i, and P_j is the CES price index,

$$P_j^{1-\sigma} = \sum_i \int_{\Omega_{ij}} \left(\bar{b}_{ij} b_{ij}(\omega)\right) \left(p_{ij}(\omega)\right)^{1-\sigma} \, d\omega. \tag{2}$$

The demand shifters, $b_{ij}(\omega)$, generate dispersion in sales across varieties conditional on prices. The term $\bar{b}_{ij}$ is the component of bilateral taste shifters that is common to all varieties.

Production. Each variety is produced by a single firm, so we refer to a variety as a firm-specific good. The production function implies that, in order to sell q units in country j, firm ω from country i incurs in a labor cost of

$$C_{ij}(\omega, q) = w_i \frac{\tau_{ij}(\omega)}{a_i(\omega)} \frac{\bar{\tau}_{ij}}{\bar{a}_i} q + w_i \bar{f}_{ij} f_{ij}(\omega).$$

The first term is the variable cost of selling q units in country j, including both firm-specific iceberg shipping costs, $\bar{\tau}_{ij}\tau_{ij}(\omega)$, and productivity, $\bar{a}_i a_i(\omega)$. The second term is the fixed cost of labor necessary to enter j. As in Melitz (2003), we specify the fixed entry cost in terms of labor in the origin country. However, we depart from Melitz (2003) by allowing firms to be different not only in their productivity, but also in their variable and fixed costs of exporting. Eaton et al. (2011) show that these additional sources of firm heterogeneity are important to generate observed patterns of firm-level exports to different countries.

We consider a monopolistic competitive environment in which firms maximize profits given the demand in (1). For firm ω of country i, the optimal price in market j is $p_{ij}(\omega) = \frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij} w_i}{\bar{a}_i} \frac{\tau_{ij}(\omega)}{a_i(\omega)}$ with an associated revenue of

$$R_{ij}(\omega) = \bar{r}_{ij} r_{ij}(\omega) \left[\left(\frac{w_i}{P_j}\right)^{1-\sigma} E_j\right] \qquad \text{where} \tag{3}$$

$$r_{ij}(\omega) \equiv b_{ij}(\omega) \left(\frac{\tau_{ij}(\omega)}{a_i(\omega)}\right)^{1-\sigma} \quad \text{and} \quad \bar{r}_{ij} \equiv \bar{b}_{ij} \left(\frac{\sigma}{\sigma - 1} \frac{\bar{\tau}_{ij}}{\bar{a}_i}\right)^{1-\sigma}. \tag{4}$$

We refer to $r_{ij}(\omega)$ as the *revenue potential* in j of firm ω from i and $\bar{r}_{ij}$ as the bilateral revenue shifter that is common to all firms. Conditional on entering market j, $r_{ij}(\omega)$ is the ω-specific sales shifter in j that combines different sources of firm heterogeneity.

The firm's entry decision depends on the profit generated by the revenue in (3), $(1/\sigma)R_{ij}(\omega)$, and the fixed-cost of entry, $w_i\bar{f}_{ij}f_{ij}(\omega)$. Specifically, firm ω of i enters j if, and only if, $\pi_{ij}(\omega) = \frac{1}{\sigma}R_{ij}(\omega) - w_i\bar{f}_{ij}f_{ij}(\omega) \geq 0$. This yields the set of firms from i selling in j, Ω_{ij}:

$$\omega \in \Omega_{ij} \quad \Leftrightarrow e_{ij}(\omega) \equiv \frac{r_{ij}(\omega)}{f_{ij}(\omega)} \geq \sigma \frac{\bar{f}_{ij}}{\bar{r}_{ij}} \left[\left(\frac{w_i}{P_j} \right)^{\sigma} \frac{P_j}{E_j} \right]. \tag{5}$$

We refer to $e_{ij}(\omega)$ as the *entry potential* of firm ω of i in j. Among firms with identical revenue potential, heterogeneity in the fixed-cost of entry generates heterogeneity in entry potentials and, therefore, in decisions to enter different destination markets. The difference between revenue and entry potentials of firms allows for imperfect cross-firm correlation between entry and sales across markets.

General Equilibrium. Firms in country i hire $\bar{F}_i$ units of domestic labor to create a new variety whose characteristics are a random draw from an arbitrary distribution:

$$v_i(\omega) \equiv \{a_i(\omega), b_{ij}(\omega), \tau_{ij}(\omega), f_{ij}(\omega)\}_j \sim G_i(v). \tag{6}$$

In the free entry equilibrium, N_i enter and receive an expected profit of zero,

$$\sum_j E\left[\max\{\pi_{ij}(\omega);\ 0\}\right] = w_i\bar{F}_i. \tag{7}$$

As in Dekle et al. (2008), we allow for exogenous international transfers $\{\bar{T}_i\}_i$, so that

$$E_i = w_i\bar{L}_i + \bar{T}_i, \quad \sum_i \bar{T}_i = 0. \tag{8}$$

Since labor is the only factor of production, labor income in i equals the total revenue of firms from i: $w_iL_i = \int_{\omega\in\Omega_{ij}} R_{ij}(\omega)\,d\omega$. Given the expression in (3),

$$w_i\bar{L}_i = \bar{r}_{ij}\left(\frac{w_i}{P_j}\right)^{1-\sigma} E_j\left[\int_{\omega\in\Omega_{ij}} r_{ij}(\omega)\,d\omega\right]. \tag{9}$$

Summarizing, given the arbitrary distribution in (6), the equilibrium is defined as the vector $\{P_i, \{\Omega_{ij}\}_j, N_i, E_i, w_i\}_i$ satisfying equations (2), (5), (7), (8), (9) for all i.

2.2 The Extensive and Intensive Margins of Firm Exports

We now use two aggregate bilateral outcomes, the share of firms from i selling in j, $n_{ij} = Pr\left[\omega \in \Omega_{ij}\right]$, and their average sales, $\bar{x}_{ij} \equiv E\left[R_{ij}\left(\omega\right)|\omega \in \Omega_{ij}\right]$, to define two corresponding aggregate elasticity functions that control how the intensive and extensive margins of firm-level exports respond to changes in bilateral trade costs. In the next section, we use these elasticity functions in our sufficient statistics characterization of the counterfactual impact of trade shocks on aggregate outcomes. In the rest of the paper, we refer to n_{ij} and $\bar{x}_{ij}$ as exporter firm share and average firm exports.

We consider the CDF of $(r_{ij}(\omega), e_{ij}(\omega))$ implied by $G_i(v)$. Without loss of generality, this CDF can be decomposed as

$$r_{ij}(\omega) \sim H^r_{ij}\left(r|e\right), \qquad \text{and} \qquad e_{ij}(\omega) \sim H^e_{ij}(e). \tag{10}$$

Intuitively, firms draw their entry potential e from $H^e_{ij}(e)$. Conditional on having an entry potential of e, firms draw their revenue potential r from $H^r_{ij}\left(r|e\right)$. We impose the following regularity condition on the distribution of entry potentials.

Assumption 1. *Assume that $H^e_{ij}(e)$ is continuous and strictly increasing in $\mathbb{R}_+$ with* $\lim_{e\to\infty} H^e_{ij}(e) = 1$.

This assumption implies that that H^e_{ij} has no mass points, which guarantees that any trade cost change induces a positive mass of firms to switch entry decisions. This is central for the change of variables necessary for our characterization of the equilibrium below.[8]

Our specification allows for any pattern of heterogeneity and correlation in $(r_{ij}(\omega), e_{ij}(\omega))$. It thus encompasses several distributional assumptions in the literature. For instance, in Melitz (2003), the only source of firm heterogeneity is productivity such that $r_{ij}(\omega) = e_{ij}(\omega) = \left(a_i(\omega)\right)^{\sigma-1}$. The single source of heterogeneity implies a strict hierarchy of entry across destinations and a perfect cross-firm correlation between intensive and extensive margins of export. The distribution of e_{ij} can be specified to be Pareto, as in Chaney (2008) and Arkolakis (2010), truncated Pareto, as in Helpman et al. (2008) and Melitz and Redding (2015), or log-normal, as in Head et al. (2014) and Bas et al. (2017). Multiple papers incorporate additional sources of heterogeneity across firms that yield dispersion in both $r_{ij}(\omega)$ and $e_{ij}(\omega)$. For example, the demand and entry cost heterogeneity in Eaton et al. (2011) are modeled so that $a_i(\omega)$ is Pareto distributed while $b_{ij}(\omega)$ and $f_{ij}(\omega)$ are joint log-normally distributed while Fernandes et al. (2019) assume a multivariate log-normal

[8] The assumption of no upper bound on the support of e also implies positive trade flows between all origin-destination pairs. However, this is not essential. In Section 3.4, we allow for the possibility of zero bilateral trade by imposing that there exists $\bar{e}_{ij} < \infty$ such that $H^e_{ij}(\bar{e}_{ij}) = 1$.

distribution of bilateral sale shifters across destinations. Arkolakis et al. (2019b) extend these setups by considering a further layer of product-firm heterogeneity.

We now use Assumption 1 to characterize the extensive and intensive margins of firm exports. Focusing on the extensive margin first, we define the inverse distribution of entry potential: $\epsilon_{ij}(n) \equiv \left(H^e_{ij}\right)^{-1}(1-n)$ where $\bar{\epsilon}_{ij}(n)$ is strictly decreasing, $\epsilon_{ij}(1) = 0$, and $\lim_{n\to 0} \epsilon_{ij}(n) = \infty$. Applying this definition to (5),

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left(\sigma \bar{f}_{ij}/\bar{r}_{ij}\right) + \ln\left(w_i^{\sigma}\right) - \ln\left(E_j P_j^{\sigma-1}\right). \tag{11}$$

This expression is a semiparametric gravity equation since it relates the function $\epsilon_{ij}(n_{ij})$ to a log-linear combination of exogenous bilateral trade shifters and endogenous outcomes in the origin and destination markets. Notice that its elasticity, $\varepsilon_{ij}(n_{ij}) \equiv \left.\frac{\partial \ln \epsilon_{ij}(n)}{\partial \ln n}\right|_{n=n_{ij}}$, controls the sensitivity of the exporter firm share to changes in bilateral trade costs (holding constant other endogenous variables),

$$\frac{\partial \ln n_{ij}}{\partial \ln \bar{\tau}_{ij}} = \frac{\sigma - 1}{\varepsilon_{ij}(n_{ij})} < 0. \tag{12}$$

Turning to the intensive margin, we define the average revenue potential when a share n_{ij} of $i's$ firms sell in j as

$$\rho_{ij}(n_{ij}) \equiv \frac{1}{n_{ij}} \int_0^{n_{ij}} E[r|e = \epsilon_{ij}(n)]\; dn \tag{13}$$

where $E[r|e = \bar{\epsilon}_{ij}(n)]$ is the mean revenue potential of firms in quantile n of the entry potential distribution. Average bilateral firm sales can be written as

$$\bar{x}_{ij} = \bar{r}_{ij}\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma} E_j\right] \int_{e^*_{ij}}^{\infty} E\left[r|e\right] \frac{dH^e(e)}{1 - H^e(e^*_{ij})}, \quad \text{where} \quad e^*_{ij} \equiv \sigma \frac{\bar{f}_{ij}}{\bar{r}_{ij}} \left[\left(\frac{w_i}{P_j}\right)^{\sigma} \frac{P_j}{E_j}\right].$$

We now implement a change of variables in this expression by defining $n = 1 - H^e_{ij}(e)$ such that $n_{ij} = 1 - H^e_{ij}(e^*_{ij})$. This yields our second semiparametric gravity equation:

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln\left(\bar{r}_{ij}\right) + \ln\left(w_i^{1-\sigma}\right) + \ln\left(E_j P_j^{\sigma-1}\right), \tag{14}$$

The expression relates the composition-adjusted average firm exports to a linear combination of exogenous bilateral trade shifters and endogenous outcomes in the origin and destination markets. The elasticity of $\rho_{ij}(n)$, $\varrho_{ij}(n_{ij}) \equiv \left.\frac{\partial \ln \rho_{ij}(n)}{\partial \ln n_{ij}}\right|_{n=n_{ij}}$, controls the sensitivity of average firm exports to changes in bilateral trade costs (holding constant other endogenous variables):

$$\frac{\partial \ln \bar{x}_{ij}}{\partial \ln \bar{\tau}_{ij}} = -(\sigma - 1) + \varrho_{ij}(n_{ij}) \frac{\partial \ln n_{ij}}{\partial \ln \bar{\tau}_{ij}}. \tag{15}$$

This elasticity combines two well-known forces. The first term is the reduction in the sales of the initial set of exporters in j arising from the constant elasticity of substitution across varieties. The second term measures how the change in the number of exporters affects the average revenue potential of firms selling in j. The sign of this term depends on how different marginal and infra-marginal exporters are in terms of revenue potential. Specifically,

$$\varrho_{ij}(n_{ij}) = \frac{E[r|e = \epsilon_{ij}(n_{ij})]}{\frac{1}{n_{ij}} \int_0^{n_{ij}} E[r|e = \epsilon_{ij}(n)]\ dn} - 1. \tag{16}$$

Notice that $\varrho_{ij}(n_{ij}) < 0$ if, and only if, the mean revenue potential of marginal exporters, $E[r|e = \epsilon_{ij}(n_{ij})]$, is lower than that of the average exporter, $\rho_{ij}(n_{ij}) = \frac{1}{n_{ij}} \int_0^{n_{ij}} E[r|e = \epsilon_{ij}(n)]\ dn$.[9]

The bilateral trade flow between i and j, X_{ij}, combines the extensive and intensive margins of firm exports, $X_{ij} = N_i n_{ij} \bar{x}_{ij}$. Thus, by definition, $\epsilon_{ij}(n_{ij})$ and $\rho_{ij}(n_{ij})$ determine the elasticity of bilateral trade flows to changes in bilateral trade costs (holding constant endogenous variables in the origin and destination countries):

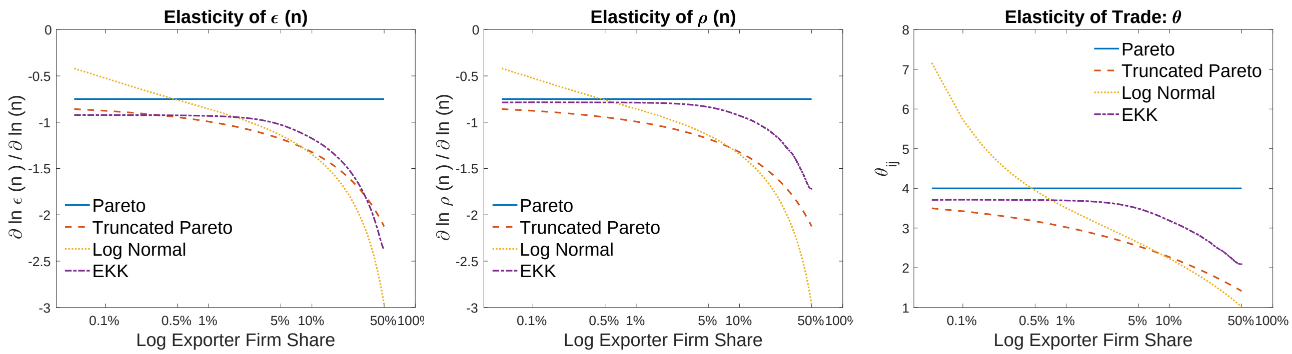
$$\theta_{ij}(n_{ij}) \equiv -\frac{\partial \ln X_{ij}}{\partial \ln \bar{\tau}_{ij}} = -\left(\frac{\partial \ln n_{ij}}{\partial \ln \bar{\tau}_{ij}} + \frac{\partial \ln \bar{x}_{ij}}{\partial \ln \bar{\tau}_{ij}} \right) = (\sigma - 1) \left(1 - \frac{1 + \varrho_{ij}(n_{ij})}{\varepsilon_{ij}(n_{ij})} \right). \tag{17}$$

This expression indicates that n_{ij} acts like a state variable that determines the elasticity of bilateral trade flows to changes in trade costs – the so-called trade elasticity. This occurs because n_{ij} controls the elasticity of the extensive and intensive margins of firm exports. Notice that the trade elasticity is positive for all n_{ij} since $\varepsilon_{ij}(n_{ij}) < 0$ and $\varrho_{ij}(n_{ij}) > -1$.

Distributional assumptions and elasticity of trade flows. Overall, one of our key findings is that the extensive and intensive margin elasticities, and thus the bilateral trade elasticity, are univariate functions of the exporter firm share, n_{ij}. We do not impose restrictions on these functions other than the regularity conditions in Assumption 1. Before characterizing the non-parametric sufficient statistics in our model, we show how different parametric assumptions on the distribution of firm fundamentals yield specific patterns of dependence between the extensive and intensive margin elasticities of firm exports and the exporter firm share. In particular, Figure 1 illustrates the elasticity functions implied by productivity distributions from the Pareto family (Chaney, 2008), the truncated Pareto family (Melitz and Redding, 2015), and the log-normal family (Head et al., 2014). We also

[9] In Melitz (2003), the single source of firm heterogeneity ($r_{ij}(\omega) = e_{ij}(\omega)$) implies that $E[r|e = \epsilon_{ij}(n_{ij})] = \epsilon_{ij}(n_{ij})$. In this case, marginal exporters are worse than existing infra-marginal exporters since $\frac{\partial \rho_{ij}(n)}{\partial n} = \frac{1}{n^2} \int_0^n (\epsilon_{ij}(n) - \epsilon_{ij}(x))\, dx < 0$ and $\epsilon_{ij}(n) < \epsilon_{ij}(x)$ for all $x < n$.

Figure 1: Distributional assumptions and Elasticity of different margins of trade flows



Note. Left panel reports the elasticity of $\epsilon_{ij}(n)$. Center panel reports the elasticity of $\rho_{ij}(n)$. Right panel reports the trade elasticity $\theta_{ij}(n)$ in (17). Each line corresponds to the elasticity as a function of n implied by different parametric restrictions on the distribution of firm fundamentals. See main text for a description of each parametrization.

consider the specification in Eaton et al. (2011) where productivity has a Pareto distribution and shifters of demand and entry costs have a joint log-normal distribution. In all cases, we use the baseline parameters reported in each paper.

The first plot indicates that the Pareto assumption yields constant elasticities of all margins. The other parameterizations yield a declining elasticity of $\epsilon(n)$, which, by equation (12), implies that the extensive margin elasticity is more sensitive when the exporter firm share is low. Similarly, all other parameterizations yield a declining elasticity of $\rho(n)$, indicating that new entrants and incumbents are more similar to each other when n_{ij} is small. This implies that composition effects are weaker when few firms export to a particular destination. The third panel combines these two margins to show that the trade elasticity is higher when n_{ij} is low. In all parametrizations, the trade elasticity falls below two when n_{ij} is above 50%. We show below that this implies a low elasticity in the domestic market where n_{ii} is high, which has important implications for the measurement of the gains from trade.

2.3 Sufficient Statistics of Firm Heterogeneity

Having characterized individual firm decisions and aggregate trade flows as a function of $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$, the following lemma establishes that all aggregate variables in equilibrium only depend on firm heterogeneity through $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Specifically, we show in Appendix A.1 that the equilibrium equations determining price indices, number of entrants, expenditures and wages can be expressed in terms of model fundamentals $\{\bar{T}_i, \bar{L}_i, \bar{F}_i, \bar{f}_{ij}, \bar{r}_{ij}\}$, the elasticity of substitution σ, and the elasticity functions $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_{i,j}$.

Lemma 1. *Suppose Assumption 1 holds. The equilibrium vector* $\{n_{ij}, \bar{x}_{ij}, X_{ij}, P_i, N_i, E_i, w_i\}_{i,j}$ *can be fully characterized as a function of (i) country fundamentals* $\{\bar{T}_i, \bar{L}_i, \bar{F}_i, \bar{f}_{ij}, \bar{r}_{ij}\}_{i,j}$*, (ii) the elasticity of substitution* σ*, and (iii) the bilateral elasticity functions,* $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_{i,j}$.

Proof. See Appendix A.1.

The intuition for this lemma follows from the equilibrium characterization in Melitz (2003). We first show that all aggregate variables depend on the set of firms operating in each country pair and their average sales, which can be written as functions of the equilibrium entry cutoff. We then use the same inversion argument used in Section 2.2 to establish a one-to-one mapping between these functions of the entry cutoff and $(\epsilon_{ij}(n), \rho_{ij}(n))$.

The results of this section indicate that the functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$ summarize the aggregate implications of the different dimensions of firm heterogeneity in the model. Thus, any parametric restriction on the distribution of firm fundamentals affects the economy's equilibrium insofar it determines the shape of these functions. We summarize this discussion in the following remark.

Remark 1. *All dimensions of heterogeneity can be folded into the functions* $\epsilon_{ij}(n)$ *and* $\rho_{ij}(n)$ *that govern the semiparametric gravity equations for the extensive and intensive margins of firm exports, (11) and (14), and all aggregate variables in general equilibrium.*

The rest of the paper exploits this insight in two ways. In the next section, we build directly on Lemma 1 by using $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_{i,j}$ to provide ex-ante and ex-post sufficient statistics for the counterfactual impact of trade shocks on aggregate outcomes and welfare given bilateral trade and exporter entry data. We also show how to invert changes in the economy's fundamentals given knowledge of $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_{i,j}$ and observed changes in bilateral trade and exporter entry. In Section 4, we will exploit the semiparametric gravity equations in (11) and (14) to estimate $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$ without imposing any parametric restrictions on the distribution of firm fundamentals.

3 Nonparametric Counterfactual Analysis and Identification of Economic Fundamentals

This section uses the equilibrium characterization above for counterfactual analysis and identification of changes in fundamentals without imposing parametric restrictions on the distribution of firm fundamentals. We first show how to construct nonparametric counterfactual changes in aggregate outcomes in response to changes in fundamentals using observed trade data in the initial equilibrium and $(\sigma, \epsilon_{ij}(n), \rho_{ij}(n))$. We then discuss how to nonparametrically identify shocks to economic fundamentals from changes in observed outcomes given knowledge of $(\sigma, \epsilon_{ij}(n), \rho_{ij}(n))$. Finally, we use $(\sigma, \epsilon_{ij}(n), \rho_{ij}(n))$ and trade data to construct sufficient statistics for welfare gains implied by shocks in trade cost.

3.1 Counterfactual Responses to Changes in Fundamentals

We start by characterizing how to compute counterfactual changes in aggregate outcomes following exogenous shocks in the economy's fundamentals. We consider shocks in international transfers $\bar{T}_i$, population $\bar{L}_i$, entry costs $\bar{F}_i$, fixed costs of exporting $\bar{f}_{ij}$, and bilateral revenue shifters $\bar{r}_{ij}$. We use $\hat{y}_j \equiv y'_j/y^0_j$ to express changes in any variable between its level in the initial observed equilibrium, y^0_i, and the counterfactual equilibrium, y'_i. We also use bold letters to denote vectors, $\boldsymbol{y} = [y_i]_i$, and bold bar variables to denote matrices, $\bar{\boldsymbol{y}} = [y_{ij}]_{i,j}$.

Proposition 1. *Consider any change in fundamentals, $\{\hat{\boldsymbol{T}}, \hat{\boldsymbol{L}}, \hat{\boldsymbol{F}}, \hat{\bar{\boldsymbol{f}}}, \hat{\bar{\boldsymbol{r}}}\}$. Given the matrices of exporter firm shares and bilateral trade flows in the initial equilibrium $(\bar{\boldsymbol{n}}^0, \bar{\boldsymbol{X}}^0)$, the substitution elasticity σ and the functions $(\bar{\boldsymbol{\epsilon}}(\bar{\boldsymbol{n}}), \bar{\boldsymbol{\rho}}(\bar{\boldsymbol{n}}))$ (up to a scalar) are "sufficient statistics" that characterize the counterfactual changes in equilibrium outcomes, $\{\hat{\bar{\boldsymbol{n}}}, \hat{\bar{\boldsymbol{x}}}, \hat{\bar{\boldsymbol{X}}}, \hat{\boldsymbol{P}}, \hat{\boldsymbol{N}}, \hat{\boldsymbol{w}}\}$.*

Proof. See Appendix A.2.

The proposition generalizes the sufficient statistics result of Arkolakis et al. (2012) beyond the class of constant-elasticity gravity models by exploiting the characterization of the equilibrium in Lemma 1. It outlines two sufficient requirements to compute counterfactual changes in aggregate outcomes. First, it is not necessary to know the entire distribution of firm fundamentals in the initial equilibrium. Instead, one only needs to obtain the exporter firm share n_{ij}, and the two functions controlling the extensive and intensive margins of firm exports, $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Second, it is also necessary to obtain bilateral trade flows between countries, as in the original "hat algebra" methodology in Dekle et al. (2008), and the elasticity of substitution σ, as in the "hat algebra" for heterogeneous firm models (see Costinot and Rodriguez-Clare (2013)).[10]

According to Proposition 1, the trade elasticity, $\theta_{ij}(n_{ij})$, does not play a direct role in the model's counterfactual predictions without parametric restriction on the distribution of firm fundamentals. However, as the next proposition shows, the trade elasticity is the channel through which the extensive and intensive margin elasticities of firm exports affect counterfactual changes in aggregate outcomes in response to trade cost shocks.

Proposition 2. *Consider a small shock to the bilateral revenue shifters between origin o and destination d, $d \ln \bar{r}_{od}$. Denote the vector of aggregate outcomes as $Y_i \equiv \{\{X_{ij}\}_j, P_i, N_i, w_i\}$.*

1. The elasticity of any element of Y_i to $\bar{r}_{od}$ is a function of $(\sigma, \bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0), \bar{\boldsymbol{X}}^0)$:

$$\frac{d \ln Y_i}{d \ln \bar{r}_{od}} = \Psi^z_{i,od}\left(\sigma, \bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0), \bar{\boldsymbol{X}}^0\right), \tag{18}$$

[10]The elasticity of substitution is necessary when the entry cost is set in terms of the origin country wage. In this case, gravity origin fixed-effects contain wages with an elasticity determined by σ.

where $\theta_{ij}(n)$ is the trade elasticity function defined in (17).

2. The elasticity of n_{ij} to $\bar{r}_{od}$ is a function of $(\sigma, \bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0), \bar{\boldsymbol{X}}^0)$ and $\varepsilon_{ij}(n_{ij}^0)$:

$$\frac{d\ln n_{ij}}{d\ln \bar{r}_{od}} = \tilde{\Psi}^z_{ij,od}\left(\sigma, \bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0), \bar{\boldsymbol{X}}^0, \varepsilon_{ij}(n_{ij}^0)\right). \tag{19}$$

Proof. See Appendix A.3.

The first part of the proposition establishes that the elasticity of aggregate outcomes, $\{\{X_{ij}\}_j, P_i, N_i, E_i, w_i\}$, to bilateral trade costs is a function of the initial aggregate trade matrix, $\bar{\boldsymbol{X}}^0$, the elasticity of substitution σ, and the bilateral trade elasticity matrix, $\bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0)$. Thus, in contrast to Proposition 1, the first part of Proposition 2 indicates that, at least for local responses, separate knowledge of the extensive and intensive margin elasticities –and thus the distribution of firm-specific fundamentals–is not required conditional on knowing the trade elasticity matrix, $\bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}}^0)$.[11] The second part of Proposition 2 explains exactly why computing counterfactual responses for large changes in parameters requires separate knowledge of $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. The extensive margin elasticity, $\varepsilon_{ij}(n_{ij}^0)$, pins down the impact of the shock on the exporter firm share, n_{ij}. Along the path to the new equilibrium, changes in n_{ij} then imply changes in the trade elasticity $\theta_{ij}(n_{ij})$ through both the intensive and extensive margins of firm exports – see equation (17). This determines the local responses of aggregate outcomes for any sequence of small shocks (as shown in the first part of the proposition).

Propositions 1 and 2 provide a synthesis of the results in Melitz and Redding (2015), who stress the importance of varying trade elasticities and firm entry, with the results of Arkolakis et al. (2012), who stress the sufficient role of the constant trade elasticity for responses in aggregate outcomes to trade shocks. According to Proposition 2, firm heterogeneity matters to the extent that it controls how much trade shocks affect the trade elasticity through (large) changes in the exporter firm share, ultimately regulated by the functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$.[12]

This point is illustrated by the special case in which the intensive and extensive margin elasticities are constant,

[11] In Appendix A.3, we show that the same data and elasticity requirements are sufficient to compute the elasticity of aggregate outcomes to changes in population $d\ln \bar{L}_o$ and trade imbaliances $d\ln \bar{T}_o$. Howerver, to compute elasticities to changes in fixed cost of exporting $d\ln \bar{f}_{od}$ and entry cost $d\ln \bar{F}_o$, we need also the initial share of the country's labor force employed to cover fixed costs of exporting, which depends on the functions $(\epsilon_{ij}(n), \rho_{ij}(n))$ and initial firm export shares $\{n_{ij}^0\}_{i,j}$ in equilibrium.

[12] It is also useful to compare Proposition 2 to the results of Arkolakis et al. (2019a) showing that knowledge of the trade elasticity is sufficient, locally, for nonparametric counterfactuals in the case of two symmetric countries. We show that their results hold locally, for many asymmetric countries, but they are not true when we consider large shocks.

$$\rho_{ij}(n) = n^{\varrho_{ij}} \quad \text{and} \quad \epsilon_{ij}(n) = n^{\varepsilon_{ij}}, \tag{20}$$

for $\varrho_{ij} > -1$ and $\varepsilon_{ij} < 0$. This specification is a flexible extension of the constant elasticity Pareto variant of Melitz (2003) studied by Chaney (2008).[13] In this case, Proposition 2 implies that all aggregate outcomes can be computed by integrating the local responses in (18) without tracking the changes in the exporter firm share n_{ij} (even for large shocks). Thus, firm heterogeneity does not matter for aggregate outcomes conditional on knowing the trade elasticity matrix. We summarize the conclusions of this section under the following remark.

Remark 2. *Firm heterogeneity only matters for counterfactual responses in aggregate outcomes to trade shocks through σ and the shape of $(\bar{\boldsymbol{\epsilon}}(\bar{\boldsymbol{n}}), \bar{\boldsymbol{\rho}}(\bar{\boldsymbol{n}}))$. For small trade cost shocks, $(\bar{\boldsymbol{\epsilon}}(\bar{\boldsymbol{n}}), \bar{\boldsymbol{\rho}}(\bar{\boldsymbol{n}}))$ matter only through their combined effect on the trade elasticity matrix $\bar{\boldsymbol{\theta}}(\bar{\boldsymbol{n}})$. If $(\bar{\boldsymbol{\epsilon}}(\bar{\boldsymbol{n}}), \bar{\boldsymbol{\rho}}(\bar{\boldsymbol{n}}))$ have constant elasticities, then trade elasticities θ_{ij} are constant and sufficient to compute counterfactual responses of aggregate outcomes to trade cost shocks.*

3.2 Identification of Changes in Fundamentals

In the preceding discussion, we established how to conduct counterfactual experiments for any given change in economic fundamentals, like changes in productivity and trade costs. An important challenge in conducting these counterfactuals is to measure actual changes in economic fundamentals hitting the world economy. We thus show how to identify such changes in fundamentals from observed aggregate data without imposing parametric restrictions on the distribution of firm fundamentals or symmetry restrictions on trade cost changes.

Given data on n_{ij}^0 and $\{\hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{w}_i^t\}$ for all i and j, we can invert changes in fixed costs of exporting and relative revenue shifters using $\{\sigma, \epsilon_{ij}(n), \rho_{ij}(n)\}$, along with data on bilateral sales, firm entry, and wages. In particular, equations (11) and (14) immediately yield

$$\hat{\bar{f}}_{ij}^t = \frac{\hat{\bar{x}}_{ij}^t}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)/\rho_{ij}(n_{ij}^0)} \frac{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij}^t)/\epsilon_{ij}(n_{ij}^0)}{\hat{w}_i^t}, \tag{21}$$

and equation (14) implies that

$$\frac{\hat{\bar{r}}_{ij}^t}{\hat{\bar{r}}_{jj}^t} = \frac{\hat{\bar{x}}_{ij}^t}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)/\rho_{ij}(n_{ij}^0)} \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}^t)/\rho_{jj}(n_{jj}^0)}{\hat{\bar{x}}_{jj}^t} \left(\frac{\hat{w}_i^t}{\hat{w}_j^t}\right)^{\sigma-1}. \tag{22}$$

[13]The model in Chaney (2008) entails $r_{ij}(\omega) = \epsilon_{ij}(\omega) = (a_i(\omega))^{\sigma-1}$ and $a_i(\omega) \sim 1 - a^{-\theta}$, leading to $\varrho_{ij} = \varepsilon_{ij} = (\sigma-1)/\theta$. This implies that average firm exports do not respond to changes in variable trade costs because $\ln \bar{x}_{ij}$ does not depend on $\ln \bar{r}_{ij}$.

Notice that we obtain these expressions without imposing the symmetry of trade cost changes used by Head and Ries (2001). We achieve this more general result at the cost of also requiring data on wage changes.[14] In Appendix A.4, we show that one can use the free entry condition to invert changes in fixed entry costs, $\hat{\hat{F}}_i$, using the same information.

Intuitively, equation (21) states that fixed entry costs must have increased whenever we observe that the composition-adjusted firm sales, $\frac{\hat{\hat{x}}_{ij}^t}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)/\rho_{ij}(n_{ij}^0)}$, has increased relative to the change in the entry cost of marginal firms, $\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij}^t)/\epsilon_{ij}(n_{ij}^0)\hat{w}_i^t$. Similarly, in market j, the revenue shifter of i must have increased more than that of j whenever the composition-adjusted sales of firms from i in j increase more than those of domestic firms.

We use the expressions in (21)–(22) to formally establish the identification of changes in economic fundamentals from observed changes in aggregate outcomes.

Proposition 3. *Consider matrices of exporter firm shares and bilateral trade flows in the initial equilibrium* $(\bar{\boldsymbol{n}}^0, \bar{\boldsymbol{X}}^0)$*, the substitution elasticity* σ*, and the elasticity functions* $(\bar{\boldsymbol{\epsilon}}(\bar{\boldsymbol{n}}), \bar{\boldsymbol{\rho}}(\bar{\boldsymbol{n}}))$*. Observed changes in* $\{\hat{\bar{\boldsymbol{n}}}, \hat{\hat{\boldsymbol{x}}}, \hat{\hat{\boldsymbol{X}}}, \hat{\boldsymbol{w}}\}$ *between two equilibria uniquely identify the set of shocks in fundamentals* $\{\hat{\boldsymbol{T}}, \hat{\boldsymbol{L}}, \hat{\boldsymbol{F}}, \hat{\hat{\boldsymbol{f}}}, \hat{\hat{\bar{\boldsymbol{r}}}}\}$ *with* $\hat{\hat{\bar{r}}}_{ij} = \hat{\hat{r}}_{ij}/\hat{\hat{r}}_{jj}$*. Observing also the change in the price index* $\hat{P}_j$ *in country* j *between the two equilibria uniquely identifies the domestic revenue shock* $\hat{\hat{r}}_{jj}$ *in country* j*.*

Proof. See Appendix A.4.

It is important to note that the full separation of location specific shocks (e.g. productivity gains) from bilateral revenue shocks (e.g. bilateral trade shocks) requires knowledge of price index changes (as the second part of the proposition shows). Without price index information, the two could not be separated because changes in relative output of countries can be generated by either a uniform change in exporting advantage or different combinations of changes in bilateral revenue shifters – a formal result that follows from the insights of Eaton et al. (2016).[15]

3.3 Sufficient Statistics for Welfare Gains

We finally derive expressions for welfare gains caused by shocks to fundamentals, $\{\hat{\boldsymbol{T}}, \hat{\boldsymbol{L}}, \hat{\boldsymbol{F}}, \hat{\hat{\boldsymbol{f}}}, \hat{\hat{\boldsymbol{r}}}\}$, under trade balance (i.e., $T_i = 0$). When $\hat{\hat{f}}_{ii} = \hat{\hat{r}}_{ii} = 1$, our expressions link real wage changes to observable variables and measurable functions, leading to sufficient statistics for the

[14]Our result differs from that in Adao et al. (2017) who use an invertible aggregate demand system to nonparametrically recover the effective bilateral prices in each destination – i.e., in our notation, $\hat{\hat{r}}_{ij}/\hat{\hat{r}}_{jj}$. We instead combine demand and supply equations to invert fixed costs and relative revenue shifters.

[15]See also Allen et al. (2014) and Costinot et al. (2010). Our identification result provides a formal treatment of that insight, extends it to cases where elasticities are not constant, and adapts it to monopolistic competition environments.

welfare consequences of trade shocks. We discuss here the main formulas and present their derivations in Appendix A.5.

We first express changes in the real wage in terms of the change in the share of domestic active firms, $\hat{n}_{ii}$, and the domestic elasticity function, $\epsilon_{ii}(n)$. From (11),

$$\ln\left(\frac{\hat{w}_i}{\hat{P}_i}\right) = \frac{1}{\sigma-1}\ln\left(\frac{\hat{\bar{r}}_{ii}}{\hat{\bar{f}}_{ii}}\right) + \frac{1}{\sigma-1}\ln\left(\frac{\epsilon_{ii}(n_{ii}\hat{n}_{ii})}{\epsilon_{ii}(n_{ii})}\right). \tag{23}$$

Since $\epsilon_{ii}(n)$ is decreasing and $\sigma > 1$, the real wage increases if, and only if, n_{ii} falls (i.e., $\hat{n}_{ii} < 1$). The second term in this expression illustrates the main new source of gains from trade in Melitz (2003): the consumption-equivalent gain of reallocating resources from domestic firms with a lower entry potential to firms with a higher entry potential. Given the change in the share of domestic active firms $\hat{n}_{ii}$, welfare gains are higher if $\bar{\epsilon}_{ii}(n)$ is more elastic. Intuitively, for a steeper $\epsilon_{ii}(n)$, the difference in entry potential between incumbent and marginal firms is larger, leading to larger reallocation gains.[16]

We obtain an alternative welfare formula by combining (11) and (14) to solve for the change in n_{ii} as a function of the changes in domestic spending share x_{ii} and number of entrants N_i. Locally,

$$d\ln\left(\frac{w_i}{P_i}\right) = \frac{1}{\sigma-1}d\ln\frac{\bar{r}_{ii}}{\bar{f}_{ii}} - \frac{1}{\theta_{ii}(n_{ii})}\left(d\ln\frac{x_{ii}}{N_i} - d\ln\bar{f}_{ii}\right), \tag{24}$$

which can be written in terms of $(\sigma, \rho_{ii}(n), \epsilon_{ii}(n))$ due to the definition of $\theta_{ii}(n)$ in (17).

This expression shows that, for any given $d\ln(x_{ii}/N_i)$, the real wage change is stronger whenever the domestic trade elasticity $\theta_{ii}(n_{ii})$ is lower. Intuitively, the lower trade elasticity implies that it is harder for the economy to substitute consumption from foreign varieties to domestic varieties (through both the extensive and the intensive margins). This amplifies the cost of reducing the spending share on foreign varieties.

This expression is closely related to the welfare formulas in Arkolakis et al. (2009) (footnote 17) and Melitz and Redding (2013) (equation 33). Our characterization shows that the trade elasticity is a function of the observable share of active firms n_{ii}. Thus, for large shocks, the computation of welfare gains must account for the correlation between changes in the domestic trade elasticity and the domestic spending share (normalized by the number of domestic entrants). Such a correlation arises from endogenous changes in the share of domestic active

[16]Expression (23) is related to the characterization of the gains from trade in terms of the domestic productivity cutoff in Melitz (2003). Notice however that such characterization lacks empirical analogs due to the lack of measures of the productivity cutoff. Instead, our formula expresses the gains from trade in terms of the change in the share of active domestic firms $\hat{n}_{ii}$ and the extensive margin elasticity function $\epsilon_{ii}(n)$. The next section shows that it is possible to obtain measures of both of these elements.

firms and its implied effect on the domestic trade elasticity.[17] Most importantly, perhaps, the domestic trade elasticity function in equation (24) has a precise empirical content, as we discuss in the next section.

Remark 3. *The elasticity of substitution, σ, and the domestic elasticity functions, $\epsilon_{ii}(n)$ and $\rho_{ii}(n)$, can be used to compute nonparametric sufficient statistics for the welfare gains implied by trade shocks.*

Gains from Trade. It is possible to use the welfare expressions above to compute the welfare gains of moving from autarky to the trade equilibrium. This requires solving for the counterfactual changes in n_{ii} and N_i. In Appendix A.6, we show that such changes are the solution of a nonlinear system of two equations and two unknowns. This system is a special case of the general system used to compute the nonparametric counterfactual changes in Proposition 1. It depends on three ingredients: (i) data on exporter firm shares and export flows of country i in the initial equilibrium $\{n_{ij}^0, X_{ij}^0\}_j$, (ii) the elasticity of substitution σ, and (iii) the two elasticity functions of firm exports for country i, $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_j$.

3.4 Extensions

Online Supplemental Material A presents five extensions of our baseline framework.

Multiple sectors, multiple factors, input-output links. We extend our baseline model to include multiple factors of production and input-output links between multiple sectors. Specifically, we extend the multi-sector multi-factor gravity model of Costinot and Rodriguez-Clare (2013) in which, as in our baseline, firms in each sector are heterogeneous with respect to productivity, preferences, and variable and fixed trade costs. We restrict all firms in a sector to have the same nested constant elasticity of substitution (CES) production technology that uses multiple factors and multiple sectoral composite goods. In this setting, we derive sector-specific analogs of (11) and (14) that can be used to perform nonparametric counterfactual analysis with respect to trade cost shocks.[18]

Allowing for zero bilateral trade flows. We extend our baseline framework to allow for zero trade flows between two countries. As in Helpman et al. (2008), we allow the support

[17] As shown in Appendix A.5, expression (24) reduces to the sufficient statistic for the gains from trade in Arkolakis et al. (2012) when the trade elasticity is restricted to be constant –as in our CES benchmark– and also identical across all destinations –as in the Melitz-Pareto baseline. This equivalence arises because N_i does not depend on trade costs in this case.

[18] This framework also accommodates nested CES preferences for the goods produced by firms with different observable characteristics – for instance, sector affiliation or country of origin.

of the entry potential distribution to be bounded: $H^e_{ij}(e)$ has full support over $[0, \bar{e}_{ij}]$. The bounded support does not affect the intensive margin gravity equation (14), but it introduces a censoring structure in the extensive margin equation (11). Under the assumption that zero trade flows remain equal to zero, we use these extended gravity equations to compute the model's counterfactual predictions.

Allowing for import tariffs. Third, we follow Costinot and Rodriguez-Clare (2013) and introduce bilateral import tariffs in the model. In this setting, market clearing and spending must account for the fact that tariff revenue remains in the destination country. We show that the semiparametric gravity equations above still hold, but now bilateral trade costs also include ad-valorem import tariffs. We then characterize the system of equations that determines the model's counterfactual predictions without parametric distributional assumptions. It depends on the same elements as before, with one addition, the tariff levels in the initial equilibrium.

Multi-product firms. We extend our model to allow firms to produce multiple products as in Bernard et al. (2011). As in the baseline, we allow for an arbitrary distribution of firm-specific fundamentals. In this extension however, firms face a convex labor cost of increasing the number of product varieties that they will sell in each destination (see e.g. Arkolakis et al. (2019b)). This yields three semiparametric gravity equations: two extended versions of the baseline expressions for the extensive and intensive margins of firm exports, and one additional equation for the average number of products sold per firm across destinations. We establish that the elasticity functions in these equations are sufficient to construct counterfactual predictions given the same information required by our baseline setting (i.e., initial trade flows and firm exporter shares, and the elasticity of substitution across varieties).

Non-CES Preferences. We adapt our framework to general Marshallian demand functions that can be written as a function of the destination's price aggregator and income level. Our demand system subsumes the settings in Arkolakis et al. (2019a) and Matsuyama and Ushchev (2017). In this case, we abstract from fixed entry costs and incorporate endogenous firm entry through a choke price in demand. We show how to extend our inversion argument to derive an extensive margin gravity equation analogous to the one in (11). Because revenue and entry potentials are identical, the same function in the extensive margin gravity equation determines the intensive margin of average firm exports.

4 Estimation Strategy: Semiparametric Gravity

In the previous section, we derived two main insights. First, the functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ are sufficient to summarize how firm heterogeneity affects counterfactual responses of aggregate outcomes to trade shocks. Second, these two functions control semiparametric gravity equations for the extensive and intensive margins of firm exports. This section outlines a strategy to estimate $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ using these gravity equations. We then use these estimates for counterfactual analysis.

4.1 Semiparametric Gravity Equations of Firm Exports

The gravity equations in (11) and (14) imply the following semiparametric specifications:

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left(\bar{f}_{ij}\bar{\tau}_{ij}^{\sigma-1}\right) + \tilde{\delta}_i^{\epsilon} + \tilde{\zeta}_j^{\epsilon}, \tag{25}$$

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln\left(\bar{\tau}_{ij}^{1-\sigma}\right) + \tilde{\delta}_i^{\rho} + \tilde{\zeta}_j^{\rho}. \tag{26}$$

where $\tilde{\delta}_i^{\epsilon} \equiv \ln(\sigma^{\sigma}(\sigma-1)^{1-\sigma}\bar{a}_i^{1-\sigma}w_i^{\sigma})$, $\tilde{\zeta}_j^{\epsilon} \equiv -\ln(E_j P_j^{\sigma-1})$, $\tilde{\delta}_i^{\rho} \equiv \ln(\sigma w_i) - \tilde{\delta}_i^{\epsilon}$, and $\tilde{\zeta}_j^{\rho} \equiv -\tilde{\zeta}_j^{\epsilon}$. Without loss of generality, we normalize $\bar{b}_{ij} \equiv 1$ since bilateral shifters of demand and trade costs are isomorphic in the model – i.e., the equilibrium only depend on $\bar{\tau}_{ij}^{1-\sigma}\bar{b}_{ij}$.

These two equations form the basis of our empirical strategy. They link average firm exports and the two functions of the exporter firm share to bilateral shifters of variable and fixed costs of exporting, as well as exporter and importer fixed-effects. We can then estimate $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ using these equations along with bilateral data on average firm exports, exporter firm shares, and trade cost shifters.

Remark 4. $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ *can be estimated with the semiparametric specifications (25)–(26).*

In the rest of this section, we first describe sufficient assumptions on the data generating process that allow us to estimate $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ using the semiparametric equations in (25) and (26). We then outline an estimator of $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ based on cross-country variation in bilateral trade cost shifters.

4.2 Data Generating Process

Our goal is to estimate the functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$. Throughout our analysis, we use estimates in the literature to calibrate the elasticity of substitution, σ. In particular, we set $\sigma = 3.9$ to match the median estimate in Hottman et al. (2016).

We start by describing the observed and unobserved variables in the economy. For a set of origin-destination pairs (i, j), we observe the share of firms of i selling in j, n_{ij}, and their average sales, $\bar{x}_{ij}$. We make assumptions on the relationship between the variable trade costs in the model and their observed counterparts (denoted by τ_{ij}). We also assume that there is an exogenous shifter of the variable and fixed trade costs (denoted by z_{ij}).

Assumption 2. *Assume that we observe a component of variable trade costs, τ_{ij}, such that*

$$\ln \bar{\tau}_{ij} = \ln \tau_{ij} + \eta^{u}_{ij}. \tag{27}$$

Assume also that there exists an observed bilateral trade shifter, z_{ij}, such that

$$\begin{aligned} \ln \bar{\tau}_{ij} &= z_{ij}\kappa^{\tau} + \delta^{\tau}_{i} + \zeta^{\tau}_{j} + \eta^{\tau}_{ij}, \\ \ln \bar{f}_{ij} &= z_{ij}\kappa^{f} + \delta^{f}_{i} + \zeta^{f}_{j} + \eta^{f}_{ij}. \end{aligned} \tag{28}$$

These equations are the first-stage of the estimation of the semiparametric gravity equations. They link variable and fixed trade costs to an observed shifter while accounting for the fact that we may not observe all components of trade costs.[19] Previewing our empirical application, we use data on bilateral freight costs to measure τ_{ij}, and data on bilateral distance to measure z_{ij}.

We further restrict the data generating process of trade shocks.

Assumption 3. *Assume that $E[\eta^{\tau}_{ij}|z_{ij}, D_{ij}] = E[\eta^{f}_{ij}|z_{ij}, D_{ij}] = E[\eta^{u}_{ij}|z_{ij}, D_{ij}] = 0$, where D_{ij} is a vector of origin and destination fixed-effects.*

This orthogonality assumption is the basis of the estimation of constant elasticity gravity equations of international trade flows – for a review, see Head and Mayer (2014). Conditional on origin and destination fixed-effects, the observed shifter must be mean independent from unobserved shifters of trade costs.

Finally, we impose the following restrictions on the functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ that reduce their dimensionality but allow for their flexible estimation.

Assumption 4. *Assume that origin-destination pairs are divided into groups $(g = 1, ..., G)$ such that, for all $(i, j) \in g$,*

$$\begin{bmatrix} \ln \rho_{ij}(n) \\ \ln \epsilon_{ij}(n) \end{bmatrix} = \begin{bmatrix} \ln \rho_{g}(n) \\ \ln \epsilon_{g}(n) \end{bmatrix} = \sum_{k=1}^{K} \begin{bmatrix} \gamma^{\rho}_{g,k} f_k(\ln n) \\ \gamma^{\epsilon}_{g,k} f_k(\ln n) \end{bmatrix} \tag{29}$$

[19] This specification allows z_{ij} to affect the fixed cost of entering foreign markets and, therefore, it is weaker than the requirement in Helpman et al. (2008) that the instrument cannot affect entry costs.

where $f_k(\ln n)$ denotes restricted cubic splines over knots $k = 1, ..., K$.

This assumption imposes two types of restrictions on $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$. First, these functions are identical among origin-destination pairs in the same group g. This allows us to estimate $\rho_g(n)$ and $\epsilon_g(n)$ using variation in the observed shifters of trade costs across origin-destination pairs at a point in time.[20] In our empirical application, we specify that all countries belong to a single group, so that $\epsilon_{ij}(n) = \epsilon(n)$ and $\rho_{ij}(n) = \rho(n)$ for all i and j. In Supplementary Material B, we provide estimates with multiple country groups defined in terms of characteristics of origin and destination countries.

Second, Assumption 4 specifies a flexible function basis for $\rho_g(n)$ and $\epsilon_g(n)$. We approximate the shape of these functions with a series of restricted cubic polynomials. We specify K knots that form intervals, $\mathcal{U}_k \equiv [u_k, u_{k+1}]$, over which a cubic spline function governs the behavior of the elasticity function – see Appendix B.2. We discuss below how our functional form choice affects the interpretation and generality of our estimates.

4.3 Estimating Moment Conditions

We now construct moment conditions for the estimation of $\rho_g(n)$ and $\epsilon_g(n)$.

Pass-through from observed shifter to variable trade cost. We first specify an equation for the estimation of the pass-through from the observed cost shifter z_{ij} to the observed component of variable trade costs τ_{ij}. Assumption 2, by combining equations (27) and (28), implies that

$$v^{\tau}_{ij} = \ln \tau_{ij} - z_{ij}\kappa^{\tau} - \delta^{\tau}_i - \zeta^{\tau}_j. \tag{30}$$

Here, $v^{\tau}_{ij} = \eta^{\tau}_{ij} - \eta^{u}_{ij}$, implying that $E[v^{\tau}_{ij}|z_{ij}, D_{ij}] = 0$ by Assumption 3. We exploit this condition to estimate κ^{τ} using the linear equation in (30).

Semiparametric gravity equations of firm exports. To estimate $\bar{\rho}_g(n)$ and $\bar{\epsilon}_g(n)$, we show in Appendix A.7 that, under Assumptions 2 and 4, equations (25)–(26) yield

$$\begin{bmatrix} v^{\epsilon}_{ij} \\ v^{\rho}_{ij} \end{bmatrix} = \begin{bmatrix} z_{ij} \\ \ln \bar{x}_{ij} + \tilde{\sigma}\kappa^{\tau} z_{ij} \end{bmatrix} - \sum_{k=1}^{K} \begin{bmatrix} \kappa^{\epsilon}\gamma^{\epsilon}_{g,k} f_k(\ln n) \\ \gamma^{\rho}_{g,k} f_k(\ln n) \end{bmatrix} - \begin{bmatrix} \delta^{\epsilon}_i + \zeta^{\epsilon}_j \\ \delta^{\rho}_i + \zeta^{\rho}_j \end{bmatrix}, \tag{31}$$

where $\tilde{\sigma} \equiv \sigma - 1$ and $\kappa^{\epsilon} \equiv 1/\left(\tilde{\sigma}\kappa^{\tau} + \kappa^{f}\right)$.

[20] Our notation allows groups to be defined as destination-origin country pairs over different years. In this case, one can easily extend our strategy to exploit variation over time in the observed trade cost shifter to obtain bilateral-specific estimates of the elasticity functions.

In terms of the structural unobserved shifters introduced above, $v_{ij}^{\epsilon} \equiv \kappa^{\epsilon}(v_{ij}^{\rho} - \eta_{ij}^{f})$ and $v_{ij}^{\rho} \equiv -\tilde{\sigma}\eta_{ij}^{\tau}$. Thus, Assumption 3 implies that $E[v_{ij}^{\epsilon}|z_{ij}, D_{ij}] = E[v_{ij}^{\rho}|z_{ij}, D_{ij}] = 0$. Combined with (31), these moment conditions can be used to estimate the parameters $\gamma_{g,k}^{\rho}$ and $\gamma_{g,k}^{\epsilon}$.[21]

Pass-through from observed shifter to fixed entry cost. To estimate the scale parameter κ^{ϵ}, we exploit the restriction imposed by the specification of entry costs in terms of labor in the origin country. Under this assumption, we show in Appendix A.7 that

$$v_j^f = \kappa^{\epsilon}\zeta_j^{\rho} - \zeta_j^{\epsilon}. \tag{32}$$

Here, $v_j^f \equiv \kappa^{\epsilon}\zeta_j^f$ is the destination fixed-effect in the first-stage specification for the entry cost in (28) and, because (28) contains a constant, $E[v_j^f] = 0$. We use this moment condition to estimate κ^{ϵ}.

Estimator. Expressions (30)–(32) can be used to compute $(v_{ij}^{\tau}, v_{ij}^{\epsilon}, v_{ij}^{\rho}, v_j^f)$ conditional on our main parameters of interest, $\boldsymbol{\Theta} \equiv \left(\kappa^{\epsilon}, \kappa^{\tau}, \left\{\gamma_{g,k}^{\rho}, \gamma_{g,k}^{\epsilon}\right\}_{g,k=1}^{G,K}\right)$, as well as the set of origin fixed-effects, $\boldsymbol{\delta} \equiv \{\delta_i^{\tau}, \delta_i^{\epsilon}, \delta_i^{\rho}\}_{i=1}^{N}$ and destination fixed-effects, $\boldsymbol{\zeta} \equiv \left\{\zeta_i^{\tau}, \zeta_j^{\epsilon}, \zeta_j^{\rho}\right\}_{j=1}^{N}$. We use the recovered structural residuals to construct the following GMM estimator for $(\boldsymbol{\Theta}, \boldsymbol{\delta}, \boldsymbol{\zeta})$:

$$\min_{(\Theta,\delta,\zeta)} h\left(\boldsymbol{\Theta}, \boldsymbol{\delta}, \boldsymbol{\zeta}\right)' \hat{\Omega} h\left(\boldsymbol{\Theta}, \boldsymbol{\delta}, \boldsymbol{\zeta}\right), \quad \text{where} \quad h\left(\boldsymbol{\Theta}, \boldsymbol{\delta}, \boldsymbol{\zeta}\right) \equiv \begin{bmatrix} \sum_{ij}\left(v_{ij}^{\tau} z_{ij},\ v_{ij}^{\tau} D_{ij}\right)' \\ \sum_{ij}\left(v_{ij}^{\epsilon} F(z_{ij}), v_{ij}^{\epsilon} D_{ij}\right)' \\ \sum_{ij}\left(v_{ij}^{\rho} F(z_{ij}), v_{ij}^{\rho} D_{ij}\right)' \\ \sum_j v_j^f \end{bmatrix}, \tag{33}$$

and $\hat{\Omega}$ is the two-step optimal matrix of moment weights and $F(z_{ij})$ is the following function,

$$F(z_{ij}) \equiv \left\{\mathbb{I}_{(ij\in g)}\mathbb{I}_{(n\in\mathcal{U}_k)}(z_{ij})^d\right\}_{g=1,k=1,d=1}^{G,K,3}.$$

There are two ways of perceiving our estimation procedure to recover $\rho_g(n)$ and $\epsilon_g(n)$. First, imposing that $\rho_g(n)$ and $\epsilon_g(n)$ are given by the flexible functional form in Assumption 4 implies that identification, consistency, and inference follow from usual results for GMM. As such, identification requires the typical GMM rank condition. Our functional form choice fits in the general class of sieve functions, which can be represented as $h(x) = \sum_{d=1}^{D} \alpha_d f_d(x)$ where $\{f_d(x)\}_d$ is a set of known functional basis. We choose a series of cubic polynomials

[21] Conditional on observing κ^{τ}, the assumption that $E[\eta_{ij}^{u}|z_{ij}, D_{ij}] = 0$ is not necessary for the estimation of $\rho_g(n)$ and $\epsilon_g(n)$ using (31). Accordingly, as in Adao et al. (2017), it is possible to estimate $\rho_g(n)$ and $\epsilon_g(n)$ with the alternative assumption of perfect pass-through from z_{ij} to $\bar{\tau}_{ij}$ (i.e., $\kappa^{\tau} \equiv 1$).

over intervals, as in Ryan (2012), but restrict the bottom and upper intervals to have a linear slope to improve on the precision of our estimates at extreme values of the support.

Alternatively, the functional form in Assumption 4 can be seen as a functional basis for the nonparametric estimation of $\rho_g(n)$ and $\epsilon_g(n)$. Under this interpretation, our estimator is the sieve nonparametric instrumental variable (NPIV) estimator in Chen and Qiu (2016), Chen and Christensen (2018), and Compiani (2019). In this case, identification requires the assumption of completeness in Newey and Powell (2003) or, in the case of our model with a linear component, the weaker version of this assumption in Florens et al. (2012).[22] Chen and Christensen (2018) derive converge rates and confidence intervals for this type of sieve NPIV estimator. In Appendix B.4, we implement their procedure to compute confidence bands accounting for the fact that the function basis approximates $\rho_g(n)$ and $\epsilon_g(n)$.

4.3.1 The Constant Elasticity Benchmark

To gain intuition for the estimation strategy, we return to the constant elasticity benchmark where $\epsilon_{ij}(n) = \epsilon(n)$ and $\rho_{ij}(n) = \rho(n)$ are log-linear (as in (20)). In this case, the equations in (31) yield two constant-elasticity gravity equations, one for the exporter firm share n_{ij}, and another for the average firm exports $\bar{x}_{ij}$:

$$\begin{aligned} \ln n_{ij} &= \beta^\epsilon z_{ij} + \tilde{\delta}_i^\epsilon + \tilde{\zeta}_j^\epsilon + \tilde{\eta}_{ij}^\epsilon \\ \ln \bar{x}_{ij} &= \beta^\rho z_{ij} + \tilde{\delta}_i^\rho + \tilde{\zeta}_j^\rho + \tilde{\eta}_{ij}^\rho, \end{aligned} \tag{34}$$

where

$$\beta^\epsilon \equiv (\kappa^\epsilon \varepsilon)^{-1} \quad \text{and} \quad \beta^\rho \equiv -(\sigma - 1)\kappa^\tau + \varrho\beta^\epsilon. \tag{35}$$

These expressions illustrate how our estimation strategy identifies the parameters controlling the functions $\epsilon_{ij}(n) = \epsilon(n)$ and $\rho_{ij}(n) = \rho(n)$. The parameter ε of $\epsilon(n)$ is proportional to β^ϵ: the constant gravity elasticity of the bilateral exporter firm share to the bilateral cost shifter (conditional on origin and destination fixed-effects). Given the fact that $\tilde{\eta}_{ij}^\epsilon \equiv \beta^\epsilon \kappa^\epsilon (\eta_{ij}^f + \tilde{\sigma}\eta_{ij}^\tau)$, the consistent estimation of β^ϵ requires the orthogonality between the observed shifter z_{ij} and the unobserved shifters of bilateral trade costs $(\eta_{ij}^f, \eta_{ij}^\tau)$ in (28). This is the same type of assumption necessary for the causal interpretation of the estimates of any gravity equation. Notice that, in order to independently recover ε from the definition of β^ϵ in (35), we need the pass-through of the cost shifter to the fixed entry cost κ^ϵ, which we estimate using the relationship described above between the destination fixed-effects in the

[22] The completeness assumption is not testable (Canay et al., 2013), but it is generically satisfied (Andrews, 2011; Chen and Christensen, 2018). By imposing that $\rho_g(n)$ and $\epsilon_g(n)$ are bounded, identification can be achieved by the weaker condition of bounded completeness (Blundell et al., 2007).

two gravity equations.[23]

In addition, the parameter ϱ in $\rho(n)$ regulates β^{ρ}: the constant gravity elasticity of the bilateral average firm exports to the bilateral cost shifter (conditional on origin and destination fixed-effects). Again, because $\tilde{\eta}^{\rho}_{ij}$ is a linear combination of the unobserved shifters of variable and fixed trade costs in (28), the consistent estimation of β^{ρ} requires the observed shifter z_{ij} to be orthogonal to $(\eta^{f}_{ij}, \eta^{\tau}_{ij})$. Finally, to separately recover ϱ from β^{ρ} in (35), we need the pass-through of z_{ij} to the variable entry cost κ^{τ}, which we estimate from the log-linear specification in (30) with an observable component of variable trade costs, τ_{ij}.

Our GMM estimator in (33) is an extension of the estimator for this log-linear specification. The only difference is that, instead of imposing a unique interval with a log-linear function, we specify that the elasticities are determined by cubic splines that may differ over K intervals: the parameters $\left\{\gamma^{\rho}_{g,k}, \gamma^{\epsilon}_{g,k}\right\}^{G,K}_{g,k=1}$ governing the functions $\epsilon_{ij}(n) = \epsilon_g(n)$ and $\rho_{ij}(n) = \rho_g(n)$ in Assumption 4. Thus, the parameters controlling $\epsilon_g(n)$ and $\rho_g(n)$ are identified from the fact that the extensive and intensive margin elasticities may differ across country pairs (in the same group) that have different levels of the exporter firm share.

5 Empirical Estimation

We use the strategy above to estimate $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ using the semiparametric gravity equations for the extensive and intensive margins of firm exports. Our results show how the two elasticity functions of firm exports vary with the exporter firm share. In the next section, we use our estimates to evaluate how much firm heterogeneity matters for computing counterfactual changes in aggregate outcomes following trade shocks.

5.1 Data

Our baseline data source for bilateral trade flows is the 2016 release of the World Input-Output Database (WIOD). It contains domestic sales, X_{ii}, as well as bilateral trade flows, X_{ij}, for 43 countries. The first columns in Table OA.1 in Appendix B.1 presents the list of countries with trade flows in the WIOD. Our sample of countries accounts for 90% of world trade and entails positive bilateral flows for almost all exporter-importer pairs.[24]

The estimator in equation (33) requires four bilateral variables: (i) the exporter firm share, n_{ij}; (ii) the average firm revenue, $\bar{x}_{ij}$; (iii) the trade cost shifter, z_{ij}; and (iv) the

[23] In this case, this relationship can be written as $v^{f}_{j} = \kappa^{\epsilon}\left(\tilde{\zeta}^{\rho}_{j} - \varrho\tilde{\zeta}^{\epsilon}_{j}\right) + \beta^{\epsilon}\tilde{\zeta}^{\epsilon}_{j}$ with $E[v^{f}_{j}] = 0$.

[24] This attenuates concerns related to the estimation of gravity equations with zero trade flows, as in Helpman et al. (2008) and Silva and Tenreyro (2006).

observed component of trade costs, τ_{ij}. We now describe how we construct these variables for a subset of the countries in the WIOD in 2012.

We use various sources to construct n_{ij} and $\bar{x}_{ij}$ for a subset of 37 origin countries in the WIOD – for the full list of countries, see columns (2) and (3) of Table OA.1 in Appendix B.1. We construct the data in two steps. We first use the OECD Trade by Enterprise Characteristics (TEC) database to obtain the number of manufacturing firms from i selling in j, N_{ij} for $i \neq j$. For origin-destination pairs not in the OECD TEC database, we obtain N_{ij} from the World Bank Exporter Dynamics Database (EDD). These datasets also contain the total exports of the same set of firms from i exporting to j. We use this information to compute the average revenue of firms from i selling to j, $\bar{x}_{ij}$.[25]

The second step is the construction of the number of entrants N_i, which is not readily available in national statistics. Together with the number of exporters N_{ij}, we use N_i to construct the exporter firm shares, $n_{ij} = N_{ij}/N_i$.[26] We compute the number of entrants as $N_i = N_{ii}/n_{ii}$ where, in country i, N_{ii} is the number of active manufacturing firms and n_{ii} is the survival probability of new manufacturing firms. Our approach assumes that a low survival rate represents a large pool of entrants that pay the sunk entry cost but fail to be productive enough to survive. A high survival rate reflects instead that most firms paying the entry cost are successful in production. To maximize country coverage, we obtain N_{ii} from several datasets: the OECD Demographic Business Statistics (SDBS), the OECD Structural Statistics for Industry and Services (SSIS), and the World Bank Enterprise Surveys. In addition, we obtain n_{ii} from the one-year survival rate of manufacturing firms in the OECD SDBS.[27]

Our measure of the bilateral trade cost shifter, z_{ij}, is the log of bilateral distance (population-weighted) in the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII). This dataset includes not only distance between countries, but also distances within a country z_{ii} due to the nature of population weighting. We use this information to include observations associated with domestic trade in our baseline sample.

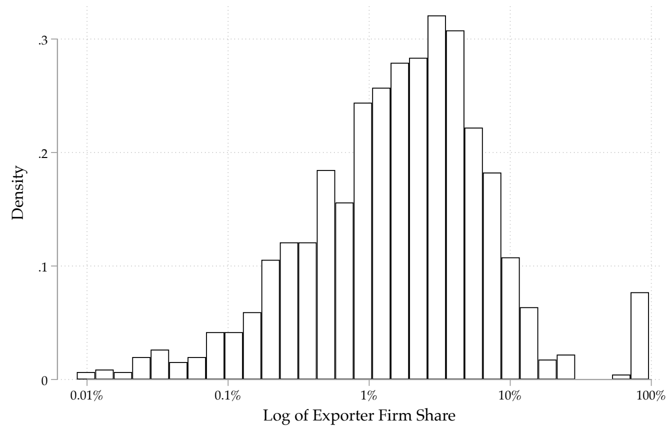
Finally, we use the bilateral freight cost to measure the observed component of variable trade costs τ_{ij}. This is only necessary for the estimation of κ^{τ} using the linear specification in (30). We consider a subset of countries for which we observe CIF/FOB import margins in

[25] For Australia and China, we directly construct N_{ij} and $\bar{x}_{ij}$ from national statistic agencies.

[26] Prior research circumvents this data requirement by assuming that $N_i = N_{ii}$ and $n_{ii} = 1$ – e.g., see Fernandes et al. (2019). However, this limits the potential sources of gains from trade in our model by shutting down welfare gains implied by changes in domestic firm composition – see Section 3.3.

[27] This data is only available for 80% of the origin countries in our sample. We impute the survival rate for the remaining countries using the simple average of the survival rate for countries with available data. We show that our results are robust to excluding from the sample countries without survival rate data. We also show that our results are similar when we use two-year or three-year survival rates.

Figure 2: Empirical distribution of $\ln n_{ij}$, 2012



Note. Empirical distribution of $\ln(n_{ij})$ in the cross-section of origin-destination pairs in 2012. For the list of countries, see Table OA.1 in Appendix B.1.

the OECD freight cost database. For each country pair, we compute τ_{ij} as the ratio of the total CIF and FOB imports for manufacturing products. Columns (4) and (5) of Table OA.1 in Appendix B.1 report the list of countries with available data on bilateral freight costs.

The availability of data on $\bar{x}_{ij}$, n_{ij}, and z_{ij} determines our sample for the estimation of the last three moment conditions in (33). Table OA.1 in Appendix B.1 reports the list of countries in our baseline sample. Figure 2 summarizes the distribution of $\ln(n_{ij})$ for all bilateral pairs in 2012.[28] The empirical distribution of n_{ij} is central for our analysis: because n_{ij} is the only input of the elasticity functions $\rho_g(n)$ and $\epsilon_g(n)$, we are only able to precisely estimate these functions in the part of the support in which we observe values of n_{ij}.

5.2 Pass-Through of Distance to Freight Costs

We start by estimating the pass-through parameter κ^τ from the linear specification in (30). We consider the pooled sample of exporter-importer-year triples in 2008-2014. Table 1 reports our pass-though estimates along with standard errors clustered at the destination-origin level. We estimate an elasticity of trade costs to distance of roughly 0.35. We obtain similar pass-through estimates in the presence of different sets of fixed-effects. This is reassuring given that the fixed-effects absorb a great deal of variation in freight costs in our sample – the R^2 increases from 0.48 in columns (1) to 0.81 in column (3).

[28] We obtain a similar country coverage for ($\bar{x}_{ij}$, n_{ij}, z_{ij}) in every year between 2010 and 2014. In addition, 2012 is the year with the most observations of the freight cost τ_{ij} used in estimation. In the appendix, we show that our results are similar when we use data for 2010 and 2014.

Table 1: Estimation of κ^τ

	Dep. Var.: Log of Freight Cost		
	(1)	(2)	(3)
Log of Distance	0.349***	0.360***	0.369***
	(0.057)	(0.083)	(0.097)
R^2	0.483	0.726	0.813
Fixed-Effects:			
Year	Yes	Yes	No
Origin, Destination	No	Yes	No
Origin-Year, Destination-Year	No	No	Yes

Note. Sample of 547 origin-destination-year triples described in Table OA.1 of Appendix B.1. Standard errors clustered by origin-destination pair. *** p $<$ 0.01

5.3 Results: Constant Elasticity Gravity

As a benchmark, Table 2 presents the estimates of ε and ϱ obtained with the GMM estimator in (33) under the constant elasticity assumption in (20). Using (11), our estimate of ε indicates that a 1% increase in bilateral trade costs triggers a change of $(\sigma - 1)/\varepsilon = -3.5\%$ in the exporter firm share (conditional on origin and destination fixed-effects). This is consistent with a skewed distribution of firm entry potentials such that the exporter firm share falls sharply with the increase in cost caused by higher bilateral distance. Our estimate of ρ indicates that a 1% increase in the exporter firm share is associated with a 0.21% decline in the mean revenue potential of exporters, implying that marginal exporters are worse than infra-marginal exporters. The combination of the definition in (17) and our estimates of ε and ϱ implies a trade elasticity of 5, which is within the range of estimates in the literature – see Costinot and Rodriguez-Clare (2013).

As shown in Section 4.3.1, our structural estimates are closely related to the gravity elasticity of the extensive and intensive margin of firm exports to distance, which we report in Table OA.2 of Appendix B.3.1. Columns (1) and (2) indicate that 1% higher distance is associated with declines of $\beta^\epsilon = 1.2\%$ in exporter firm shares and $\beta^\rho = 0.8\%$ in average firm exports. The extensive margin gravity coefficient is $\beta^\epsilon = 1/\kappa^\epsilon\varepsilon$, which yields $\varepsilon = 1.1$ because the ratio of the average destination fixed-effects is $\kappa^\epsilon \approx 0.75$. The intensive margin gravity coefficient is $\beta^\rho \equiv \varrho\beta^\epsilon - (\sigma - 1)\kappa^\tau$, which yields $\rho = -0.21$ because we set $\sigma = 3.9$ and estimate $\kappa^\tau = 0.36$ in Table 1. Thus, due to the non-zero intensive margin elasticity, we reject the restriction of $\varrho = \varepsilon$ imposed in the Melitz-Pareto model of Chaney (2008).

Table 2: Constant Elasticity Gravity of Firm Exports with $\epsilon_{ij}(n) = n^{\varepsilon}$ and $\rho_{ij}(n) = n^{\varrho}$

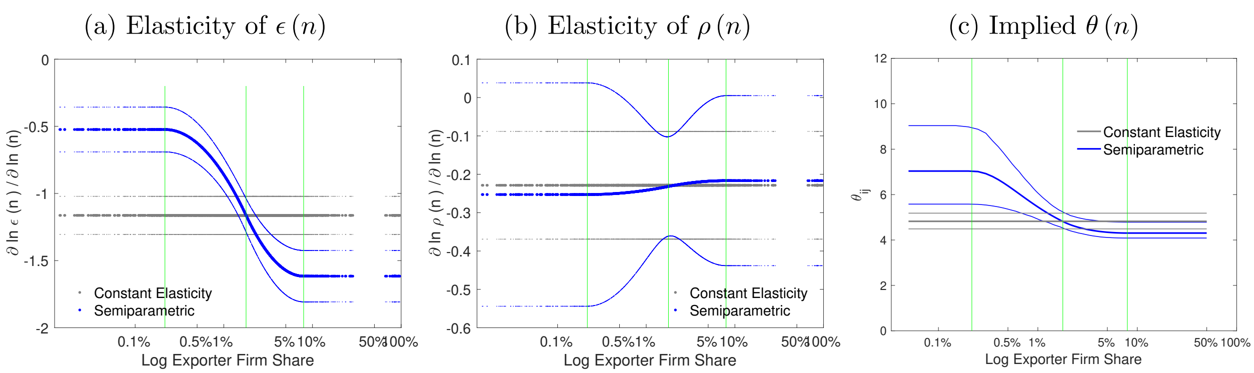
ε	ϱ	θ	κ^{ϵ}	κ^{τ}
-1.16	-0.23	4.82	0.75	0.36
(0.030)	(0.029)	(0.180)	(0.038)	(0.083)

Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,522 origin-destination pairs described in Table OA.1 of Appendix B.1. Calibration of $\tilde{\sigma} = \sigma - 1 = 2.9$ from Hottman, Redding and Weinstein (2016). Standard errors in parenthesis and 95% confidence intervals in brackets. For ϱ, ε and κ^{ϵ}, we report robust standard errors implied estimation of (33). For θ, we report standard error obtained from 1,000 bootstrapped draws. For κ_{τ}, we report standard errors clustered by origin-destination, as in Table 1.

5.4 Results: Semiparametric Gravity

We now turn to our semiparametric estimates of $\epsilon(n)$ and $\rho(n)$ for a single group pooling all countries. Figure 3 presents estimates of the elasticities of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share. We use green bars to denote the estimation knots, and overlay our baseline estimates with the estimates of the constant elasticity specification presented in Table 2. We report the elasticity of $\epsilon(n)$ in Panel (a), the elasticity of $\rho(n)$ in Panel (b), and the implied trade elasticity $\theta(n)$ in Panel (c) (obtained from (17)).

Figure 3: Semiparametric Gravity of Firm Exports with $\epsilon_{ij}(n) = \epsilon(n)$ and $\rho_{ij}(n) = \rho(n)$



Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,522 origin-destination pairs described in Table OA.1 of Appendix B.1. Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = \sigma - 1 = 2.9$ from Hottman, Redding and Weinstein (2016). Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 1,000 bootstrap draws for $\theta(n)$.

Our estimates show that the extensive margin elasticity varies with the share of firms exporting to a market. Since $\frac{\partial \ln n_{ij}}{\partial \ln \bar{\tau}_{ij}}$ is inversely proportional to $\varepsilon(n)$ (see equation (12)), Panel (a) shows that the extensive margin becomes less responsive as more firms serve a market. For low levels of entry, the expression in (12) implies that a 1 log-point increase in trade costs reduces the share of exporting firms by 5.8 log-points. This elasticity is lower for higher levels of firm entry. In the top knot, a 1 log-point increase in trade costs reduces the exporter firm share by 2.1 log-point. This implies that, at first, exporters are very sensitive to changes in trade frictions. However, high levels of entry potential are rare in the economy: when many firms export, small changes in trade frictions lead to smaller responses in the share of firms that decide to export.

Panel (b) indicates that selection patterns do not change substantially with the exporter firm share. For all levels of n_{ij}, the entry of 1% more exporters into a market induces a reduction in the mean revenue potential of around 0.2%. Thus, for different levels of entry potential, marginal exporters are worse than incumbent exporters in terms of average sales by a similar amount.

Panel (c) shows what the elasticities of the two margins of firm exports imply for the response of bilateral trade flows to changes in bilateral trade costs. Due to the declining extensive margin, the trade elasticity is lower when the exporter firm share is high, falling from 7 in pairs with n_{ij} in the lower knot to 4 in pairs with n_{ij} in the top knot.[29] Thus, in line with the product-level evidence in Kehoe and Ruhl (2013) and Kehoe et al. (2015), the trade elasticity tends to be lower when trade volumes are high.

As discussed in Section 4.3.1, our estimation strategy extends the constant elasticity gravity specification widely used in the literature. As such, we obtain variation in the estimated coefficients around the values obtained when we restrict the elasticities to be constant for all values of n_{ij}. The patterns shown in Figure 3 arise because the impact of distance on firm level exports varies across different parts of the support of n_{ij}. In Table OA.3 of Appendix OA.2, we show that the extensive margin elasticity is less sensitive to distance among country pairs with high n_{ij}, while the intensive margin elasticity varies little with n_{ij}. The decreasing the extensive margin elasticity then implies that bilateral trade flows become less sensitive to distance among countries with high values of n_{ij}.

In Figure OA.10 of Appendix B.5.1, we compare our estimated trade elasticity function to that implied by parametric assumptions and their associated estimates about the distribution of firm fundamentals in the existing literature. The log-normal assumption in Bas et al. (2017) and Head et al. (2014) implies a much steeper trade elasticity function. The trade elasticity for high levels of n_{ij} is below two, while it is above twelve for low levels of n_{ij}. The truncated Pareto assumption in Melitz and Redding (2015) yields a trade elasticity function that is uniformly low. It is always below four and falls below two for high levels of n_{ij}. In both cases, the sharp decline in the trade elasticity is driven by strong reductions in $\varepsilon(n_{ij})$

[29] It is possible that existing trade elasticity estimates are average treatment effects obtained from variation in particular parts of the support of exporter firm shares. Our approach then just captures how the trade elasticity varies across the support of values of n_{ij}.

and $\varrho(n_{ij})$ when n_{ij} increases.

These comparisons highlight the difference between our approach based on semiparametric gravity equations and approaches based on parametrizations of cross-section variation in firm-level outcomes. While we directly estimate the elasticity functions driving the model's aggregate predictions, the parametric micro approach extrapolates from heterogeneity in firm-level outcomes to obtain these elasticity functions. Our results indicate that this extrapolation may lead to elasticity functions that are substantially different from those implied by estimates of the semiparametric gravity equations of firm exports. In the next section, we investigate the quantitative implications of such differences for the model's counterfactual predictions.

Robustness of baseline estimates. In Appendix B.4, we investigate the robustness of the baseline estimates in Figure 3. We show that confidence intervals are similar when we implement the inference method in Chen and Christensen (2018) that accounts for the fact that the function basis in Assumption 4 approximates the nonparametric functions $\epsilon(n)$ and $\rho(n)$. We show that estimates are similar when (i) we use data for alternative years with similar country coverage, (ii) we exclude observations associated with domestic sales, (iii) we measure n_{ii} using two-year or three-year survival rates, (iv) we exclude from the sample origin countries with imputed survival rates, (v) we assume that all entrants sell in the domestic market (i.e., $n_{ii} = 1$ and $\bar{f}_{ii} = 0$), and (vi) we use a higher elasticity of substitution given by $\sigma = 5$.

Additional estimates for multiple country groups. Our baseline estimates impose identical elasticity functions across all exporter-destination pairs ($G = 1$). In the Online Supplementary Material B, we allow the elasticity functions to vary across country groups. Specifically, we investigate whether the trade elasticity functions vary with the country's per capita income, as in Adao et al. (2017). For developed origin countries, all elasticity functions are similar to those reported in Figure 3. However, for developing origins, the extensive margin elasticity almost does not vary with n_{ij}, leading to a trade elasticity that is roughly constant around six. We further show that the elasticity estimates do not vary with the destination's development level. Finally, we also find that the elasticity functions do not differ across country pairs that belong to free trade areas, and share a common language or currency.

6 Quantifying The Importance of Firm Heterogeneity for The Impact of Trade Shocks on Welfare

We conclude by applying our methodology to study the welfare consequences of two types of trade shocks. In Section 6.1, we use the results in Sections 3.1 and 3.3 to quantify the impact of reducing trade costs from infinity to the levels observed in the current trade equilibrium – i.e., the so-called gains from trade. In Section 6.2, we use the results in Section 3.2 to identify the changes in economic fundamentals in the world economy between 2003 and 2012. We then use the inverted shocks to quantify the welfare consequences of good market integration among countries in the European single market. In both cases, we measure the importance of firm heterogeneity for the shock's impact by comparing results implied by our semiparametric estimates of the gravity equations of firm exports and those implied by the benchmark constant elasticity gravity model of bilateral trade flows.

6.1 Welfare Impact of Moving to Autarky

To investigate how firm heterogeneity affects the gains from trade, we consider results under two scenarios. First, the benchmark constant elasticity specification in Table 2, where the gains are given by the aggregate sufficient statistics in Arkolakis et al. (2012), and firm heterogeneity does not play any role for the gains from trade. Second, the semiparametric gravity estimates in Figure 3, where the gains are given by the nonparametric sufficient statistics in Section 3.3, and depend on the export decisions of firms.[30]

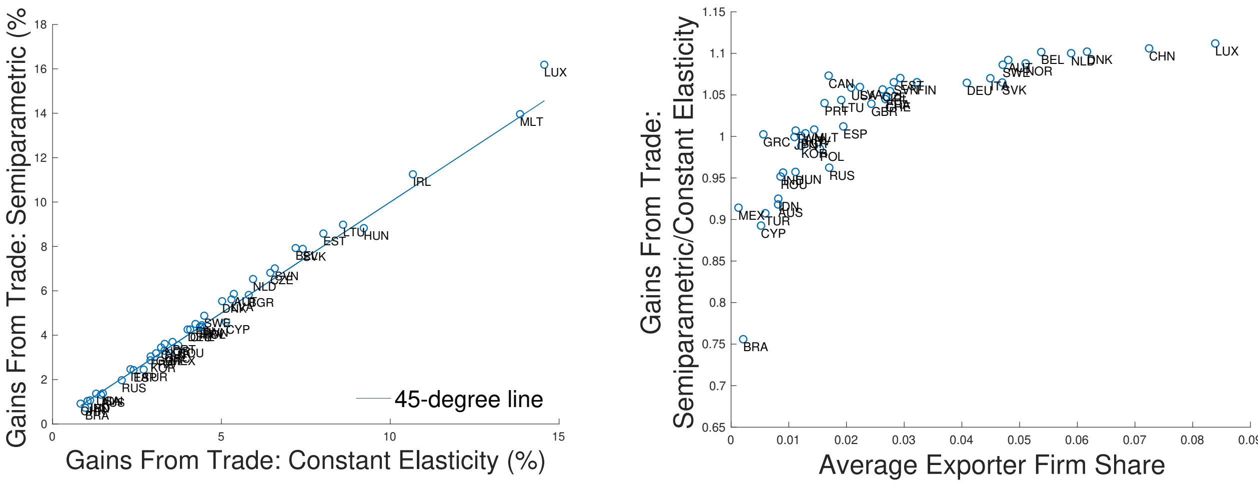
Figure 4a compares the gains under the two scenarios implied by a move from the counterfactual autarky equilibrium to the trade equilibrium of 2012. They yield highly correlated gains from trade – the cross-country correlation is 0.99. As pointed out in Section 3.3, this is a consequence of the fact that the domestic trade share remains an important driver of the gains from trade in our general specification. Notice that firm heterogeneity may still have a substantial impact on the gains from trade of some countries. It yields gains from trade that are roughly 10% higher for Luxembourg, China, Denmark, Belgium and Netherlands. However, the gains are more than 10% lower for Brazil, Mexico, Turkey, India and Australia. Overall, when we account for firm heterogeneity, the absolute average change in the gains from trade is 6%.

In Figure 4b, we investigate how firm heterogeneity affects the ratio between the gains

[30]We solve $\hat{n}_{ii}$ and $\hat{N}_i$ using the system in Appendix A.6. We have data on n_{ij} for 81% of country pairs in our baseline sample, accounting for 88% world trade flows in 2012. To compute gains from trade, we impute the exporter firm share for the subset of pairs with missing data using estimates of the constant elasticity gravity equations reported in column of (1) of Table OA.2.

Figure 4: Importance of Firm Heterogeneity for the Gains from Trade

(a) Gains from Trade　　　　(b) Gains from Trade and Exporter Firm Share



Note. Gains from trade is the percentage change in the real wage implied by moving from autarky to the observed equilibrium in 2012. Gains from trade for semiparametric specification computed with the formula in Section 3.3 for $\hat{n}_{ii}$ and $\hat{N}_i$ solving the system in Appendix A.6 and the baseline semiparametric estimates in Figure 3. Gains from trade for constant elasticity specification computed with the formula in Section 3.3 with $\hat{N}_i = 1$ and $\theta_{ii}(n) = \theta$ reported in Table 2.

from trade implied by the semiparametric and the constant elasticity specifications. For each country, we compute this ratio and plot it against the average exporter firm share (i.e., average n_{ij} across $j \neq i$). The use of this statistic is motivated by our welfare gains derivation, equation (23), and a large tradition in international trade that stresses the link between firm heterogeneity and exporter success (see e.g. Bernard and Jensen (1999); Bernard et al. (2007); Melitz and Redding (2014)). The plot shows that firm heterogeneity amplifies the gains from trade in countries with a higher share of exporting firms. Economies of scale drive this relationship. When the fraction of firms exporting is high, more resources are allocated to covering the fixed cost of entering foreign markets. This strengthens competition in the domestic labor market and, therefore, amplifies the decline in domestic survival rate. This in turn creates higher welfare gains, as discussed in Section 3.3.

The Importance of Parametric Assumptions. In Appendix B.5.1, we investigate the quantitative importance of using our semiparametric approach to recover the two elasticity functions that are sufficient for counterfactual analysis. We compare our baseline estimates of the gains from trade to those implied by parametric assumptions about the firm productivity distribution and their associated estimates in the literature – specifically, the Truncated Pareto distribution in Melitz and Redding (2015) and the Log-normal distribution in Bas et al. (2017). Both parametric assumptions have quantitatively large impacts on the gains from trade. They yield gains from trade that are too large because they imply a low trade elasticity for all levels of the exporter firm share – see Figure OA.10 of Appendix B.5.1. These results indicate that one should be cautious when extrapolating elasticity functions from cross-sectional dispersion in firm-level outcomes.

6.2 The Welfare Impact of Observed Changes in Trade Costs

Our second exercise investigates the welfare consequences of actual changes in trade costs that recently hit the world economy. We use Proposition 3 to recover the changes in fixed costs of exporting $\hat{\bar{f}}_{ij}$ and revenue shifters $\hat{\bar{r}}_{ij}$ between 2003 and 2012.[31] We focus on measuring the welfare consequences of changes in trade costs between countries in the European single market, including the admission of 12 new members that took place in this period. Specifically, we ask "For any country j, how much higher (or lower) would welfare have been in 2012 if fixed costs of exporting and revenue shifters between members of the European single market were those of 2003 rather than those of 2012?".[32]

We first invert the changes in economic fundamentals between the EU-28, Norway, and Switzerland. In our model, they can be recovered from changes in bilateral trade outcomes, wages and prices. In fact, up to a first order approximation, (21)–(22) imply that

$$\begin{aligned} \log \hat{\bar{f}}^t_{ij} &\approx \log \hat{\bar{x}}^t_{ij} + (\varepsilon(n^0_{ij}) - \varrho(n^0_{ij})) \log \hat{n}^t_{ij} - \log \hat{w}^t_i \\ \log \hat{\bar{r}}^t_{ij} &\approx \log \hat{\bar{x}}^t_{ij} - \varrho(n^0_{ij}) \log \hat{n}^t_{ij} + (\sigma - 1) \log \hat{w}^t_i + \zeta^t_j \end{aligned} \tag{36}$$

where ζ^t_j is a destination-specific component. Given our baseline estimates, these expressions clearly illustrate what features in the data are associated with changes in economic fundamentals through the lens of our model. The fixed cost of exporting must have increased whenever we observe firm-level exports increasing ($\hat{\bar{x}}^t_{ij}$ increasing), but firms are not entering foreign markets ($\hat{n}^t_{ij}$ decreasing) or production costs are not increasing ($\hat{w}^t_i$ decreasing). Similarly, we recover an increase in the bilateral revenue shifter whenever we observe increases in either firm

[31] We construct changes in trade outcomes using the same data described in Section 5.1. The inversion of economic fundamentals also requires changes in wages and prices (relative to the economy's numeraire). We construct these variables using data from the Penn Tables while normalizing variables so that the U.S. wage is the numeraire. Since our model entails a single factor, our wage measure is the output-side real GDP at current PPPs divided by the number of engaged persons. Our price index measure is the corresponding GDP price deflator PPP-adjusted to U.S. prices of 2011. We normalize both variables by the U.S. per-worker GDP in each year.

[32] Our exercise differs from that in Caliendo et al. (2017) for two main reasons. First, we focus on the consequences of firm heterogeneity and ignore changes in labor mobility frictions across countries. Second, we consider all changes in economic fundamentals instead of changes in tariffs associated with the EU enlargement, so our exercise accounts for changes in (variable and fixed) trade costs and productivity experienced by both old and new members.

average exports (higher $\hat{\tilde{x}}_{ij}^t$) or wages (higher $\hat{w}_i^t$) given the change in firm composition (lower $\varrho(n_{ij}^0) \log \hat{n}_{ij}^t$). Note that such changes in revenue shifters are proportional to productivity and inversely proportional to variable trade costs.

Table 3 displays the results of regressing inverted changes in bilateral fundamentals on the changes in firm average exports, firm entry share, and origin's labor cost. Not surprisingly, all coefficient estimates are close to the values in the first-order approximation above. These variables explain most of the variation in fundamentals across country pairs: 93% for $\hat{\tilde{f}}_{ij}^t$ and 72% for $\hat{\tilde{r}}_{ij}^t$. More interesting, we use a Shapley decomposition of the R^2 to uncover which variables drive most of the cross-country variation in changes in fundamentals. For fixed exporting costs, the main driver is the growth in firm average exports, with a smaller relevant contribution of firm entry shares. In contrast, growth in firm average exports and labor costs have equal contributions for the variation in the growth of bilateral revenue shifters.

Table 3: Determinants of Changes in Bilateral Shifters

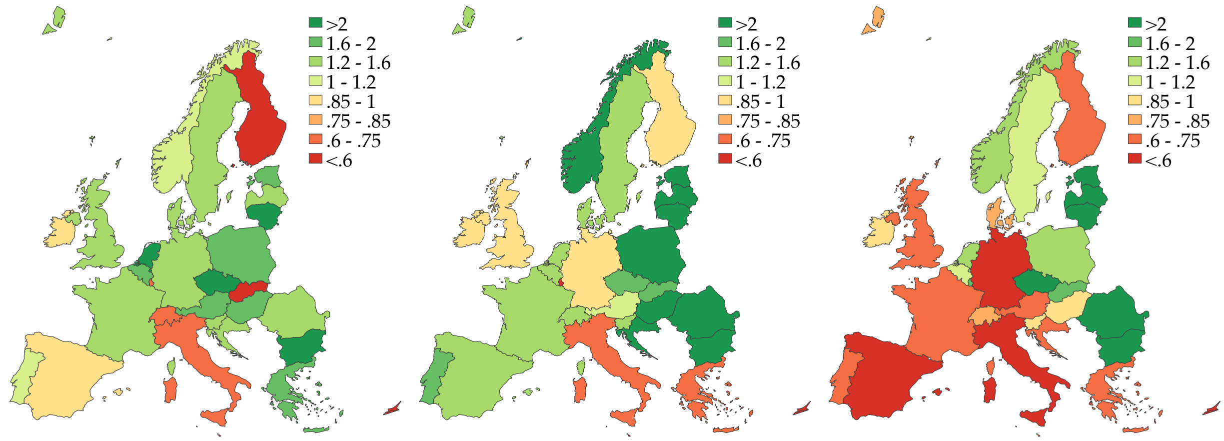
Dep. Var.	$\log \hat{\tilde{f}}_{ij}$	$\log \hat{\tilde{r}}_{ij}$
	(1)	(2)
$\log \hat{\tilde{x}}_{ij}$	0.976***	0.723***
	(0.015)	(0.027)
	[79.1%]	[49.6%]
$\log \hat{n}_{ij}$	-0.498***	-0.092*
	(0.052)	(0.046)
	[19.3%]	[2.1%]
$\log \hat{w}_i$	-0.853***	3.499***
	(0.005)	(0.123)
	[1.7%]	[48.3%]
Constant	-0.039**	0.753***
	(0.013)	(0.037)
R^2	0.934	0.719

Note. Sample of changes in bilateral shifters between 870 country pairs in the European single market (where origin and destination are different). We report the robust standard errors in parenthesis, and the Shapley contribution to the R^2 of each regressor in brackets. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$

Figure 5 displays each origin's average inverted shock for the 30 countries of our sample in the European single market. Panel (a) shows that, through the lens of our model, firms in most countries had to hire more labor to cover the fixed cost necessary to enter foreign markets. This follows from the fact that, in almost every country, the increase in bilateral average firm exports was much stronger than average growth in firm export shares and wages – see Figure OA.12 in Appendix B.5.2. Panel (b) reports the average change in bilateral revenue shifters for each origin country. This panel shows that revenue shifters increased substantially in most countries mainly because of their strong growth in average firm export

Figure 5: Average Change in Bilateral Shifters of Exporting to Countries in The European Single Market, 2003-2012

(a) Average $\hat{\hat{f}}_{ij}$ for $i \neq j$ (b) Average $\hat{\hat{r}}_{ij}$ for $i \neq j$ (c) Average $\hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$ for $i \neq j$



Note: Changes in economic fundamentals computed with the expressions in Proposition 3 using changes in outcomes between 2003 and 2012. For each origin country i, we report the simple average of the changes in bilateral fundamentals across all other destination countries j in the European single market (with $i \neq j$).

and wages – see Figure OA.12 in Appendix B.5.2.[33] This growth was particularly strong in Eastern Europe, exceeding 100% in several countries. Notice that, by the definition in (4), $\hat{\hat{r}}_{ij}$ incorporates changes in both variable trade costs and productivity. To measure only changes in trade costs, panel (c) presents the average change in the normalized bilateral revenue shifter, $\hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii} = \hat{\hat{b}}_{ij}(\hat{\hat{\tau}}_{ij})^{1-\sigma}$. Our results indicate that countries in Eastern Europe experienced strong average increases in normalized revenue shifters (and thus reductions in trade costs). In contrast, countries in Western Europe had reductions in their average normalized revenue shifter, indicating that the export growth was particularly weak compared to the domestic performance of firms.

We now turn to the welfare consequences of the changes in bilateral trade shifters between members of the European single market. The top panels of Figure 6 report the negative of the percentage change in the real wage of each country implied by a different set of shocks. A positive (negative) number corresponds to welfare gains (losses). Panel (a) shows that changes in both $\hat{\hat{f}}_{ij}$ and $\hat{\hat{r}}_{ij}$ caused welfare gains in all member countries (except Greece). The average gain was 1.7% among all countries. The map also shows that, because of stronger increases in revenue shifters, gains were higher for countries in East Europe that experienced an average gain of 3.4%. The largest gains occurred in Slovakia (4.1%), Lithuania (4.5%), Estonia (5.1%), and Hungary (5.2%).

[33]Notable exceptions are Greece and Italy. For these countries, the revenue shifter reduction follows from their weak growth in firm average exports and, more importantly, in wages (relative to the U.S.).

In panel (b), we display the real wage changes only considering the impact of changes in revenue shifters. Because we find that fixed exporting costs tended to increase in this period, welfare gains are typically higher when we do not feed shocks in fixed exporting costs. However, the similarity between panels (a) and (b) indicates that the change in fixed exporting costs had a quantitatively small impact on real wages throughout this period.

Finally, panel (c) reports the impact of changes in variable trade costs. The stark difference between panels (b) and (c) indicates that Western and Eastern Europe had different source of welfare gains. For countries in West Europe, gains are much lower in panel (c) than in panel (b) – the average gain is 0.5% in panel (b) and -0.1% in panel (c). This implies that these countries mostly benefited from productivity gains in their trade partners. In fact, as discussed above, they seem to have experienced average increases in their variable export costs (reduction in normalized revenue shifters). For countries in East Europe, the gains are similar in both panels. This indicates that their gains came mostly from gaining access to the European single market. Interestingly, the only Eastern country to suffer a welfare loss, Croatia, is also the only in country in our sample to join the EU after 2012.

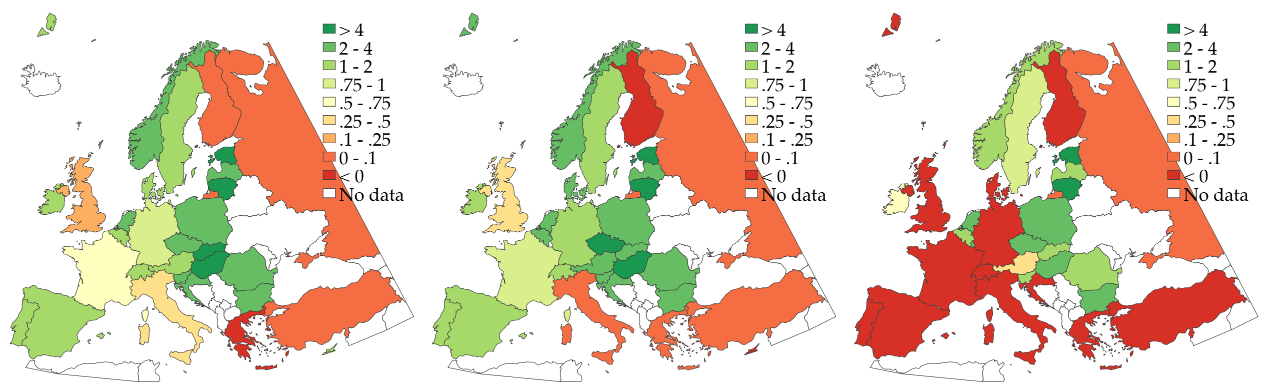
Panels (e)–(f) of Figure 6 conclude our analysis by comparing the real wage change obtained by our baseline model based on the semiparametric estimates in Figure 3 and those obtained by the constant elasticity benchmark in Table 2. We consider the same set of changes in economic fundamentals. For the three sets of shocks, our baseline estimates yield larger gains than those implied by the constant elasticity benchmark: the median ratio is 1.08 in panel (a), 1.04 in panel (b) and 1.04 in panel (c). This difference combines two consequences of the non-linearity in the estimates of the elasticity function: different recovered shocks ($\hat{\bar{f}}_{ij}$and $\hat{\bar{r}}_{ij}$), as well as different firm-level counterfactual responses. In particular, it is mainly driven by the fact that the changes in revenue shifters affect many country pairs for which the firm export share is in the range with a semiparametric trade elasticity above the benchmark constant trade elasticity.

The Importance of Recovering Asymmetric Trade Costs. Appendix B.5.2 investigates the quantitative importance of using our methodology to recover shocks in bilateral trade costs. We consider two alternative approaches in the existing literature. First, we measure trade cost shocks using tariff changes triggered by the EU enlargement, as in Caliendo et al. (2017). Since tariffs are only a fraction of trade costs, this approach yields shocks that are much smaller and more homogeneous (especially for pre-2003 EU members). This translates into much smaller welfare changes, with small gains for all countries. Finally, we consider a non-parametric extension of the approach in Head and Ries (2001) that imposes symmetry of trade cost shocks. This yields smaller welfare losses in West Europe and smaller welfare

Figure 6: The Welfare Impact of Changing Bilateral Shifters of Exporting to Countries in The European Single Market, 2003-2012

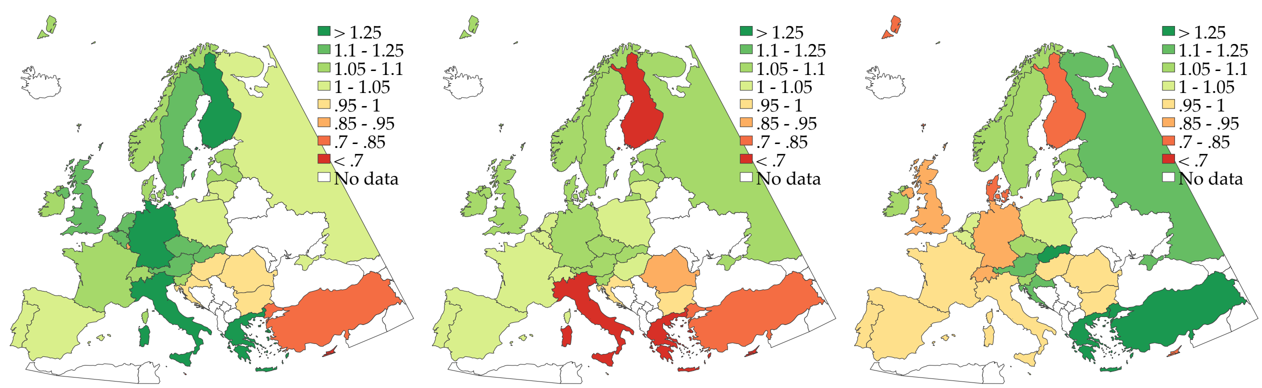
Percentage Change in Real Wage: Semiparametric Gravity

(a) $\hat{\hat{f}}_{ij}$ and $\hat{\hat{r}}_{ij}$ for $i \neq j$ (b) $\hat{\hat{r}}_{ij}$ for $i \neq j$ (c) $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = \hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$ for $i \neq j$



Ratio of Percentage Change in Real Wage: Semiparametric and Constant Elasticity Gravity

(d) $\hat{\hat{f}}_{ij}$ and $\hat{\hat{r}}_{ij}$ for $i \neq j$ (e) Impact of $\hat{\hat{r}}_{ij}$ for $i \neq j$ (f) $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = \hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$for $i \neq j$



Note: Real wage changes caused by changes in bilateral trade shifters between 2003 and 2012 among 30 countries in the European single market. Bilateral trade shifters computed with the expressions in Proposition 3 for observed changes in outcomes between 2003 and 2012. Counterfactual exercise: starting from the equilibrium in 2012, we compute the impact of changing fundamentals back to their level of 2003 (while holding $\hat{\hat{r}}_{ii} = \hat{\hat{f}}_{ii} = 1$). Panels (a)–(c) report minus the real wage change obtained with the baseline semiparametric estimates in Figure 3. Panel (d)–(f) report the ratio of real wage changes obtained with the baseline semiparametric estimates in Figure 3 and the constant elasticity estimates in Table 2.

gains in East Europe because the symmetry assumption attenuates the increase (decline) in bilateral trade costs for West (East) Europe shown in Figure 5.

7 Conclusion

We propose a new way of modeling and estimating the aggregate implications of firm heterogeneity in the workhorse monopolistic competition framework of international trade that dispenses parametric restrictions on the distribution of firm fundamentals. We use this approach to revisit a number of open questions about the role of firm heterogeneity for the economy's response to trade shocks through the lens of a new *not*-parametric point of view.

Instead of focusing on parametrically specifying the distribution of various firm-specific wedges, we show that they can be folded into two elasticity functions in the model's semi-parametric gravity equations that intuitively shape how trade costs affect firm-level entry and sales across country pairs. Given the initial equilibrium, the different sources of firm heterogeneity, and any associated parametric assumption imposed, only matter for counterfactual predictions through the shape of these two gravity elasticity functions of firm exports. This characterization also allows us to (i) construct nonparametric counterfactual predictions to trade shocks, (ii) nonparametrically recover changes in economic fundamentals from observed trade and macroeconomic outcomes, and (iii) obtain nonparametric ex-post sufficient statistics for the impact of trade shocks on welfare.

Our results indicate that a key new statistic for aggregate gains from trade is the share of exporting firms. We evaluate its impact on the trade elasticity and the welfare impact of trade shocks. We find that firm heterogeneity amplifies both the gains from trade in countries with more exporter firms, and the welfare gains of European good market integration in 2003-2012.

References

Adão, Rodrigo, "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from Brazil," *Unpublished paper, MIT*, 2015.

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson, "Nonparametric counterfactual predictions in neoclassical models of international trade," *The American Economic Review*, 2017, *107* (3), 633--689.

_ **, Costas Arkolakis, and Federico Esposito**, "General equilibrium effects in space: Theory and measurement," Technical Report, Tech. rep., Working Paper 2020.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085--1140.

__ , __ , **and Yuta Takahashi**, "Universal gravity," *NBER Working Paper*, 2014, (w20787).

Andrews, Donald WK, "Examples of l2-complete and boundedly-complete distributions," 2011.

Arkolakis, Costas, "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of Political Economy*, 2010, *118* (6), 1151--1199.

__ , **Arnaud Costinot, and Andrés Rodríguez-Clare**, "New Trade Models, Same Old Gains?," *NBER Working Paper*, 2009, (w15628).

__ , __ , **and Andres Rodríguez-Clare**, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94--130.

__ , __ , **Dave Donaldson, and Andrés Rodríguez-Clare**, "The elusive pro-competitive effects of trade," *The Review of Economic Studies*, 2019, *86* (1), 46--80.

__ , **Sharat Ganapati, and Marc-Andreas Muendler**, "The Extensive Margin of Exporting Products: A Firm-level Analysis," 2019.

Baqaee, David and Emmanuel Farhi, "Networks, barriers, and trade," Technical Report, National Bureau of Economic Research 2019.

Bartelme, Dominick, Ting Lan, and Andrei Levchenko, "Specialization, Market Access and Medium-Term Growth," 2019.

Bas, Maria, Thierry Mayer, and Mathias Thoenig, "From micro to macro: Demand, supply, and heterogeneity in the trade elasticity," *Journal of International Economics*, 2017, *108*, 1--19.

Bernard, Andrew B. and J. Bradford Jensen, "Exceptional Exporter Performance: Cause, Effect, or Both," *Journal of International Economics*, 1999, *47*, 1--25.

__ , **Stephen J. Redding, and Peter Schott**, "Multi-Product Firms and Trade Liberalization," *Quarterly Journal of Economics*, 2011, *126* (3), 1271--1318.

Bernard, Andrew, Bradford Jensen, Steve Redding, and Peter J. Schott, "FIrms in International Trade," *Journal of Economic Perspectives*, 2007, *21* (3), 105--130.

Berry, Steven T and Philip A Haile, "Identification in differentiated products markets using market level data," *Econometrica*, 2014, *82* (5), 1749--1797.

Blundell, Richard, Xiaohong Chen, and Dennis Kristensen, "Semi-Nonparametric IV Estimation of Shape-Invariant Engel Curves," *Econometrica*, 2007, *75* (6), 1613--1669.

Caliendo, Lorenzo and Fernando Parro, "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 2014.

__ , **Luca David Opromolla, Fernando Parro, and Alessandro Sforza**, "Goods and factor market integration: a quantitative assessment of the eu enlargement," Technical Report, National Bureau of Economic Research 2017.

Canay, Ivan A, Andres Santos, and Azeem M Shaikh, "On the testability of identification in some nonparametric models with endogeneity," *Econometrica*, 2013, *81* (6), 2535--2559.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 2008, *98* (4), 1707--1721.

Chen, Xiaohong and Timothy M Christensen, "Optimal sup-norm rates and uniform inference on nonlinear functionals of nonparametric IV regression," *Quantitative Economics*, 2018, *9* (1), 39--84.

__ **and Yin Jia Jeff Qiu**, "Methods for nonparametric and semiparametric regressions with endogeneity: A gentle guide," *Annual Review of Economics*, 2016, *8*, 259--290.

Compiani, Giovanni, "Market counterfactuals and the specification of multi-product demand: a nonparametric approach," Technical Report, Working paper. 3, 43 2019.

Costinot, Arnaud and Andres Rodriguez-Clare, "Trade theory with numbers: Quantifying the consequences of globalization," Technical Report, National Bureau of Economic Research 2013.

__ **, Dave Donaldson, and Ivana Komunjer**, "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," 2010. Forthcoming, Review of Economic Studies.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, 2008, *55* (3), 511--540.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 2011, *79* (5), 1453--1498.

__ **,** __ **, Brent Neiman, and John Romalis**, "Trade and the global recession," *The American Economic Review*, 2016, *106* (11), 3401--3438.

Fajgelbaum, Pablo D and Amit K Khandelwal, "Measuring the unequal gains from trade," *The Quarterly Journal of Economics*, 2016, *131* (3), 1113--1180.

Feenstra, Robert C, "Restoring the product variety and pro-competitive gains from trade with heterogeneous firms and bounded productivity," *Journal of International Economics*, 2018, *110*, 16--27.

Fernandes, Ana, Peter Klenow, Martha Denisse Periola, Sergii Meleshchuk, and Andres Rodriguez-Clare, "The Intensive Margin in Trade: How Big and How Important?," *Unpublished manuscript*, 2019.

Florens, Jean-Pierre, Jan Johannes, and Sébastien Van Bellegem, "Instrumental regression in partially linear models," *The Econometrics Journal*, 2012, *15* (2), 304--324.

Head, Keith and John Ries, "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, 2001, *91* (4), 858--876.

__ **and Thierry Mayer**, "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook," in Gita Gopinath, Elhanan Helpman, and Kenneth Rogoff, eds., *Handbook of International Economics*, Vol. 4 of *Handbook of International Economics*, Elsevier, 2014, pp. 131 -- 195.

— , — , **and Mathias Thoenig**, "Welfare and trade without pareto," *American Economic Review*, 2014, *104* (5), 310--16.

Helpman, Elhanan, Marc Melitz, and Yona Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2008, *2* (5), 441--487.

Hottman, Colin J, Stephen J Redding, and David E Weinstein, "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 2016, *131* (3), 1291--1364.

Jr, Frank E Harrell, *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, Springer, 2001.

Kehoe, Timothy J. and Kim J. Ruhl, "How Important Is the New Goods Margin in International Trade?," *Journal of Political Economy*, 2013, *121* (2), 358 -- 392.

Kehoe, Timothy J, Jack Rossbach, and Kim J Ruhl, "Using the new products margin to predict the industry-level impact of trade reform," *Journal of International Economics*, 2015, *96* (2), 289--297.

Kleinman, Benny, Ernest Liu, and Stephen J Redding, "International friends and enemies," Technical Report, National Bureau of Economic Research 2020.

Lind, Nelson and Natalia Ramondo, "Trade with correlation," Technical Report, National Bureau of Economic Research 2018.

Matsuyama, Kiminori and Philip Ushchev, "Beyond CES: Three Alternative Cases of Flexible Homothetic Demand Systems," *Buffett Institute Global Poverty Research Lab Working Paper*, 2017, (17-109).

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71* (6), 1695--1725.

Melitz, Marc J and Stephen J Redding, "New trade models, new welfare implications," Technical Report, National Bureau of Economic Research 2013.

— **and** — , "Heterogeneous firms and trade," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 1--54.

— **and** — , "New trade models, new welfare implications," *American Economic Review*, 2015, *105* (3), 1105--46.

Newey, Whitney K and James L Powell, "Instrumental variable estimation of nonparametric models," *Econometrica*, 2003, *71* (5), 1565--1578.

Novy, Dennis, "International trade without CES: Estimating translog gravity," *Journal of International Economics*, 2013, *89* (2), 271--282.

Redding, Stephen J., "Theories of Heterogeneous FIrms and Trade," *Annual Review of Economics*, 2011, *3*, 77--105.

Redding, Stephen J, "Goods trade, factor mobility and welfare," *Journal of International Economics*, 2016, *101*, 148--167.

Ryan, Stephen P, "The costs of environmental regulation in a concentrated industry," *Econometrica*, 2012, *80* (3), 1019--1061.

Silva, J. M. C. Santos and Silvana Tenreyro, "The Log of Gravity," *Review of Economics and Statistics*, 2006, *88* (4), 641--658.

Online Appendix

A Theory Appendix: Proofs and Additional Results

A.1 Proof of Lemma 1

Part 1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, satisfy (11) and (14) for all i and j. Together with N_i, they determine bilateral trade flows, $X_{ij} = N_i n_{ij} \bar{x}_{ij}$.

Part 2. For all i, total spending, E_i, satisfies (8).

Part 3. To derive the labor market clearing condition notice that there are three sources of demand for labor: production of goods, fixed-cost of entering a market and fixed-cost of creating a variety. Thus,

$$w_i \bar{L}_i = \sum_j N_i Pr[\omega \in \Omega_{ij}] \left(1 - \frac{1}{\sigma}\right) E\left[R_{ij}(\omega) | \omega \in \Omega_{ij}\right] + \sum_j N_i Pr[\omega \in \Omega_{ij}] w_i \bar{f}_{ij} E\left[f_{ij}(\omega) | \omega \in \Omega_{ij}\right] + N_i w_i \bar{F}_i$$

From the free entry condition, we know that

$$w_i \bar{F}_i = \sum_j \mathbb{E}\left[\max\{\pi_{ij}(\omega);\ 0\}\right] = \sum_j Pr[\omega \in \Omega_{ij}] \left(\frac{1}{\sigma} E\left[R_{ij}(\omega) | \omega \in \Omega_{ij}\right] - w_i \bar{f}_{ij} E\left[f_{ij}(\omega) | \omega \in \Omega_{ij}\right]\right),$$

which implies that

$$w_i \bar{L}_i = \sum_j N_i Pr[\omega \in \Omega_{ij}] E\left[R_{ij}(\omega) | \omega \in \Omega_{ij}\right].$$

Thus, since $\bar{x}_{ij} \equiv E\left[R_{ij}(\omega) | \omega \in \Omega_{ij}\right]$ and $n_{ij} = Pr[\omega \in \Omega_{ij}]$, this immediately implies

$$w_i \bar{L}_i = \sum_j N_i n_{ij} \bar{x}_{ij}. \tag{OA.1}$$

Thus, the only exogenous element in this expression is $\bar{L}_i$.

Part 4. Since $p_{ij}(\omega) = \frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij} w_i}{\bar{a}_i} \frac{\tau_{ij}(\omega)}{a_i(\omega)}$, the expression for $P_j^{1-\sigma}$ in (2) implies that

$$P_j^{1-\sigma} = \sum_i \left[\bar{b}_{ij} \left(\frac{\sigma}{\sigma - 1} \frac{\bar{\tau}_{ij}}{\bar{a}_i}\right)^{1-\sigma}\right] \left(w_i^{1-\sigma}\right) \int_{\Omega_{ij}} \left(b_{ij}(\omega)\right) \left(\frac{\tau_{ij}(\omega)}{a_i(\omega)}\right)^{1-\sigma} d\omega$$

Using the definitions in (4), we can write this expression as

$$P_j^{1-\sigma} = \sum_i \bar{r}_{ij} \left(w_i^{1-\sigma}\right) \int_{\Omega_{ij}} r_{ij}(\omega)\ d\omega$$

Notice that $\int_{\Omega_{ij}} r_{ij}(\omega)\ d\omega = N_i Pr[\omega \in \Omega_{ij}] E[r | \omega \in \Omega_{ij}] = N_i n_{ij} \rho_{ij}(n_{ij})$. This immediately yields

$$P_j^{1-\sigma} = \sum_i \bar{r}_{ij} w_i^{1-\sigma} \rho_{ij}(n_{ij}) n_{ij} N_i. \tag{OA.2}$$

Thus, the only exogenous elements in this expression are σ, $\bar{r}_{ij}$, and $\rho_{ij}(n)$.

Part 5. We start by writing

$$\begin{aligned}\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] &= Pr[\omega\in\Omega_{ij}]E\left[\pi_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]+Pr[\omega\notin\Omega_{ij}]0\\ &= Pr[\omega\in\Omega_{ij}]\left(\tfrac{1}{\sigma}E\left[R_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]-w_i\bar{f}_{ij}E\left[f_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]\right)\\ &= n_{ij}\left(\tfrac{1}{\sigma}\bar{x}_{ij}-w_i\bar{f}_{ij}E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]\right)\end{aligned}$$

where the second equality follows from the expression for $\pi_{ij}(\omega)=(1/\sigma)R_{ij}(\omega)-w_i\bar{f}_{ij}f_{ij}\left(\omega\right)$, and the third equality follows from the definitions of $\bar{x}_{ij}\equiv E\left[R_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]$ and $e_{ij}(\omega)\equiv r_{ij}(\omega)/f_{ij}(\omega)$.

By defining $e^*_{ij}\equiv\sigma\frac{\bar{f}_{ij}}{\bar{r}_{ij}}\left[\left(\frac{w_i}{P_j}\right)^{\sigma}\frac{P_j}{E_j}\right]$, we can write

$$E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]=\int_{e^*_{ij}}^{\infty}\frac{1}{e}\left[\int_0^{\infty}rdH^r_{ij}\left(r|e\right)\right]\frac{dH^e(e)}{1-H^e(e^*_{ij})}$$

Consider the transformation $n=1-H_{ij}(e)$ such that $e=\bar{\epsilon}_{ij}(n)$. In this case, $dH_{ij}(e)=-dn$ and $n_{ij}=1-H_{ij}(e^*_{ij})$, which implies that

$$E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]=\frac{1}{n_{ij}}\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}dn,$$

where, to simplify exposition, we define the mean revenue potential of firms in quantile n of the entry potential distribution as

$$\rho^m_{ij}(n)\equiv E\left[r|e=\epsilon_{ij}(n)\right]. \tag{OA.3}$$

Thus,

$$\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right]=\frac{1}{\sigma}n_{ij}\bar{x}_{ij}-w_i\bar{f}_{ij}\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Thus, the free entry condition is

$$\sigma w_i\bar{F}_i=\sum_j n_{ij}\bar{x}_{ij}-\sum_j\left(\sigma w_i\bar{f}_{ij}\right)\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn. \tag{OA.4}$$

Notice that the summation of (11) and (14) implies that

$$\ln\left(\sigma w_i\bar{f}_{ij}\right)=\ln\bar{x}_{ij}-\ln\rho_{ij}(n_{ij})+\ln\epsilon_{ij}(n_{ij})$$

which yields

$$\sigma w_i\bar{F}_i=\sum_j n_{ij}\bar{x}_{ij}-\sum_j\bar{x}_{ij}\frac{\epsilon_{ij}(n_{ij})}{\rho_{ij}(n_{ij})}\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

By substituting the definition of $\rho_{ij}(n)$, we can write the free entry condition as

$$\sigma w_i\bar{F}_i=\sum_j n_{ij}\bar{x}_{ij}-\sum_j n_{ij}\bar{x}_{ij}\frac{\epsilon_{ij}(n_{ij})}{\int_0^{n_{ij}}\rho^m_{ij}(n)\ dn}\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn. \tag{OA.5}$$

Using the market clearing condition in (OA.1), we have that

$$\frac{1}{N_i} = \sigma \frac{\bar{F}_i}{\bar{L}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\epsilon_{ij}(n_{ij})}{\int_0^{n_{ij}} \rho_{ij}^m(n)\, dn} \int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn, \tag{OA.6}$$

which immediately yields

$$N_i = \left[\sigma \frac{\bar{F}_i}{\bar{L}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn} \right]^{-1}. \tag{OA.7}$$

Notice that, by (16), $\rho_{ij}^m(n)$ is uniquely determined by $\rho_{ij}(n)$. Thus, the only exogenous elements in this expression are $\sigma \bar{F}_i$, $\bar{L}_i$, $\epsilon_{ij}(n)$, and $\rho_{ij}(n)$.

Part 6. The equilibrium vector $\{n_{ij}, \bar{x}_{ij}, E_i, w_i, P_i, N_i\}_{i,j}$ is determined by equations (11), (14), (8), (OA.1), (OA.2), and (OA.7). The system is thus a function of the vector of country fundamentals $\{\bar{T}_i, \bar{L}_i, \bar{F}_i, \bar{f}_{ij}, \bar{r}_{ij}\}$, the elasticity of substitution σ, and the bilateral functions, $\{\epsilon_{ij}(n), \rho_{ij}(n)\}_{i,j}$.■

A.2 Proof of Proposition 1

Part 1. We start by pointing out that equation (16) implies that knowledge of $\rho_{ij}(n)$ implies knowledge of $\rho_{ij}^m(n)$ (as defined in (OA.3)). We then use the equilibrium conditions in Proposition 1 to obtain a system of equations for the changes in $\{\{n_{ij}, \bar{x}_{ij}\}_j, P_i, N_i, E_i, w_i\}$ given changes in $\left\{\bar{T}_i, \bar{L}_i, \bar{F}_i, \{\bar{r}_{ij}, \bar{f}_{ij}\}_j\right\}_i$.

1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, in (11) and (14) imply

$$\frac{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}{\epsilon_{ij}(n_{ij})} = \frac{\hat{\bar{f}}_{ij}}{\hat{\bar{r}}_{ij}} \left[\left(\frac{\hat{w}_i}{\hat{P}_j} \right)^{\sigma} \frac{\hat{P}_j}{\hat{E}_j} \right], \tag{OA.8}$$

$$\hat{\bar{x}}_{ij} = \hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \left[\left(\frac{\hat{w}_i}{\hat{P}_j} \right)^{1-\sigma} \hat{E}_j \right]. \tag{OA.9}$$

2. Let $\iota_i \equiv w_i L_i / E_i = \left(\sum_d X_{id}\right) / \left(\sum_o X_{oi}\right)$ be the output-spending ratio in country i in the initial equilibrium. The spending equation in (8) implies

$$\hat{E}_i = \iota_i \left(\hat{w}_i \hat{\bar{L}}_i \right) + (1 - \iota_i) \hat{\bar{T}}_i, \tag{OA.10}$$

3. Let $y_{ij} \equiv (N_i n_{ij} \bar{x}_{ij}) / (w_i L_i) = X_{ij} / \left(\sum_{j'} X_{ij'} \right)$ be the share of i's revenue from sales to j. The labor market clearing condition in (OA.1) implies

$$\hat{w}_i \hat{\bar{L}}_i = \sum_j y_{ij} \left(\hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij} \right). \tag{OA.11}$$

4. The price index (OA.2) implies

$$\begin{aligned}
\hat{P}_j^{1-\sigma} &= \sum_i \frac{\bar{r}_{ij} w_i^{1-\sigma} \rho_{ij}(n_{ij}) n_{ij} N_i}{P_j^{1-\sigma}} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i \right) \\
&= \sum_i \frac{\bar{r}_{ij} w_i^{1-\sigma} \rho_{ij}(n_{ij}) n_{ij} N_i E_j P_j^{\sigma-1}}{\sum_o \bar{r}_{oj} w_o^{1-\sigma} \rho_{oj}(n_{oj}) n_{oj} N_o E_j P_j^{\sigma-1}} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i \right) \\
&= \sum_i \frac{\bar{x}_{ij} n_{ij} N_i}{\sum_o \bar{x}_{oj} n_{oj} N_o} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i \right)
\end{aligned}$$

Let $x_{ij} \equiv (N_i n_{ij} \bar{x}_{ij}) / (\sum_o \bar{x}_{oj} n_{oj} N_o) = X_{ij} / (\sum_o X_{oj})$ be the spending share of country j on country i. Thus,

$$\hat{P}_j^{1-\sigma} = \sum_i x_{ij} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i \right). \tag{OA.12}$$

5. The free entry condition in (OA.7) implies

$$N_i \hat{N}_i = \left[\sigma \frac{\bar{F}_i}{\bar{L}_i} \frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\hat{w}_i \hat{\bar{L}}_i} \frac{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\, dn} \right]^{-1}$$

Using (OA.7) to substitute for $\sigma \frac{\bar{F}_i}{\bar{L}_i}$,

$$\hat{N}_i = \left[\left(1 - \sum_j y_{ij} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, d} \right) \frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j y_{ij} \frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\hat{w}_i \hat{\bar{L}}_i} \frac{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\, dn} \right]^{-1}. \tag{OA.13}$$

A.3 Proof of Proposition 2

Part 1. We start by totally differentiating the equilibrium equations in Lemma 1. Equation (11) implies

$$\varepsilon_{ij}(n_{ij}) d\ln n_{ij} = d\ln \bar{f}_{ij} - d\ln \bar{r}_{ij} + \sigma d\ln w_i - (\sigma - 1) d\ln P_j - d\ln E_j \tag{OA.14}$$

Equation (14) yields

$$d\ln \bar{x}_{ij} = d\ln \bar{r}_{ij} + \varrho_{ij}(n_{ij}) d\ln n_{ij} - (\sigma-1) d\ln w_i + (\sigma - 1) d\ln P_j + d\ln E_j.$$

The sum of the two equations above implies that

$$d\ln \bar{x}_{ij} = d\ln \bar{f}_{ij} + d\ln w_i + (\varrho_{ij}(n_{ij}) - \varepsilon_{ij}(n_{ij}))\, d\ln n_{ij} \tag{OA.15}$$

Equation 8 implies that

$$d\ln E_j = \iota_j d\ln w_j + \iota_j d\ln \bar{L}_j + (1 - \iota_j) d\ln \bar{T}_j. \tag{OA.16}$$

By combining the market clearing condition in (OA.1) with the version of the free entry condition in (OA.4), we have that

$$\frac{\sigma \bar{F}_i}{\bar{L}_i} = \frac{1}{N_i} - \sum_j \frac{\sigma \bar{f}_{ij}}{\bar{L}_i} \int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn,$$

which implies that

$$\begin{aligned}
-\tfrac{1}{N_i} d\ln N_i \quad -\tfrac{\sigma \bar{F}_i}{\bar{L}_i} \left(d\ln \bar{F}_i / \bar{L}_i \right) = & \left(\textstyle\sum_j \frac{\sigma \bar{f}_{ij}}{\bar{L}_i} \int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn \right) \left(d\ln \bar{f}_{ij} - d\ln \bar{L}_i \right) + \textstyle\sum_j \frac{\sigma \bar{f}_{ij}}{\bar{L}_i} \frac{\rho_{ij}^m(n_{ij})}{\epsilon_{ij}(n_{ij})} n_{ij} d\ln n_{ij} \\
= & \textstyle\sum_j \left(\frac{\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\epsilon_{ij}(n_{ij})}{\rho_{ij}(n_{ij})} \int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn \right) \left(d\ln \bar{f}_{ij} - d\ln \bar{L}_i \right) + \textstyle\sum_j \frac{\bar{x}_{ij} n_{ij}}{w_i \bar{L}_i} \frac{\rho_{ij}^m(n_{ij})}{\rho_{ij}(n_{ij})} d\ln n_{ij} \\
= & \textstyle\sum_j \left(\frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn} \right) \left(d\ln \bar{f}_{ij} - d\ln \bar{L}_i \right) + \textstyle\sum_j \frac{\bar{x}_{ij} n_{ij}}{w_i \bar{L}_i} \left(1 + \varrho_{ij}(n_{ij}) \right) d\ln n_{ij} \\
= & \tfrac{1}{N_i} \textstyle\sum_j \left(y_{ij} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn} \right) \left(d\ln \bar{f}_{ij} - d\ln \bar{L}_i \right) + \tfrac{1}{N_i} \textstyle\sum_j y_{ij} \left(1 + \varrho_{ij}(n_{ij}) \right) d\ln n_{ij},
\end{aligned}$$

where the second equality uses $\bar{x}_{ij} = \frac{\rho_{ij}(n_{ij})}{\epsilon_{ij}(n_{ij})}\sigma \bar{f}_{ij} w_i$, the third equality uses (16) and (OA.3), and the fourth uses $y_{ij} \equiv N_i \bar{x}_{ij} n_{ij}/w_i \bar{L}_i$.

Thus,

$$d\ln N_i = -\pi_i d\ln \bar{F}_i/\bar{L}_i - \sum_j y_{ij}\pi_{ij}(n_{ij})\left(d\ln \bar{f}_{ij} - d\ln \bar{L}_i\right) - \sum_j y_{ij}\left(1+\varrho_{ij}(n_{ij})\right) d\ln n_{ij} \tag{OA.17}$$

where

$$\pi_{ij}(n_{ij}) \equiv \frac{\int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n_{ij})}\, dn} \tag{OA.18}$$

and, by equation (OA.6),

$$\pi_i \equiv \sigma \frac{\bar{F}_i N_i}{\bar{L}_i} = 1 - \sum_j y_{ij}\pi_{ij}(n_{ij}). \tag{OA.19}$$

Equation (OA.1) implies

$$d\ln w_i + d\ln \bar{L}_i = \sum_j y_{ij}\left(d\ln N_i + d\ln n_{ij} + d\ln \bar{x}_{ij}\right),$$

which combined with (OA.15) implies

$$-\,d\ln N_i + d\ln \bar{L}_i = \sum_j y_{ij} d\ln \bar{f}_{ij} + \sum_j y_{ij}\left(1+\varrho_{ij}(n_{ij}) - \varepsilon_{ij}(n_{ij})\right) d\ln n_{ij}. \tag{OA.20}$$

The combination of this equation and (OA.17) implies that

$$\sum_j y_{ij}\varepsilon_{ij}(n_{ij}) d\ln n_{ij} = -\pi_i d\ln \bar{F}_i + \sum_j y_{ij}\left(1-\pi_{ij}(n_{ij})\right) d\ln \bar{f}_{ij}. \tag{OA.21}$$

Finally, equation (OA.2) implies

$$(1-\sigma)d\ln P_j = \sum_i x_{ij}\left(d\ln \bar{r}_{ij} - (\sigma-1)d\ln w_i + \left(1+\varrho_{ij}(n_{ij})\right) d\ln n_{ij} + d\ln N_i\right) \tag{OA.22}$$

Equations (OA.14), (OA.17), (OA.21) and (OA.22) form a system that determines $\{d\ln n_{ij}, d\ln N_i, d\ln P_i, d\ln w_i\}_{i.j}$ for any arbitrary set of shocks. We now establish Part 1 of Proposition 2 by reducing this system to two sets of equations determining $\{d\ln P_i, d\ln w_i\}_i$ in terms of σ, $\{\theta_{ij}(n_{ij}), n_{ij}, X_{ij}\}_{i,j}$. To this end, note that the definition of $\theta_{ij}(n_{ij})$ in (17) implies that $\frac{1+\varrho_{ij}(n_{ij})}{\varepsilon_{ij}(n_{ij})} = 1 + \frac{\theta_{ij}(n_{ij})}{1-\sigma}$. Thus, equations (OA.22) and (OA.20) imply

$$\begin{array}{rcl} (1-\sigma)d\ln P_j & = & \sum_i x_{ij}\left(d\ln \bar{r}_{ij} - (\sigma-1)d\ln w_i + \varepsilon_{ij}(n_{ij})d\ln n_{ij}\right) \\ & + & \sum_i x_{ij}\left[\left(\frac{\theta_{ij}(n_{ij})}{1-\sigma}\right)\varepsilon_{ij}(n_{ij})d\ln n_{ij} + d\ln N_i\right] \end{array} \tag{OA.23}$$

$$d\ln N_i = d\ln \bar{L}_i - \sum_j y_{ij} d\ln \bar{f}_{ij} + \sum_j y_{ij}\left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\right)\varepsilon_{ij}(n_{ij})d\ln n_{ij}. \tag{OA.24}$$

By substituting the second equation into the first, we get that

$$
\begin{aligned}
(1-\sigma)d\ln P_j \quad = & \quad \textstyle\sum_i x_{ij}\left(d\ln \bar{r}_{ij} - (\sigma-1)d\ln w_i + \varepsilon_{ij}(n_{ij})d\ln n_{ij}\right) \\
- & \quad \textstyle\sum_i x_{ij}\left[\left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\right)\varepsilon_{ij}(n_{ij})d\ln n_{ij} - \sum_d y_{id}\left(\frac{\theta_{id}(n_{id})}{\sigma-1}\right)\varepsilon_{id}(n_{id})d\ln n_{id}\right] \\
+ & \quad \textstyle\sum_i x_{ij}\left(d\ln \bar{L}_i - \sum_d y_{id} d\ln \bar{f}_{id}\right)
\end{aligned}
$$

By substituting (OA.14) into this expression,

$$
\begin{aligned}
d\ln E_j \quad -\textstyle\sum_i x_{ij}d\ln w_i = & \quad \textstyle\sum_i x_{ij}\left(d\ln \bar{L}_i + d\ln \bar{f}_{ij} - \sum_d y_{id}d\ln \bar{f}_{id}\right) \\
- & \quad \textstyle\sum_i x_{ij}\left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\right)\left(d\ln \bar{f}_{ij}/\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - d\ln E_j\right) \\
+ & \quad \textstyle\sum_i x_{ij}\sum_d y_{id}\left(\frac{\theta_{id}(n_{id})}{\sigma-1}\right)\left(d\ln \bar{f}_{id}/\bar{r}_{id} + \sigma d\ln w_i - (\sigma-1)d\ln P_d - d\ln E_d\right)
\end{aligned}
$$

Substituting (OA.16) into this expression,

$$
\begin{aligned}
\iota_j d\ln w_j - \textstyle\sum_i x_{ij}d\ln w_i \quad = & \quad -\iota_j d\ln \bar{L}_j - (1-\iota_j)d\ln \bar{T}_j + \textstyle\sum_i x_{ij}\left(d\ln \bar{L}_i + d\ln \bar{f}_{ij} - \sum_d y_{id}d\ln \bar{f}_{id}\right) \\
- & \quad \textstyle\sum_i x_{ij}\frac{\theta_{ij}(n_{ij})}{\sigma-1}\left(d\ln \bar{f}_{ij}/\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - \iota_j d\ln w_j\right) \\
+ & \quad \textstyle\sum_i x_{ij}\sum_d y_{id}\frac{\theta_{id}(n_{id})}{\sigma-1}\left(d\ln \bar{f}_{id}/\bar{r}_{id} + \sigma d\ln w_i - (\sigma-1)d\ln P_d - \iota_d d\ln w_d\right) \\
+ & \quad \textstyle\sum_i x_{ij}\left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\left(\iota_j d\ln \bar{L}_j + (1-\iota_j)d\ln \bar{T}_j\right) - \sum_d y_{id}\frac{\theta_{id}(n_{id})}{\sigma-1}\left(\iota_d d\ln \bar{L}_d + (1-\iota_d)d\ln \bar{T}_d\right)\right)
\end{aligned}
$$

Thus, for market i,

$$
\sum_j v^p_{ij} d\ln P_j - \sum_j v^w_{ij} d\ln w_j = d\ln r^p_i - d\ln f^p_i + d\ln L^p_i + d\ln T^p_i \tag{OA.25}
$$

$$
\begin{aligned}
v^w_{ij} \quad \equiv & \quad 1[i=j]\left(1 - \textstyle\sum_j \frac{x_{ji}\theta_{ji}(n_{ji})}{\sigma-1}\right)\iota_i \\
+ & \quad \left[\left(\textstyle\sum_o x_{oi}y_{oj}\theta_{oj}(n_{oj})\right)\left(\frac{\iota_j}{\sigma-1}\right) - x_{ji}\left(1 - \frac{\sigma}{\sigma-1}\left(\theta_{ji}(n_{ji}) - \textstyle\sum_d y_{jd}\theta_{jd}(n_{jd})\right)\right)\right]
\end{aligned} \tag{OA.26}
$$

$$
v^p_{ij} \equiv 1[i=j]\left(\sum_o x_{oi}\theta_{oi}(n_{oi})\right) - \left(\sum_o x_{oi}y_{oj}\theta_{oj}(n_{oj})\right) \tag{OA.27}
$$

$$
d\ln r^p_i \equiv \sum_j x_{ji}\left(\frac{\theta_{ji}(n_{ji})}{1-\sigma}d\ln \bar{r}_{ji} - \sum_d y_{jd}\frac{\theta_{jd}(n_{jd})}{1-\sigma}d\ln \bar{r}_{jd}\right) \tag{OA.28}
$$

$$
d\ln f^p_i \equiv \sum_j x_{ji}\left[\left(1 - \frac{\theta_{ji}(n_{ji})}{\sigma-1}\right)d\ln \bar{f}_{ji} - \sum_d y_{jd}\left(1 - \frac{\theta_{jd}(n_{jd})}{\sigma-1}\right)d\ln \bar{f}_{jd}\right] \tag{OA.29}
$$

$$
d\ln L^p_i \equiv \left(1 - \sum_j x_{ji}\frac{\theta_{ji}(n_{ji})}{\sigma-1}\right)\iota_i d\ln \bar{L}_i - \sum_j x_{ji}\left(d\ln \bar{L}_j - \sum_d y_{jd}\frac{\theta_{jd}(n_{jd})}{\sigma-1}\iota_d d\ln \bar{L}_d\right) \tag{OA.30}
$$

$$
d\ln T^p_i \equiv \left(1 - \sum_j x_{ji}\frac{\theta_{ji}(n_{ji})}{\sigma-1}\right)(1-\iota_i)d\ln \bar{T}_i + \sum_j x_{ji}\left(\sum_d y_{jd}\frac{\theta_{jd}(n_{jd})}{\sigma-1}(1-\iota_d)d\ln \bar{T}_d\right) \tag{OA.31}
$$

Equations (OA.21) and (OA.14) imply

$$
\begin{aligned}
\textstyle\sum_j y_{ij}\left(d\ln \bar{f}_{ij} - d\ln \bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j\right) \quad - & \quad \textstyle\sum_j y_{ij}\left(\iota_j d\ln w_j + \iota_j d\ln \bar{L}_j + (1-\iota_j)d\ln \bar{T}_j\right) \\
= & \quad -\pi_i d\ln \bar{F}_i + \textstyle\sum_j y_{ij}\left(1 - \pi_{ij}(n_{ij})\right)d\ln \bar{f}_{ij}
\end{aligned}
$$

Thus,

$$\begin{aligned}\sigma d\ln w_i - \textstyle\sum_i y_{ij}\iota_j d\ln w_j - \sum_j y_{ij}(\sigma-1)d\ln P_j \quad &= \quad \textstyle\sum_j y_{ij} d\ln \bar{r}_{ij} - \sum_j y_{ij}\pi_{ij}(n_{ij}) d\ln \bar{f}_{ij} \\ &\quad -\pi_i d\ln \bar{F}_i + \textstyle\sum_j y_{ij}(1-\iota_j) d\ln \bar{T}_j \\ &\quad + \textstyle\sum_j y_{ij}\iota_j d\ln \bar{L}_j\end{aligned}$$

Thus,

$$\sum_j m_{ij}^w d\ln w_j - \sum_j m_{ij}^p d\ln P_j = d\ln r_i^w - d\ln f_i^w + d\ln F_i^w + d\ln T_i^w + d\ln L_i^w \tag{OA.32}$$

$$m_{ij}^w \equiv 1[i=j]\sigma - y_{ij}\iota_j \tag{OA.33}$$

$$m_{ij}^p \equiv y_{ij}(\sigma-1) \tag{OA.34}$$

$$d\ln r_i^w \equiv \sum_j y_{ij} d\ln \bar{r}_{ij}, \quad \text{and} \quad d\ln f_i^w \equiv \sum_j y_{ij}\pi_{ij}(n_{ij}) d\ln \bar{f}_{ij} \tag{OA.35}$$

$$d\ln F_i^w \equiv -\pi_i d\ln \bar{F}_i, \quad d\ln T_i^w \equiv \sum_j y_{ij}(1-\iota_j) d\ln \bar{T}_j, \quad d\ln L_i^w \equiv \sum_j y_{ij}\iota_j d\ln \bar{L}_j \tag{OA.36}$$

Let us use bold letters to denote vectors, $\boldsymbol{v} = [v_i]_i$ and bold bar variables to denote matrices, $\bar{\boldsymbol{v}} = [v_{ij}]_{i,j}$. Thus, equations (OA.25)–(OA.32) imply

$$\begin{aligned}\bar{\boldsymbol{v}}^p d\ln \boldsymbol{P} - \bar{\boldsymbol{v}}^w d\ln \boldsymbol{w} \quad &= \quad d\ln \boldsymbol{\psi}^p \\ -\bar{\boldsymbol{m}}^p d\ln \boldsymbol{P} + \bar{\boldsymbol{m}}^w d\ln \boldsymbol{w} \quad &= \quad d\ln \boldsymbol{\psi}^w\end{aligned}$$

where

$$\begin{aligned}d\ln \psi_i^p \quad &\equiv \quad d\ln r_i^p - d\ln f_i^p + d\ln L_i^p + d\ln T_i^p \\ d\ln \psi_i^w \quad &\equiv \quad d\ln r_i^w - d\ln f_i^w + d\ln L_i^w + d\ln T_i^w + d\ln F_i^w\end{aligned}$$

We then use the first equation to solve for the price index change,

$$d\ln \boldsymbol{P} = (\bar{\boldsymbol{v}}^p)^{-1}(\bar{\boldsymbol{v}}^w d\ln \boldsymbol{w} + d\ln \boldsymbol{\psi}^p), \tag{OA.37}$$

which we then substitute into the second equation to obtain,

$$\left[\bar{\boldsymbol{m}}^w - \bar{\boldsymbol{m}}^p (\bar{\boldsymbol{v}}^p)^{-1} \bar{\boldsymbol{v}}^w\right] d\ln \boldsymbol{w} = d\ln \boldsymbol{\psi}^w + \bar{\boldsymbol{m}}^p (\bar{\boldsymbol{v}}^p)^{-1} d\ln \boldsymbol{\psi}^p. \tag{OA.38}$$

Notice that, because of the numeraire choice, solving (OA.38) requires dropping one row and one column by setting $d\ln w_n = 0$ for some arbitrary country n.

Recall that $\{X_{ij}\}_{ij}$ immediately yields $\{\iota_j, x_{ij}, y_{ij}\}_{i,j}$. Thus, the system (OA.37)–(OA.38) determines $\{d\ln P_i, d\ln w_i\}_i$ as a function of shocks. Since $(\bar{\boldsymbol{m}}^w, \bar{\boldsymbol{m}}^p, \bar{\boldsymbol{v}}^w, \bar{\boldsymbol{v}}^p)$ and $(d\ln \boldsymbol{r}^p, d\ln \boldsymbol{r}^w)$ depend only on σ, $\{\theta_{ij}(n_{ij}), X_{ij}\}_{i,j}$, and $\{d\ln \bar{r}_{ij}\}_{i,j}$, $\frac{d\ln w_i}{d\ln \bar{r}_{od}}$ and $\frac{d\ln P_i}{d\ln \bar{r}_{od}}$ are functions of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$. By (OA.16), $\frac{d\ln E_i}{d\ln \bar{r}_{od}}$ is also a function of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$. For the other shocks, it is also necessary to know the share of country i's labor force allocated to cover fixed costs of exporting to j, $\{\pi_{ij}(n_{ij}^0)\}_{i,j}$, which immediately yields π_i as defined in (OA.19).

To obtain changes in the number of entrants, we combine equations (OA.24) and (OA.14):

$$d\ln N_i = d\ln \bar{L}_i - \sum_j y_{ij} d\ln \bar{f}_{ij} + \sum_j y_{ij}\left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\right)\left(d\ln \bar{f}_{ij}/\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - d\ln E_j\right). \tag{OA.39}$$

This implies that $d\ln N_i$ is a function of $\left\{d\ln \bar{f}_{ij}, d\ln \bar{r}_{ij}\right\}_j$, σ, $\{\theta_{ij}(n_{ij})\}_j$, $\{d\ln P_j, d\ln w_j, d\ln E_j\}_j$, and $\{X_{ij}\}_{ij}$. Thus, given that $\frac{d\ln w_i}{d\ln \bar{r}_{od}}$, $\frac{d\ln P_i}{d\ln \bar{r}_{od}}$ and $\frac{d\ln E_i}{d\ln \bar{r}_{od}}$ are functions of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$, $\frac{d\ln N_i}{d\ln \bar{r}_{od}}$ is a function of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$.

Finally,

$$\begin{aligned} d\ln X_{ij} &= d\ln N_i + d\ln n_{ij} + d\ln \bar{x}_{ij} \\ &= d\ln N_i + d\ln w_i + (1 + \varrho_{ij}(n_{ij}) - \varepsilon_{ij}(n_{ij}))\, d\ln n_{ij} \\ &= d\ln N_i + d\ln \bar{f}_{ij} + d\ln w_i - \theta_{ij}(n_{ij})\tfrac{\varepsilon_{ij}(n_{ij})}{\sigma-1} d\ln n_{ij} \end{aligned}$$

where the first equality follows from $X_{ij} \equiv N_i n_{ij}\bar{x}_{ij}$, the second equality follows from (OA.15), and the third equality follows from the definition of $\theta_{ij}(n_{ij})$ in (17).

Using (OA.14),

$$d\ln X_{ij} = d\ln N_i + d\ln \bar{f}_{ij} + d\ln w_i - \left(\frac{\theta_{ij}(n_{ij})}{\sigma-1}\right)\left(d\ln \bar{f}_{ij}/\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - d\ln E_j\right) \tag{OA.40}$$

Thus, since $\{\frac{d\ln w_i}{d\ln \bar{r}_{od}}, \frac{d\ln P_i}{d\ln \bar{r}_{od}}, \frac{d\ln E_i}{d\ln \bar{r}_{od}}, \frac{d\ln N_i}{d\ln \bar{r}_{od}}\}_j$ are functions of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$, $\frac{d\ln X_{ij}}{d\ln \bar{r}_{od}}$ is a function of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$.

Part 2. Equation (OA.14) immediately implies that

$$d\ln n_{ij} = \frac{1}{\varepsilon_{ij}(n_{ij})}\left(d\ln \bar{f}_{ij} - d\ln \bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - d\ln E_j\right)$$

Since $\{\frac{d\ln w_i}{d\ln \bar{r}_{od}}, \frac{d\ln P_i}{d\ln \bar{r}_{od}}, \frac{d\ln E_i}{d\ln \bar{r}_{od}}\}_j$ are functions of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$, then $\frac{d\ln n_{ij}}{d\ln \bar{r}_{od}}$ is a function of $(\sigma, \{\theta_{km}(n_{km})\}_{km}, \{X_{km}\}_{km})$ and $\varepsilon_{ij}(n_{ij})$.

A.4 Proof of Proposition 3

Assume that we observe trade outcomes in the initial equilibrium, $\{n^0_{ij}, X^0_{ij}\}_{i,j}$, and changes in wages and trade outcomes between two equilibria, $\{\hat{w}^t_i, \hat{n}^t_{ij}, \hat{\bar{x}}^t_{ij}, \hat{X}^t_{ij}\}_{i,j}$. We can immediately compute $y^0_{ij} = X^0_{ij}/\sum_d X^0_{id}$, $\iota^0_i = \left(\sum_d X^0_{id}\right)/\left(\sum_o X^0_{oi}\right)$, $\hat{N}^t_i = \hat{X}^t_{ij}/\hat{\bar{x}}^t_{ij}\hat{n}^t_{ij}$, $\hat{\bar{L}}_i = \sum_j y^0_{ij}\hat{X}^t_{ij}/\hat{w}^t_i$, $\hat{E}^t_i = \sum_i x^0_{ij}\hat{X}^t_{ij}$, $\hat{\iota}_i = \hat{w}^t_i\hat{\bar{L}}^t_i/\hat{E}^t_i$ and $\hat{T}^t_i = \hat{E}^t_i(1-\iota^0_i\hat{\iota}_i)/(1-\iota^0_i)$.

The first step of the proof is to establish the set of fundamental changes that rationalizes changes in bilateral trade outcomes and observed wages. In the main text we discuss the identification of $\hat{\bar{f}}_{ij}$ and $\hat{\bar{r}}^t_{ij} \equiv \hat{\bar{r}}^t_{ij}/\hat{\bar{r}}^t_{jj}$ – see equations (21) and (22). Given $\{n^0_{ij}, y^0_{ij}\}_j$ and $\{\hat{w}^t_i, \hat{n}^t_{ij}, \hat{\bar{x}}^t_{ij}, \hat{X}^t_{ij}\}_j$, equation (OA.13) for origin i implies that

$$\hat{\bar{F}}^t_i = \left[\frac{\hat{\bar{L}}^t_i}{\hat{N}^t_i} - \sum_j y^0_{ij}\frac{\hat{n}^t_{ij}\hat{\bar{x}}^t_{ij}}{\hat{w}^t_i}\frac{\int_0^{n^0_{ij}\hat{n}^t_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n^0_{ij}\hat{n}^t_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n^0_{ij}\hat{n}^t_{ij})}\, dn}\right]\left(1 - \sum_j y^0_{ij}\frac{\int_0^{n^0_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n^0_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n^0_{ij})}\, dn}\right)^{-1} \tag{OA.41}$$

The second step of the proof is to show that, given $\{n^0_{ij}, X^0_{ij}\}_{i,j}$ in the initial equilibrium, the observed changes $\{\hat{w}_i^t, \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{X}_{ij}^t\}_{i,j}$ are the equilibrium changes in our economy when

$$\hat{\bar{f}}_{ij} = \hat{\bar{f}}_{ij}^t \ \forall i,j; \quad \hat{\bar{r}}_{ij} = \hat{\bar{r}}_{ij}^t \ \forall i \neq j; \quad \hat{\bar{r}}_{ii} = 1 \ \forall i; \quad \hat{\bar{F}}_i = \hat{\bar{F}}_i^t \ \forall i; \quad \hat{\bar{L}}_i = \hat{\bar{L}}_i^t; \ \hat{\bar{T}}_i = \hat{T}_i \forall i. \tag{OA.42}$$

To this end, we use (OA.8)–(OA.13) to write the conditions that determine the equilibrium vector $\{\hat{w}_i, \hat{n}_{ij}, \hat{\bar{x}}_{ij}, \hat{N}_i, \hat{E}_j\}_{i,j}$ for the set of exogenous shocks in OA.42. By substituting $\hat{\bar{f}}_{ij}^t$ in expression (21) into (OA.8) and (OA.9), $\hat{n}_{ij}$ and $\hat{\bar{x}}_{ij}$ must satisfy

$$\frac{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij})}{\epsilon_{ij}(n_{ij}^0)} \frac{\hat{\bar{x}}_{ij}}{\hat{w}_i} \frac{\rho_{ij}(n_{ij}^0)}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij})} = \frac{\hat{\bar{x}}_{ij}^t}{\hat{w}_i^t} \frac{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij}^t)}{\epsilon_{ij}(n_{ij}^0)} \frac{\rho_{ij}(n_{ij}^0)}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)} \tag{OA.43}$$

By substituting $\hat{\bar{r}}_{ij} = \hat{\bar{r}}_{ij}^t$ in expression (22) into the ratio of equation (OA.9) for pairs (i,j) and (j,j), $\hat{\bar{x}}_{ij}/\hat{\bar{x}}_{jj}$ must satisfy

$$\frac{\hat{\bar{x}}_{ij}}{\hat{\bar{x}}_{jj}} \left(\frac{\hat{w}_i}{\hat{w}_j}\right)^{\sigma-1} \frac{\rho_{ij}(n_{ij}^0)}{\rho_{jj}(n_{jj}^0)} \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj})}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij})} = \frac{\hat{\bar{x}}_{ij}^t}{\hat{\bar{x}}_{jj}^t} \left(\frac{\hat{w}_i^t}{\hat{w}_j^t}\right)^{\sigma-1} \frac{\rho_{ij}(n_{ij}^0)}{\rho_{jj}(n_{jj}^0)} \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}^t)}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)} \tag{OA.44}$$

By substituting $\hat{\bar{r}}_{ij} = \hat{\bar{r}}_{ij}^t$ in expression (22) and $\hat{\bar{r}}_{jj} = 1$ into equation (OA.9) for (j,j), $\hat{\bar{x}}_{jj}$ must satisfy

$$\hat{\bar{x}}_{jj} = \hat{\bar{x}}_{jj}^t \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}) \left[(\hat{w}_j)^{1-\sigma} \hat{E}_j\right]}{\sum_i x_{ij}^0 \left(\hat{\bar{x}}_{ij}^t \left(\frac{\hat{w}_i^t}{\hat{w}_j^t}\right)^{\sigma-1} \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}^t)}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)} \rho_{ij}(n_{ij}^0 \hat{n}_{ij}) \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i\right)} \tag{OA.45}$$

By substituting $\hat{\bar{F}}_{ij}$ in expression (OA.41) into (OA.13), $\hat{N}_i$ must satisfy

$$\hat{N}_i = \left[\frac{1}{\hat{N}_i^t} - \sum_j y_{ij}^0 \frac{\hat{n}_{ij}^t \hat{\bar{x}}_{ij}^t}{\hat{w}_i^t \hat{\bar{L}}_i^t} \frac{\int_0^{n_{ij}^0 \hat{n}_{ij}^t} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}^0 \hat{n}_{ij}^t} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij}^t)}\, dn} + \sum_j y_{ij}^0 \frac{\hat{n}_{ij} \hat{\bar{x}}_{ij}}{\hat{w}_i \hat{\bar{L}}_i^t} \frac{\int_0^{n_{ij}^0 \hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}^0 \hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij})}\, dn}\right]^{-1} \tag{OA.46}$$

From (OA.10) and (OA.11), $\hat{E}_i$ and $\hat{w}_i$ must satisfy

$$\hat{E}_i = \iota_i^0 \left(\hat{w}_i \hat{\bar{L}}_i^t\right) + (1-\iota_i^0) \hat{\bar{T}}_i^t, \tag{OA.47}$$

$$\hat{w}_i \hat{\bar{L}}_i = \sum_j y_{ij}^0 \left(\hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij}\right). \tag{OA.48}$$

We now verify that the system (OA.43)–(OA.48) is satisfied by $\{\hat{w}_i^t, \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{N}_i^t, \hat{E}_j^t\}_{i,j}$ where $\hat{N}_i^t = \hat{X}_{ij}^t / \hat{\bar{x}}_t^t \hat{n}_{ij}^t$ and $\hat{E}_i^t = \sum_i x_{ij}^0 \hat{X}_{ij}^t$. It is straight forward to check that (OA.43), (OA.44) and (OA.46) are satisfied for $\{\hat{w}_i^t, \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{N}_i^t, \hat{E}_j^t\}$. Since $\hat{E}_i^t = \sum_i x_{ij}^0 \hat{X}_{ij}^t$, equation (OA.45) holds:

$$\hat{\bar{x}}_{jj}^t \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}^t) \left[\left(\hat{w}_j^t\right)^{1-\sigma} \hat{E}_j^t\right]}{\sum_i x_{ij}^0 \left(\hat{\bar{x}}_{ij}^t \left(\frac{\hat{w}_i^t}{\hat{w}_j^t}\right)^{\sigma-1} \frac{\rho_{jj}(n_{jj}^0 \hat{n}_{jj}^t)}{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)} \rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t) \left(\hat{w}_i^t\right)^{1-\sigma} \hat{n}_{ij}^t \hat{N}_i^t\right)} = \hat{\bar{x}}_{jj}^t \frac{\hat{E}_j^t}{\sum_i x_{ij}^0 \left(\hat{\bar{x}}_{ij}^t \hat{n}_{ij}^t \hat{N}_i^t\right)} = \hat{\bar{x}}_{jj}^t.$$

Using the definitions of $\hat{\bar{L}}_i^t$ and $\hat{\bar{T}}_i^t$, we can also show that equations (OA.47) and (OA.48) hold:

$$\iota_i^0 \hat{w}_i^t \hat{\bar{L}}_i^t + (1-\iota_i^0)\hat{\bar{T}}_i^t = \iota_i^0 \hat{w}_i^t \hat{\bar{L}}_i^t + (1-\iota_i^0)\frac{\hat{E}_i^t}{1-\iota_i^0}\left(1-\iota_i^0 \frac{\hat{w}_i^t \hat{\bar{L}}_i^t}{\hat{E}_i^t}\right) = \hat{E}_i^t$$

$$\sum_j y_{ij}^0 \left(\hat{N}_i^t \hat{n}_{ij}^t \hat{\bar{x}}_{ij}^t\right) = \sum_j y_{ij}^0 \hat{X}_{ij}^t = \hat{w}_i^t \frac{\sum_j y_{ij}^0 \hat{X}_{ij}^t}{\hat{w}_i^t} = \hat{w}_i^t \hat{\bar{L}}_i.$$

Notice that, by the definition of $\hat{N}_i^t$, $\hat{X}_{ij}^t = \hat{N}_i^t \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t$. Thus, given $\{n_{ij}^0, X_{ij}^0\}_{i,j}$ in the initial equilibrium, $\{\hat{w}_i^t, \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{X}_{ij}^t\}_{i,j}$ is an equilibrium vector of outcome changes implied by the set of exogenous shocks in OA.42.

Finally, we establish the second part of the proposition. Suppose that we also observe the change in the price index for country j, $\hat{P}_j^t$. We now derive the change in the domestic revenue shifter $\hat{\bar{r}}_{jj}^t$ that generates $\hat{P}_j^t$ given observed trade outcomes in the initial equilibrium $\{n_{ij}^0, X_{ij}^0\}_{i,j}$ and observed changes in wages and trade outcomes between two equilibria, $\{\hat{w}_i^t, \hat{n}_{ij}^t, \hat{\bar{x}}_{ij}^t, \hat{X}_{ij}^t\}_{i,j}$. Combining equation (OA.12) and $\hat{n}_{ij}^t \hat{N}_i^t = \hat{X}_{ij}^t / \hat{\bar{x}}_{ij}^t$, we get that

$$\hat{\bar{r}}_{jj}^t = \frac{\left(\hat{P}_j^t\right)^{1-\sigma}}{\sum_i x_{ij}^0 \left(\hat{\tilde{r}}_{ij}^t \frac{\rho_{ij}(n_{ij}^0 \hat{n}_{ij}^t)}{\rho_{ij}(n_{ij}^0)} (\hat{w}_i^t)^{1-\sigma} \frac{\hat{X}_{ij}^t}{\hat{\bar{x}}_{ij}^t}\right)} \tag{OA.49}$$

where $\hat{\tilde{r}}_{jj}^t = 1$, $\hat{\tilde{r}}_{ij}^t \equiv \hat{\bar{r}}_{ij}^t / \hat{\bar{r}}_{jj}^t$ is given by (22) for $i \neq j$.

A.5 Proof of the expressions in Section 3.3

Equation (23). If $T_i = 0$, then $\hat{E}_j = \hat{w}_j$ and equation (11) for $i = j$ is

$$\frac{\epsilon_{ii}(n_{ii}\hat{n}_{ii})}{\epsilon_{ii}(n_{ii})} = \frac{\hat{\bar{f}}_{ii}}{\hat{\bar{r}}_{ii}} \left(\frac{w_i}{P_i}\right)^{\sigma-1},$$

which immediately yields the expression in (23).

Equation (24). For the case of balanced trade with $\iota_i = 1$, equation (11) implies that

$$\varepsilon_{ii}(n_{ii}) d\ln n_{ii} = d\ln \bar{f}_{ii}/\bar{r}_{ii} + (\sigma-1) d\ln w_i/P_i.$$

Equation (14) implies that

$$d\ln \bar{x}_{ii} = d\ln \bar{r}_{ii} + \varrho_{ii}(n_{ii}) d\ln n_{ii} - (\sigma-1) d\ln w_i/P_i + d\ln \hat{E}_i.$$

By summing these expressions, we get that

$$d\ln \bar{x}_{ii} = d\ln \bar{f}_{ii} + \left(\varrho_{ii}(n_{ii}) - \varepsilon_{ii}(n_{ii})\right) d\ln n_{ii} + d\ln \hat{E}_i$$

Thus,

$$\begin{aligned}
d\ln x_{ii} &= d\ln N_i + d\ln n_{ii} + d\ln \bar{x}_{ii} - d\ln E_i \\
&= d\ln \bar{f}_{ii} + d\ln N_i + (1+\varrho_{ij}(n_{ij}) - \varepsilon_{ij}(n_{ij}))\, d\ln n_{ij} \\
&= d\ln \bar{f}_{ii} + d\ln N_i + \tfrac{(1+\varrho_{ij}(n_{ij})-\varepsilon_{ij}(n_{ij}))}{\varepsilon_{ij}(n_{ij})}\varepsilon_{ij}(n_{ij}) d\ln n_{ij}
\end{aligned}$$

Using the fact that $\varepsilon_{ii}(n_{ii}) d\ln n_{ii} = d\ln \bar{f}_{ii}/\bar{r}_{ii} + (\sigma-1) d\ln w_i/P_i$,

$$d\ln x_{ii}/N_i\bar{f}_{ii} = -\left[(\sigma-1)\left(1-\frac{1+\varrho_{ij}(n_{ij})}{\varepsilon_{ij}(n_{ij})}\right)\right]\left(d\ln w_i/P_i + \frac{1}{\sigma-1} d\ln \bar{f}_{ii}/\bar{r}_{ii}\right)$$

By the definition of $\theta_{ii}(n_{ii})$ in (17),

$$d\ln x_{ii}/N_i\bar{f}_{ii} = -\theta_{ii}(n_{ii})\left(d\ln w_i/P_i + \frac{1}{\sigma-1} d\ln \bar{f}_{ii}/\bar{r}_{ii}\right),$$

which immediately yields the expression in (24).

The constant elasticity benchmark.
Assume that that $\varepsilon_{ij}(n) = \varepsilon_i$ and $\varrho_{ij}(n) = \varrho_i$ for all n, i and j. By the definition of θ_{ij}, we immediately get that $\theta_{ij} \equiv \theta_i = (\sigma-1)\left(1-\frac{1+\varrho_i}{\varepsilon_i}\right)$. By equation (16), $\rho_i^m(n) = (1+\varrho_i(n))\,\rho_i(n)$ and, therefore, $\rho_i^m(n) = (1+\varrho_i)\, n^{\varrho_i}$. Consider the free entry condition in equation (OA.5):

$$\begin{aligned}
\sigma w_i \bar{F}_i &= \textstyle\sum_j n_{ij}\bar{x}_{ij}\left(1-\frac{\int_0^{n_{ij}} n^{\varrho_i-\varepsilon_i}\, dn}{n_{ij}^{-\varepsilon_i}\int_0^{n_{ij}} n^{\varrho_i}\, dn}\right) \\
&= \textstyle\sum_j n_{ij}\bar{x}_{ij}\left(1-\frac{1+\varrho_i}{1+\varrho_i-\varepsilon_i}\frac{n^{1+\varrho_i-\varepsilon_i}}{n^{1+\varrho_i-\varepsilon_i}}\right) = \left(\frac{-\varepsilon_i}{1+\varrho_i-\varepsilon_i}\right)\sum_j n_{ij}\bar{x}_{ij}.
\end{aligned}$$

The market clearing condition in (OA.1) implies that $\sum_j n_{ij}\bar{x}_{ij} = w_i\bar{L}_i/N_i$ and, therefore, $N_i = \frac{\bar{L}_i}{\sigma\bar{F}_i}\left(\frac{-\varepsilon_i}{1+\varrho_i-\varepsilon_i}\right)$.

A.6 Gains from Trade

We now compute the gains from trade in our model. We assume that $\hat{\bar{\tau}}_{ij} \to \infty$ for all $i \neq j$, and that $\hat{\bar{a}}_i = \hat{\bar{F}}_i = \hat{\bar{f}}_{ij} = \hat{\bar{\tau}}_{ii} = \hat{\bar{L}}_i = 1$ for all i and j.

Corollary 1. *Consider an economy moving from the trade equilibrium to the autarky equilibrium with $\hat{T}_i^A = 0$. The change in the real wage is given by (23) where $\hat{n}_{ii}^A$ and $\hat{N}_i^A$ solve*

$$\frac{\epsilon_{ii}\left(n_{ii}\hat{n}_{ii}^A\right)}{\epsilon_{ii}\left(n_{ii}\right)} = \left(\frac{x_{ii}}{\iota_i}\right)\left(\hat{n}_{ii}^A\hat{N}_i^A\right)\frac{\rho_{ii}\left(n_{ii}\hat{n}_{ii}^A\right)}{\rho_{ii}\left(n_{ii}\right)}, \tag{OA.50}$$

$$\left(1-\sum_j \frac{X_{ij}}{\sum_{j'} X_{ij'}}\frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, d}\right)\hat{N}_i^A = 1 - \frac{\int_0^{n_{ii}\hat{n}_{ii}^A} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ii}\hat{n}_{ii}^A} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ii}\hat{n}_{ii}^A)}\, dn}. \tag{OA.51}$$

In order to compute the gains from trade using (23), we need to compute changes in n_{ii} when the economy moves to autarky (i.e, $\hat{\bar{\tau}}_{ij} \to \infty$ for all $i \neq j$). Equation (OA.50) captures the change in the profitability of the domestic market that determines the change in the domestic survival rate of firms (given $\hat{N}_i^A$). Equation (OA.51) is the free entry condition that determines the change in the number of entrants when the country moves to autarky. The left hand size of (OA.51) is the profit/revenue ratio (inclusive of entry costs) that firms have in different markets in the initial equilibrium. The right hand size is the profit/revenue ratio that entrants have in the domestic market in the autarky equilibrium.

A.6.1 Proof of Corollary 1

To simplify the notation, we drop the superscript A and use "hat" variables to denote the change from the initial equilibrium to the autarky equilibrium. We assume that $\hat{\bar{\tau}}_{ij} \to \infty$ for all $i \neq j$, and that $\hat{\bar{a}}_i = \hat{\bar{F}}_i = \hat{\bar{f}}_{ij} = \hat{\bar{\tau}}_{ii} = \hat{\bar{L}}_i = 1$ for all i and j. We set the wage of i to be the numerarie, $w_i \equiv 1$, so that $\hat{w}_i = 1$. Equation (OA.12) implies that

$$\left(\hat{P}_i\right)^{1-\sigma} = x_{ii} \frac{\rho_{ii}\left(n_{ii}\hat{n}_{ii}\right)}{\rho_{ii}\left(n_{ii}\right)} \left(\hat{n}_{ii}\hat{N}_i\right) \tag{OA.52}$$

From equation (OA.8), we get that, for all $i \neq j$, $\epsilon_{ij}\left(n_{ij}\hat{n}_{ij}\right) \to \infty$ and, therefore, $\hat{n}_{ij} = 0$. In addition, it implies that

$$\frac{\epsilon_{ii}\left(n_{ii}\hat{n}_{ii}\right)}{\epsilon_{ii}\left(n_{ii}\right)} = \frac{\left(\hat{P}_i\right)^{1-\sigma}}{\hat{E}_i} \tag{OA.53}$$

Using the fact that $\hat{E}_i = \iota_i$, (OA.52) and (OA.53) imply that

$$\frac{\epsilon_{ii}\left(n_{ii}\hat{n}_{ii}\right)}{\epsilon_{ii}\left(n_{ii}\right)} = \frac{x_{ii}}{\iota_i} \frac{\rho_{ii}\left(n_{ii}\hat{n}_{ii}\right)}{\rho_{ii}\left(n_{ii}\right)} \left(\hat{n}_{ii}\hat{N}_i\right). \tag{OA.54}$$

From expression (OA.13),

$$\hat{N}_i = \left[\left(1 - \sum_j y_{ij} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn}\right) + y_{ii}\hat{n}_{ii}\hat{\bar{x}}_{ii} \frac{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ii}\hat{n}_{ii})}\, dn}\right]^{-1}$$

$$1 = \hat{N}_i \left(1 - \sum_j y_{ij} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn}\right) + y_{ii}\hat{N}_i\hat{n}_{ii}\hat{\bar{x}}_{ii} \frac{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ii}\hat{n}_{ii})}\, dn}$$

Recall that $\hat{x}_{ii} = \frac{\hat{N}_i\hat{n}_{ii}\hat{\bar{x}}_{ii}}{\hat{E}_i} = 1/x_{ii}$. Thus, $y_{ii}\hat{N}_i\hat{n}_{ii}\hat{\bar{x}}_{ii} = y_{ii}\frac{\iota_i}{x_{ii}} = \frac{X_{ii}}{w_i\bar{L}_i}\frac{E_i}{X_{ii}}\frac{w_i\bar{L}_i}{E_i} = 1$ and, therefore,

$$\left(1 - \sum_j y_{ij} \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn}\right) \hat{N}_i = 1 - \frac{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ii}\hat{n}_{ii}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ii}\hat{n}_{ii})}\, dn}.$$

A.7 Derivation of Equation (31)

By plugging (28) into (11)–(14) we have that

$$\ln \epsilon_{ij}\left(n_{ij}\right) = z_{ij}/\kappa^{\epsilon} + \left(\tilde{\sigma}\eta_{ij}^{\tau} + \eta_{ij}^{f}\right) + \left[\ln \sigma w_i \left(\frac{\sigma}{\sigma - 1}\frac{w_i}{\bar{a}_i}\right)^{\sigma-1} + \tilde{\sigma}\delta_i^{\tau} + \delta_i^{f}\right] - \left[\ln\left(P_j^{\sigma-1}E_j\right) - \tilde{\sigma}\zeta_j^{\tau} - \zeta_j^{f}\right]$$

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = -\tilde{\sigma}\kappa^{\tau} z_{ij} - \tilde{\sigma}\eta^{\tau}_{ij} + \left[\ln \left(\frac{\sigma}{\sigma-1} \frac{w_i}{\bar{a}_i} \right)^{1-\sigma} - \tilde{\sigma}\delta^{\tau}_j \right] + \left[\ln \left(P_j^{\sigma-1} E_j \right) - \tilde{\sigma}\zeta^{\tau}_j \right]$$

where $\tilde{\sigma} \equiv \sigma - 1$ and $\kappa^{\epsilon} \equiv 1/(\tilde{\sigma}\kappa^{\tau} + \kappa^{f})$.

This implies that

$$\begin{aligned} -\kappa^{\epsilon} \left(\tilde{\sigma}\eta^{\tau}_{ij} + \eta^{f}_{ij} \right) \quad &= \quad z_{ij} - \kappa^{\epsilon} \ln \epsilon_{ij}(n_{ij}) + \kappa^{\epsilon} \left[\ln \sigma w_i \left(\tfrac{\sigma}{\sigma-1} \tfrac{w_i}{\bar{a}_i} \right)^{\sigma-1} + \tilde{\sigma}\delta^{\tau}_i + \delta^{f}_i \right] \\ &\quad - \kappa^{\epsilon} \left[\ln \left(P_j^{\sigma-1} E_j \right) - \tilde{\sigma}\zeta^{\tau}_j - \zeta^{f}_j \right], \end{aligned}$$

$$-\tilde{\sigma}\eta^{\tau}_{ij} = \ln \bar{x}_{ij} + \tilde{\sigma}\kappa^{\tau} z_{ij} - \ln \rho_{ij}(n_{ij}) - \left[\ln \left(\frac{\sigma}{\sigma-1} \frac{w_i}{\bar{a}_i} \right)^{1-\sigma} - \tilde{\sigma}\delta^{\tau}_j \right] - \left[\ln \left(P_j^{\sigma-1} E_j \right) - \tilde{\sigma}\zeta^{\tau}_j \right].$$

We can then write

$$\begin{aligned} v^{\epsilon}_{ij} &= z_{ij} - \kappa^{\epsilon} \ln \epsilon_{ij}(n_{ij}) - \delta^{\epsilon}_i - \zeta^{\epsilon}_j \\ v^{\rho}_{ij} &= \ln \bar{x}_{ij} + \tilde{\sigma}\kappa^{\tau} z_{ij} - \ln \rho_{ij}(n_{ij}) - \delta^{\rho}_i - \zeta^{\epsilon}_j \end{aligned} \tag{OA.55}$$

where

$$v^{\epsilon}_{ij} \equiv -\kappa^{\epsilon} \left(\tilde{\sigma}\eta^{\tau}_{ij} + \eta^{f}_{ij} \right) \quad \text{and} \quad v^{\rho}_{ij} \equiv -\tilde{\sigma}\eta^{\tau}_{ij},$$

$$\delta^{\epsilon}_i \equiv -\kappa^{\epsilon} \left[\ln \sigma w_i \left(\frac{\sigma}{\sigma-1} \frac{w_i}{\bar{a}_i} \right)^{\sigma-1} + \tilde{\sigma}\delta^{\tau}_i + \delta^{f}_i \right] \quad \text{and} \quad \zeta^{\epsilon}_j \equiv \kappa^{\epsilon} \left[\ln \left(P_j^{\sigma-1} E_j \right) - \tilde{\sigma}\zeta^{\tau}_j \right] - \kappa^{\epsilon}\zeta^{f}_j,$$

$$\delta^{\rho}_i \equiv \ln \left(\frac{\sigma}{\sigma-1} \frac{w_i}{\bar{a}_i} \right)^{1-\sigma} - \tilde{\sigma}\delta^{\tau}_j \quad \text{and} \quad \zeta^{\rho}_j \equiv \ln \left(P_j^{\sigma-1} E_j \right) - \tilde{\sigma}\zeta^{\tau}_j.$$

We obtain expression (31) by plugging the functional form assumptions in (29) into (OA.55). Finally, notice that the definitions of ζ^{ϵ}_j and ζ^{ρ}_j above immediately imply that

$$\kappa^{\epsilon}\zeta^{f}_j = \kappa^{\epsilon}\zeta^{\rho}_j - \zeta^{\epsilon}_j.$$

B Empirical Appendix

B.1 Sample Statistics in Table OA.1

B.2 Restricted Cubic Spline Implementation

We follow Harrell Jr (2001) in setting up our restricted cubic splines. Formally we use a restricted cubic spline with knot values u_k for $k = 1, ..., K$:

$$f_1(\ln n) = \ln n$$

$$f_{k+1}(\ln n) = \frac{(\ln n - \ln u_k)^3_+ - \frac{(\ln n - \ln u_{k-1})^3_+ (\ln u_K - \ln u_{k+1}) - (\ln n - \ln u_k)^3_+ (\ln u_{k-1} - \ln u_k)}{(\ln u_k - \ln u_{k-1})}}{(\ln u_K - \ln u_1)^2},$$

with the auxiliary function $(n)_+ = n$ if $n > 0$ and zero otherwise.

Table OA.1: Data Availability

Country	$\{\bar{x}_{ij}, n_{ij}, z_{ij}\}$		$\{\tau_{ij}\}$		
Country Name	Origin	Dest.	Origin	Dest.	Developed
AUS	43	32	3	0	1
AUT	43	37	2	0	1
BEL	43	36	3	0	1
BGR	43	37	3	0	0
BRA	43	31	3	0	0
CAN	41	34	3	0	1
CHE	0	37	0	0	1
CHN	43	37	3	0	0
CYP	24	36	1	0	1
CZE	43	36	2	31	0
DEU	37	37	2	0	1
DNK	43	37	3	0	1
ESP	43	36	3	0	1
EST	43	36	3	0	0
FIN	37	36	3	0	1
FRA	43	37	3	0	1
GBR	43	36	3	0	1
GRC	43	36	3	0	1
HRV	43	29	3	0	0
HUN	37	36	3	0	0
IDN	0	31	0	0	0
IND	0	37	0	0	0
IRL	37	37	3	0	1
ITA	43	36	3	0	1
JPN	0	36	0	0	1
KOR	42	32	3	0	1
LTU	43	35	3	0	0
LUX	35	35	3	0	1
LVA	38	36	3	0	0
MEX	43	36	3	0	0
MLT	42	34	3	0	1
NLD	43	37	3	0	1
NOR	43	36	3	0	1
POL	43	36	2	0	0
PRT	43	36	3	0	1
ROU	43	36	3	0	0
RUS	0	37	0	0	0
SVK	42	36	2	36	0
SVN	42	35	3	0	1
SWE	40	37	3	0	1
TUR	43	36	3	0	0
TWN	0	29	0	0	1
USA	39	37	2	36	1
Count > 0	37	43	37	3	26
Observations	1522		103		

Table OA.2: Constant Elasticity Gravity of Firm Exports

Dep. Var.:	$\ln n_{ij}$	$\ln \bar{x}_{ij}$	$\ln X_{ij}$
	(1)	(2)	(3)
Panel A: Constant elasticity gravity estimation			
Log of to Distan	-1.148***	-0.778***	-1.927***
	(0.0478)	(0.0420)	(0.0708)
R^2	0.913	0.691	0.870
Panel B: Implied structural parameters			
	$\kappa^\epsilon \times \epsilon$	ρ	
	-0.84	-0.56	

Note. Sample of 1,524 origin-destination pairs in 2012 – see Table OA.1 of Appendix B.1. All specifications include origin and destination fixed-effects. Implied structural parameters computed with $\tilde{\sigma} = 2.9$ and $\kappa^\tau = 0.36$ from column (3) of Table 1. Standard errors clustered by origin-destination pair. *** p < 0.01

In our main specification, we choose $K = 4$ knots. To determine the values u_i, we follow Harrell Jr (2001) and divide the data into the 5th, 35th, 65th, and 95th percentiles.

B.3 The Impact of Distance on the Extensive and Intensive Margins of Firm Exports

B.3.1 Constant-Elasticity Gravity Equations

Table OA.2 presents the estimates of (34). Column (1) indicates that the exporter firm share falls sharply with distance: a 1% higher bilateral distance leads to a 1.2% decline in exporter firm share. Column (2) indicates that average sales also decline with distance. As pointed out by Fernandes et al. (2019), this evidence is inconsistent with the lack of average revenue responses in the Melitz-Pareto model (Chaney, 2008). Finally, column (3) reports an elasticity of bilateral trade flows to distance of -2, which is slightly lower than the typical estimates in the literature reviewed by Head and Mayer (2014).

In Panel B of Table OA.2, we use the expressions in (35) to recover $\kappa^\epsilon \varepsilon$ and ϱ. We use our baseline calibration of $\tilde{\sigma}\kappa^\tau = 1.04$ that sets $\tilde{\sigma} = \sigma - 1 = 2.9$ from Hottman et al. (2016) and $\kappa^\tau = 0.36$ from column (3) of Table 1. The negative extensive margin elasticity implies that $\kappa^\epsilon \varepsilon < 0$. Thus, in line with our model, $\varepsilon < 0$ whenever distance increases trade costs, $\kappa^\epsilon > 0$. In addition, the implied value of ϱ indicates that the average revenue potential of all exporters falls by 0.2% when the exporter firm share increases by 1%. Hence, marginal exporters have a lower revenue potential than incumbent exporters in each market.

B.3.2 Piece-wise Linear Approximation of Baseline Estimates

In this section, we extend the specification in (34) by allowing the coefficients β^ϵ and β^ρ to differ across country pairs in different ranges of the support of n_{ij}. This specification sheds light on how the gravity elasticities of the extensive and intensive margins of firms exports vary with the exporter firm share. Accordingly, it illustrates the main features of the data behind our estimates of the non-linear elaticities reported in Figure 3.

Table OA.3: Constant Elasticity Gravity of Firm Exports

	(1) $\ln n_{ij}$	(2) $\ln \bar{x}_{ij}$	(3) $\ln X_{ij}$
Log of Distance $\times$ $n_{ij} \in (P_0(n), P_5(n)]$	-0.992***	-0.701***	-1.693***
	(0.0360)	(0.0469)	(0.0680)
Log of Distance $\times$ $n_{ij} \in (P_5(n), P_{35}(n)]$	-0.864***	-0.724***	-1.588***
	(0.0399)	(0.0479)	(0.0729)
[1em] Log of Distance $\times$ $n_{ij} \in (P_{35}(n), P_{65}(n)]$	-0.778***	-0.702***	-1.480***
	(0.0414)	(0.0515)	(0.0774)
Log of Distance $\times$ $n_{ij} \in (P_{95}(n), P_{100}(n)]$	-0.692***	-0.667***	-1.358***
	(0.0456)	(0.0575)	(0.0848)
R^2	0.912	0.696	0.886

Note. Sample of 1,524 origin-destination pairs in 2012 – see Table OA.1 of Appendix B.1. All specifications include origin and destination fixed-effects. The four quantiles of n_{ij} are defined by the knots used in the estimation of the semiparametric specification in Figure 3, with $P_c(n)$ denoting percentile c of the empirical distribution of n_{ij}. Standard errors clustered by origin-destination pair. *** p < 0.01

Table OA.3 reports estimates of the gravity elasticity over four ranges of the support of n_{ij} defined by the same knots used to in our semi-parametric estimates in Figure 3. Column (1) shows that the extensive margin elasticity is less sensitive to distance among country pairs with high n_{ij}. In contrast, column (2) indicates that the intensive margin elasticity is roughly constant across country pairs with different levels of n_{ij}. Finally, column (3) shows that, because of the declining extensive margin elasticity, bilateral trade flows become less sensitive to distance among countries with high values of n_{ij}.

B.4 Robustness of Baseline Estimates in Section 5.4

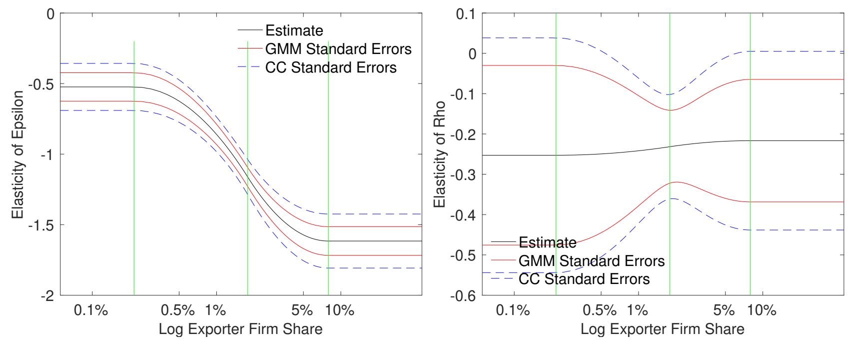
In this section, we investigate the robustness of the baseline estimates of $\epsilon(n)$ and $\rho(n)$ presented in Section 5.4. First, we implement the inference procedure in Chen and Christensen (2018) for the estimation of NPIV sieve estimators. Second, we show that results are similar when we use data for different years that have a similar country coverage. Third, we show that results are similar when we ignore observations associated with domestic sales. Fourth, we investigate how our results depend on the assumptions used to compute the number of domestic entrants, N_i. Fifth, we report our estimates with alternative calibrations of the elasticity of substitution, σ.

B.4.1 Alternative Inference

In this section, we investigate the robustness of our standard estimates of $\epsilon(n)$ and $\rho(n)$. We estimate standard errors with a criterion value estimated using the bootstrapped procedure in Chen and Christensen (2018). This method accounts for the fact that the function basis used in estimation are intended to approximate the true nonparametric functions $\epsilon(n)$ and $\rho(n)$. Our results show that the confidence intervals implied by the method in Chen and Christensen (2018) are similar to those implied by the robust standard errors of the parameter estimates.

Figure OA.1: Semiparametric Gravity of Firm Exports – Alternative Inference Procedure

(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$

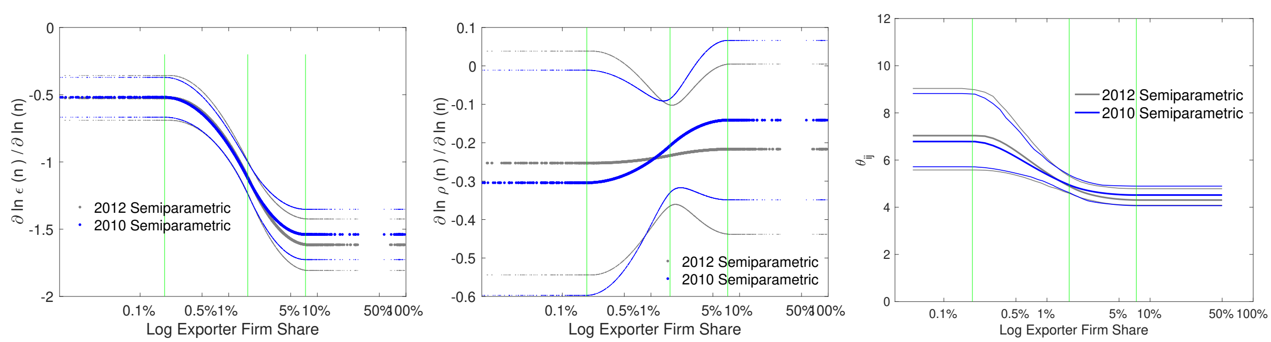


Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,522 origin-destination pairs described in Table OA.1 of Appendix B.1. Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$. CC Standard Errors obtained using the bootstrapped procedure in Chen and Christensen (2018).

B.4.2 Alternative Sample Years

Our baseline estimates use the sample of country pairs for 2012. Our data has a similar country coverage for all years between 2010 and 2014. We thus estimate the model with alternative samples for 2010 and 2014. Figures OA.2 and OA.3 show that results are broadly consistent with the baseline estimates obtained from the sample for 2012.

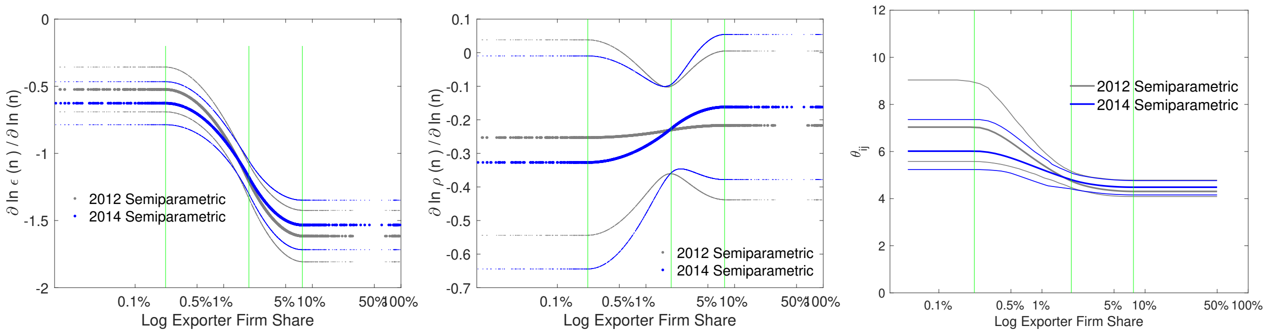
Figure OA.2: Semiparametric Gravity of Firm Exports – 2010 Sample



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 1,000 bootstrap draws for $\theta(n)$.

Figure OA.3: Semiparametric Gravity of Firm Exports – 2014 Sample



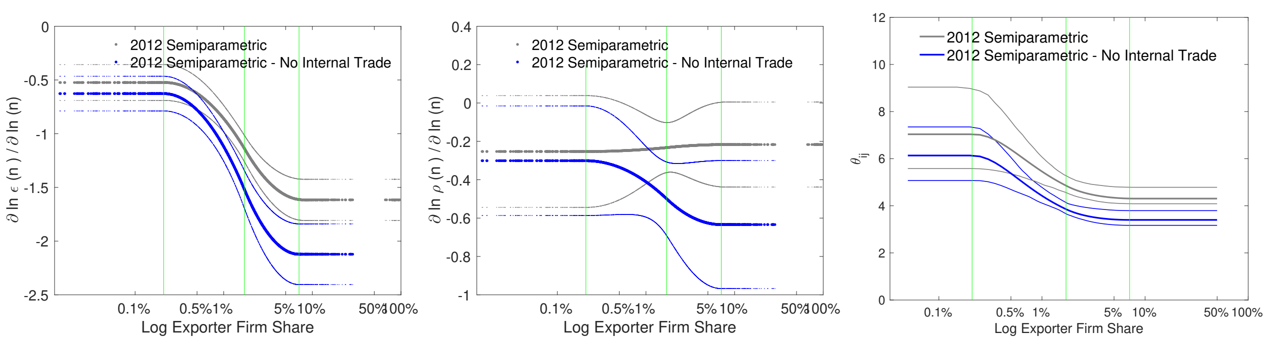
(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

B.4.3 Baseline Sample Excluding Domestic Trade Observations

Our main estimation combines data on international trade with domestic sales. This requires not only n_{ij} and z_{ij} for $i \neq j$, but also n_{ii} and z_{ii}. Importantly, domestic sales are a high fraction of observations in the top knot where the trade elasticity is lower. To access whether these estimates depend on domestic sales, we re-run our estimation procedure in an alternative sample without domestic trade observations. Figure OA.4 shows that this has only a small impact on our baseline estimates of the trade elasticity function.

Figure OA.4: Semiparametric Gravity of Firm Exports – Sample excluding domestic trade observations



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

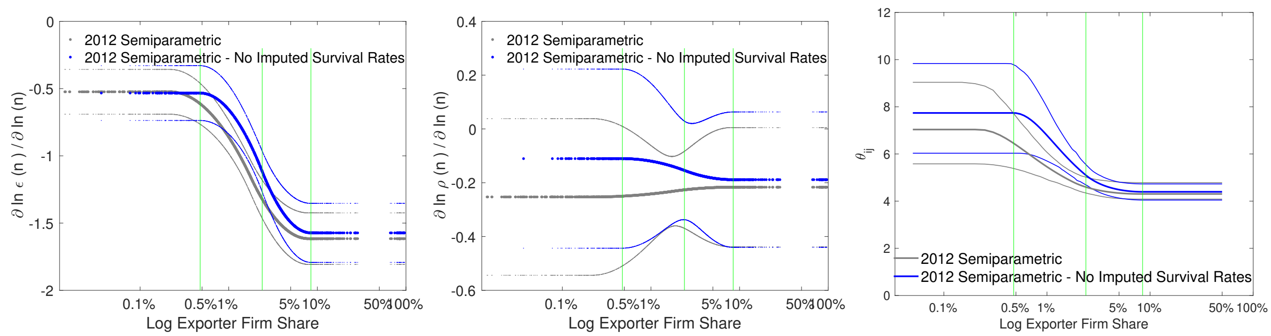
Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

B.4.4 Alternative Measures of the Number of Entrants

In our data construction, we measure the share of successful entrants using one-year survival rates for manufacturing firms. This accounts for the fact that not all entrant firms are successful in entry and may leave the market (in the spirit of Melitz (2003)). We now conduct four robustness checks with respect to the construction of n_{ii}. We first exclude all countries with imputed values of n_{ii} from our baseline sample. We also re-estimate the elasticity functions under the alternative assumptions that n_{ii} is either one (survival rate of 100%), the 2-year firm survival rate, or the 3-year firm survival rate.

Alternative sample excluding origin countries with imputed survival rate. In Figure OA.5, we replicate our baseline estimation in a sample that excludes all observations for origin countries with imputed values of the one-year survival rate. Results are roughly similar to our baseline estimates.

Figure OA.5: Semiparametric Gravity of Firm Exports – Sample excluding countries with imputed survival rates

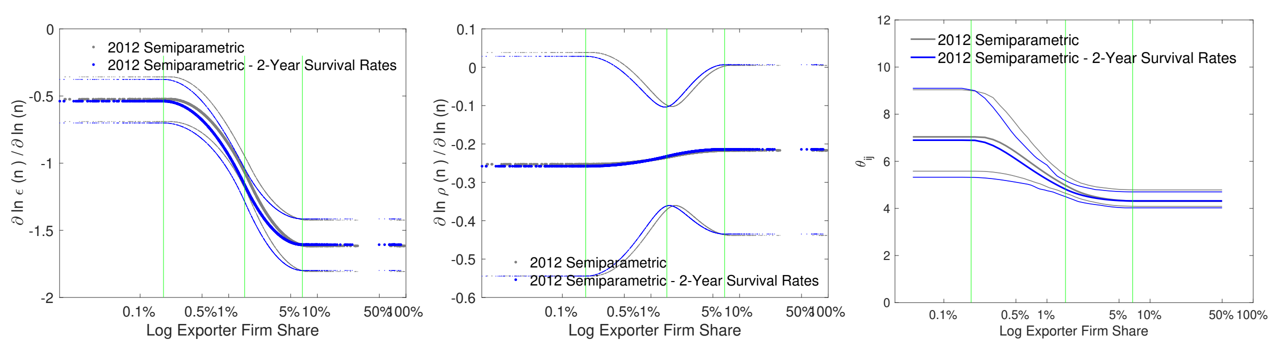


(a) Elasticity of $\epsilon\,(n)$ (b) Elasticity of $\rho\,(n)$ (c) Implied $\theta\,(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

Alternative sample with n_{ii} measured with survival rates over different horizons. We now investigate how our baseline estimates change when we measure n_{ii} using 2-year and 3-year survival rates in manufacturing. Figures OA.6 and OA.7 show that the estimated elasticity functions are almost identical in both cases.

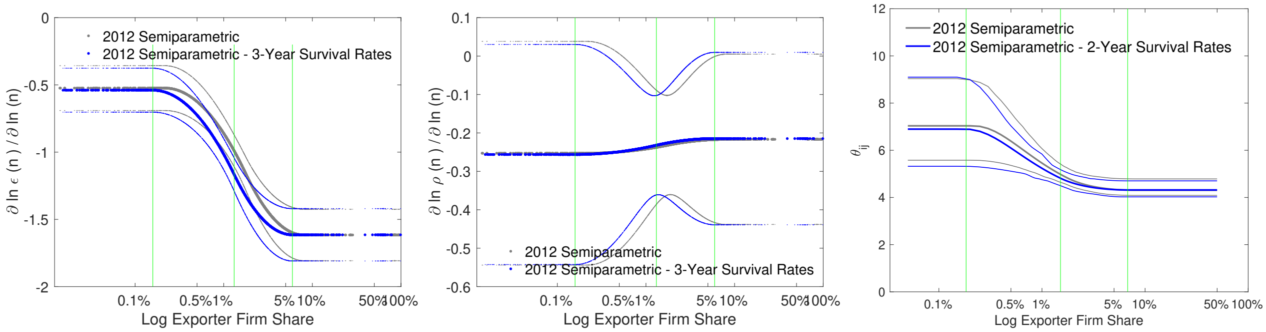
Figure OA.6: Semiparametric Gravity of Firm Exports – n_{ii} is two-year survival rate in manufacturing



(a) Elasticity of $\epsilon\,(n)$ (b) Elasticity of $\rho\,(n)$ (c) Implied $\theta\,(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

Figure OA.7: Semiparametric Gravity of Firm Exports – n_{ii} is three-year survival rate in manufacturing

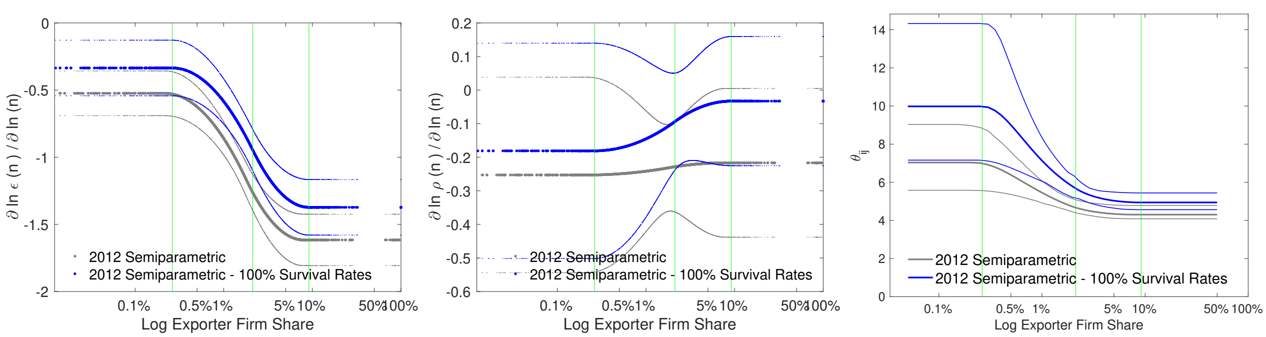


(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

Alternative sample with $n_{ii} = 1$ (survival rate of 100%). In Figure OA.8, we replicate our baseline estimates under the assumption that all entrants produce for the domestic market (i.e., $\bar{f}_{ii} = 0$ and $n_{ii} = 1$). Panel (a) shows that the estimates of $\epsilon(n)$ are robust to imposing $n_{ii} = 1$. Panel (b) indicates that point estimates for the elasticity of $\rho(n)$ are slightly increasing, but are not statistically different from our baseline estimates. The lower elasticity of $\rho(n)$ yields estimates of the trade elasticity that are close to 10 in the bottom knot.

Figure OA.8: Semiparametric Gravity of Firm Exports – $n_{ii} = 1$



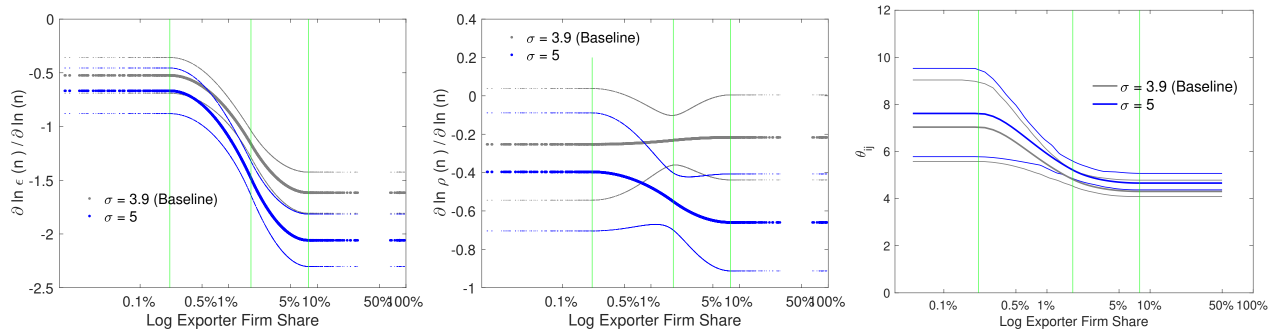
(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

B.4.5 Alternative Calibrations of the Elasticity of Substitution

In our baseline specification, we follow Hottman et al. (2016) to specify $\sigma = 3.9$. In Figure OA.9, we investigate how our our estimates depend on the value of σ. In particular, we implement our estimator for a higher elasticity of $\sigma = 5$. Our estimates for $\epsilon(n)$ are broadly similar, with a small upward level shift. Out estimates for $\rho(n)$ are slightly lower, but still close to the baseline of -0.2. Panel (c) shows that the implied $\theta(n)$ are nearly identical, with similar point estimates and confidence intervals.

Figure OA.9: Semiparametric Gravity of Firm Exports – Alternative values of σ



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

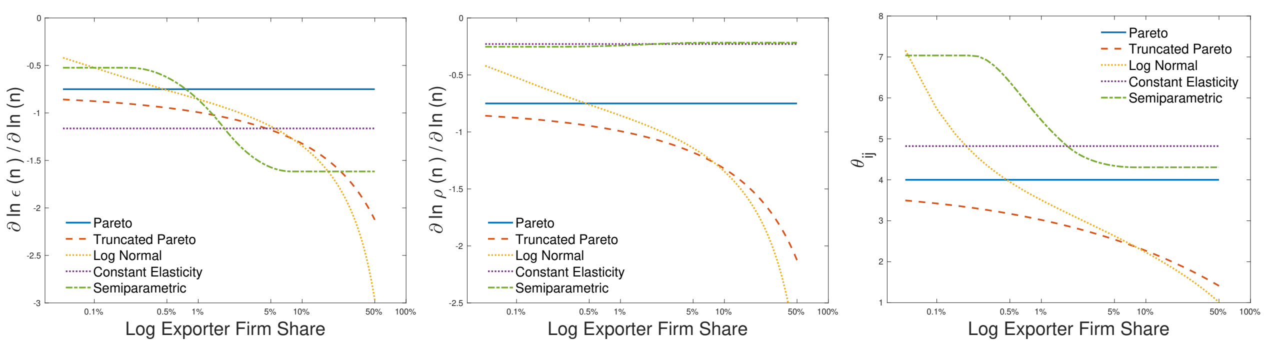
Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a single group ($G = 1$). Baseline calibration of $\tilde{\sigma} = \sigma - 1 = 2.9$ and and alternative calibration of $\tilde{\sigma} = \sigma - 1 = 4$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

B.5 Counterfactual Analysis: Additional Results

B.5.1 Gains From Trade: Comparison to the Literature

In Figure OA.10, we compare the baseline estimates in Figure 3 to calibrated elasticity functions obtained from estimates in literature of parametric distributions of firm fundamentals.

Figure OA.10: Baseline Estimates and Parametric Distributions in the Literature

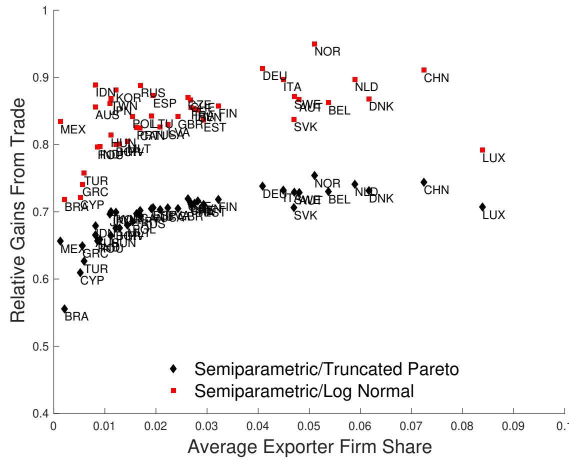


(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Pareto is the Melitz-Pareto model in Chaney (2008) with a trade elasticity of four. Truncated Pareto uses the productivity distribution in Melitz and Redding (2015). Log-normal uses the baseline estimate of the productivity distribution in Head et al. (2014). Constant elasticity and semiparametric specifications correspond to the baseline estimates in Section 5.4.

We compute the ratio between the gains from trade implied by our semiparametric gravity specification and the gains implied by specifications based on the assumption that firm productivity has either the Truncated Pareto distribution in Melitz and Redding (2015) or the Log-normal distribution in Bas et al. (2017). We use the parameter estimates reported on these papers. Figure OA.11 presents the cross-country relationship between these ratios and initial trade outcomes.

Figure OA.11: Importance of Functional Form Assumptions for the Gains from Trade



Note. Gains from Trade is the percentage change in the real wage implied by moving from autarky to the observed equilibrium in 2012. For each specification, gains from trade are computed with the formula in Section 3.3 for $\hat{n}_{ii}$ and $\hat{N}_i$ solving the system in Appendix A.6. Gains for semiparametric specification computed with the semiparametric estimates in Figure 3. Gains for Truncated Pareto specification computed with elasticity functions implied by the productivity distribution in Melitz and Redding (2015). Gains for Log-normal specification computed with elasticity functions implied by the productivity distribution in Head et al. (2014). Vertical axis is the ratio between the gains from trade implied by our baseline semiparametric estimates and those obtained with each parametric assumption.

The diamond-shaped dots in Figure OA.11 show that the Truncated Pareto specification leads to much higher gains from trade for all countries when compared to those implied by our baseline estimates. This is a direct consequence of the low trade elasticities implied by the parametrization in Melitz and Redding (2015) – see Figure OA.10. The square-shaped dots in Figure OA.11 show that the gains from trade are also lower for the Log-normal specification. Again, this follows from the average trade elasticity implied by the productivity distribution in Bas et al. (2017). Figure OA.10 shows that the implied trade elasticity in the log-normal case is slightly lower than our baseline estimate for all values of the exporter firm share.

B.5.2 The Welfare Impact of Observed Changes in Trade Costs: Additional Results

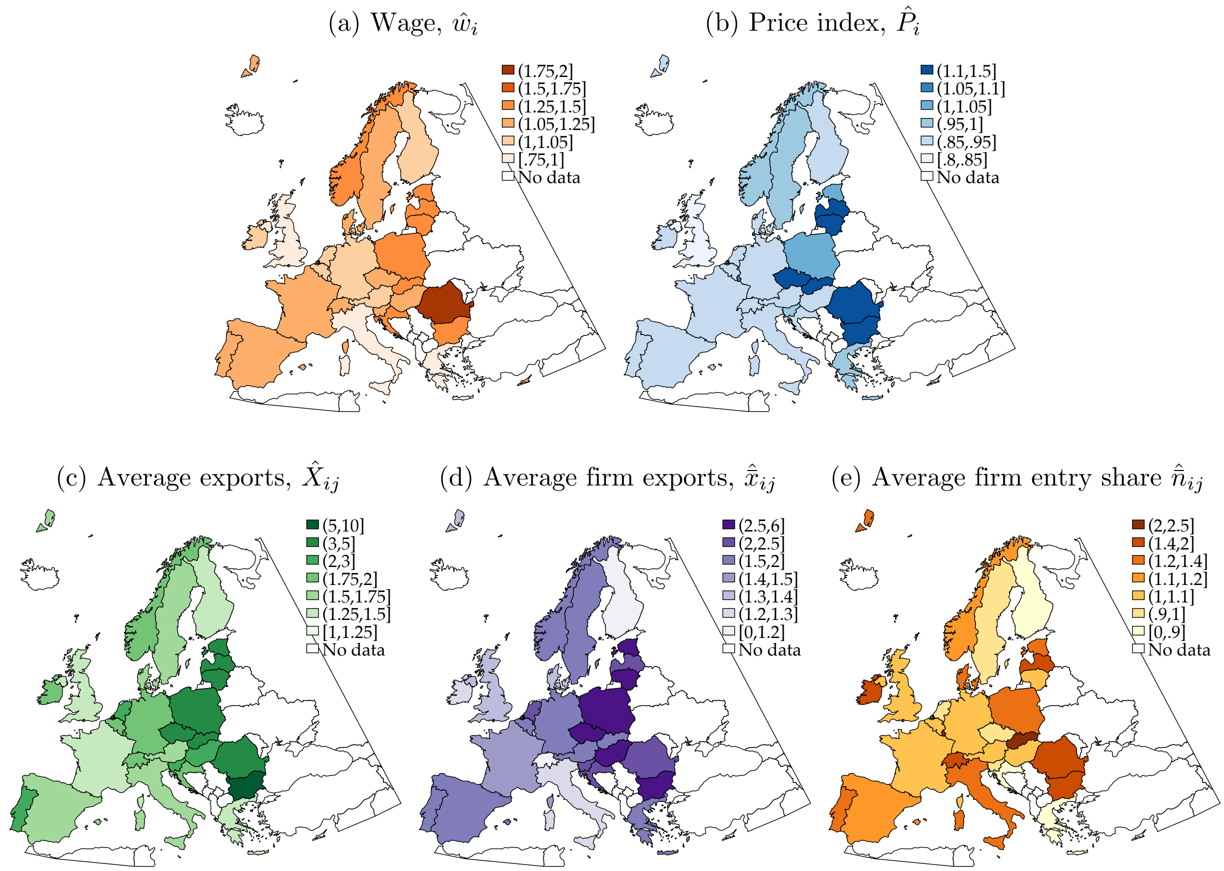
This section complements the quantitative results in Section 6.2 regarding the welfare impact of changes in trade costs among members of the European single market.

Summary statistics: changes in observed variables. Figure OA.12 depicts the change in the outcomes used in the inversion of changes in economic fundamentals for each origin country. Panel (a) shows that countries experienced vary different changes in their wage relative to the U.S. wage. We observe strong wage gains in Eastern Europe, which are 50% higher than U.S. wage gains for several countries. For western countries, wage gains are often small, with relative wages falling by 0.8% in Italy, 2.0% in the UK, and 10.2% in Greece. These wage patterns contribute to the weak growth in relative revenue shifters that we observe in West Europe. Panel (b) reports the price index changes used to measure the domestic revenue shifters, $\hat{\bar{r}}^t_{jj}$. Relative to U.S. wage growth, the price index fell slightly in West Europe, but increased in East Europe.

Panels (c)–(e) illustrate the changes in export outcomes for each origin country. For most countries, we observe strong overall growth in total exports (panel (c)), firm average exports (panel (d)), and firm export shares (panel (e)). Again, export growth was substantially stronger in East Europe than in West Europe.

The different scales in panels (d) and (e) indicate that the growth in firm average exports was stronger than the growth in firm export shares. In our model, this fact is an important force contributing to the observed rises of both inverted fixed export costs and inverted revenue shifters – see the expressions in (36) and the companion discussion.

Figure OA.12: Change in Outcomes by Origin Country, 2003-2012



Note: Each panel reports the origin's outcome in 2012 divided by the origin's outcome in 2003. For bilateral variables in panels (c)-(e), we report the simple average of the ratio across all other destination countries j in the European single market (with $i \neq j$).

Comparison to the literature: alternative measures of trade cost shocks. We now compare the welfare impact of shocks in bilateral trade costs measured with different methodologies. Figure OA.15 reports the changes in real wages implied by three different approaches to measure trade cost shocks. In panel (a), we use the baseline changes in bilateral trade costs, $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = \hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$, implied by Proposition 3. In panel (b), we measure the changes in trade costs using observed tariff changes between countries – in our application, we consider tariff changes caused by the EU enlargement, as in Caliendo et al. (2017). Finally, in panel (c), we measure trade cost shocks using a non-parametric extension of the approach in Head and Ries (2001) that imposes symmetry of trade cost shocks. Specifically, by assuming that $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = (\hat{\hat{\tau}}_{ji})^{1-\sigma}$, we set trade cost shocks to

$$(\hat{\hat{\tau}}_{ij})^{1-\sigma} = (\hat{\hat{\tau}}_{ji})^{1-\sigma} = \left(\frac{\hat{\hat{x}}_{ij}}{\hat{\hat{x}}_{ii}} \frac{\hat{\hat{x}}_{ji}}{\hat{\hat{x}}_{jj}} \left[\frac{\rho_{ij}(n_{ij})}{\rho_{ii}(n_{ii})} \frac{\rho_{ji}(n_{ji})}{\rho_{jj}(n_{jj})} \right] \left[\frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ii}(n_{ii}\hat{n}_{ii})} \frac{\rho_{ji}(n_{ji}\hat{n}_{ji})}{\rho_{jj}(n_{jj}\hat{n}_{jj})} \right]^{-1} \right)^{1/2}, \quad \text{(OA.56)}$$

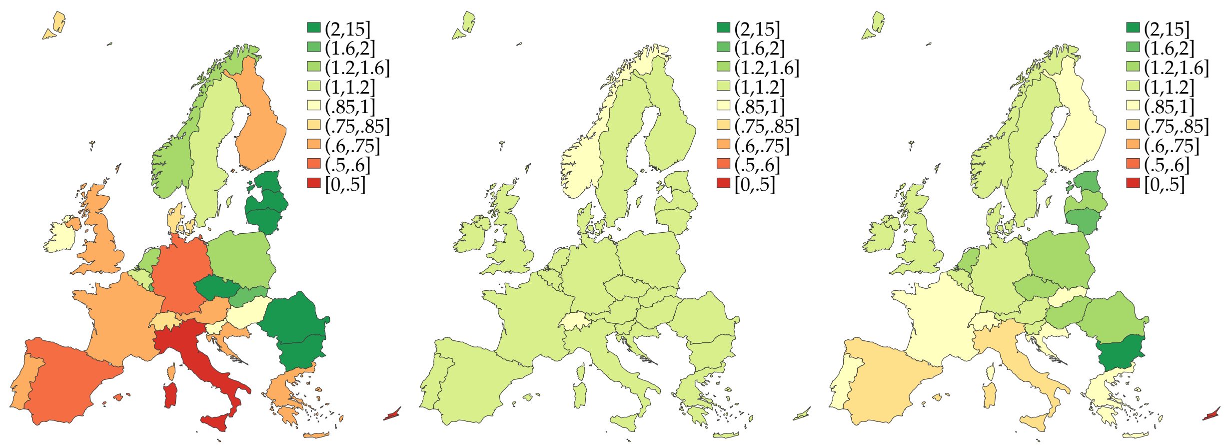
as implied by the semiparametric equation of firm average exports in (14).

Figures OA.13 and OA.14 compare the shocks measured with the different approaches. Panel (a) shows great cross-country dispersion in our baseline inverted trade shocks – the standard deviation across origins of the average log-change was 1.04. In Panel (b), tariff changes created much smaller changes in bilateral revenue shifters that were on average positive for all origin countries – the average log-change was 0.075, with a standard deviation of 0.11. Figure OA.14 indicates that tariff changes and our inverted shocks have a correlation close to zero. Finally, panel (c) shows the average revenue shifters implied by the symmetric shocks measured with (OA.56). Because it is "approximately" a bilateral average of our inverted trade shocks, expression (OA.56) yields smaller increases (declines) in bilateral revenue shifters for East (West) Europe. Figure OA.14 shows that, despite this attenuation bias, the two measures are positively correlated – their correlation is 0.56.

Figure OA.15 reports the welfare changes implied by the three alternative ways of measuring trade cost shocks. The comparison between panels (a) and (b) indicates that the welfare impact of the tariff changes is much smaller than that of our inverted trade cost shocks. The average effect is 0.7% in panel (a), but only 0.3% in panel (b). This follows from the fact that tariff changes are much smaller and more homogeneous than the inverted changes in trade costs. Lastly, when we use the symmetric trade cost shocks in panel (c), we get smaller welfare losses in West Europe and smaller welfare gains in East Europe. This attenuation of trade cost shocks is a natural implication by the symmetry assumptions (as discussed above).

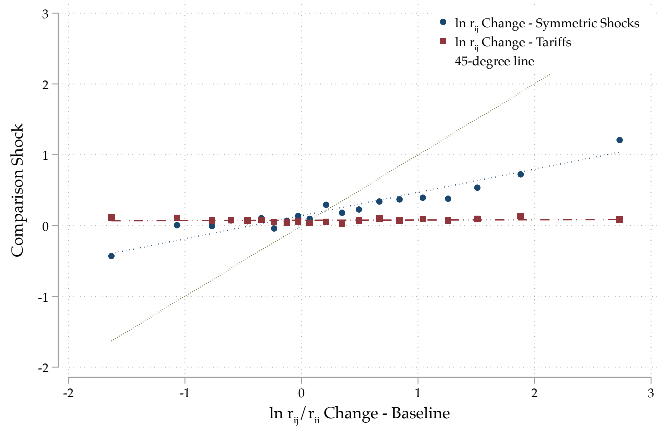
Figure OA.13: Average Change in Bilateral Revenue Shifters in The European Single Market, 2003-2012 – Alternative Approaches to Measure Shocks

(a) Baseline, $\hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$ (b) Tariff change, $(\widehat{1+\text{tariff}_{ij}})^{1-\sigma}$ (c) Head-Ries, $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = (\hat{\hat{\tau}}_{ji})^{1-\sigma}$



Note: Changes in variable trade costs recovered from changes in outcomes between 2003 and 2012. For each origin country i, we report the simple average of the change in bilateral revenue shifter across all other destination countries j in the European single market (with $i \neq j$). In Panel (a), we set $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = \hat{\hat{r}}_{ij}/\hat{\hat{r}}_{ii}$ for $i \neq j$ with $\hat{\hat{r}}_{ij}$ computed using Proposition 3. In Panel (b), we set $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = (\widehat{1+\text{tariff}_{ij}})^{1-\sigma}$ to be the change in tariffs between EU members, as in Caliendo et al. (2017). In Panel (c), for $i \neq j$, we use $(\hat{\hat{\tau}}_{ij})^{1-\sigma} = (\hat{\hat{\tau}}_{ji})^{1-\sigma}$ given by the generalization of the approach in Head and Ries (2001) in expression (OA.56).

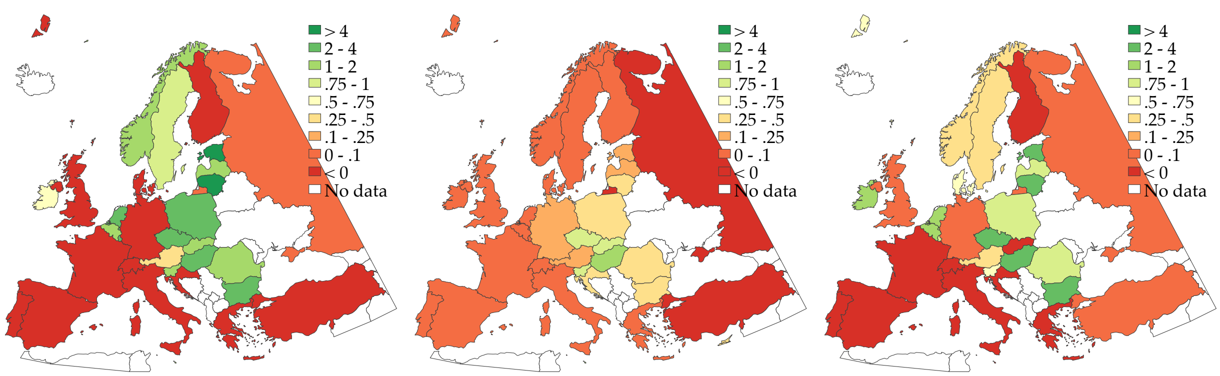
Figure OA.14: Bin Scatter Plot - Comparison of Changes in Bilateral Revenue Shifters Implied by Different Approaches



Note: Changes in variable trade costs recovered from changes in outcomes between 2003 and 2012. The horizontal axis is the baseline shock $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = \hat{\bar{r}}_{ij}/\hat{\bar{r}}_{ii}$ with $\hat{\bar{r}}_{ij}$ computed using Proposition 3. The vertical axis is the change in bilateral revenue shifter implied by an alternative approach. For tariff change, we set $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = (\widehat{1+\text{tariff}}_{ij})^{1-\sigma}$ to be the change in tariffs between EU members, as in Caliendo et al. (2017). For symmetric shock, we use $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = (\hat{\bar{\tau}}_{ji})^{1-\sigma}$ given by the generalization of the approach in Head and Ries (2001) in expression (OA.56).

Figure OA.15: The Welfare Impact of Changing Bilateral Variable Trade Costs in The European Single Market, 2003-2012 – Alternative Approaches to Measure Shocks

(a) Baseline, $\hat{\bar{r}}_{ij}/\hat{\bar{r}}_{ii}$ (b) Tariff change, $(\widehat{1+\text{tariff}}_{ij})^{1-\sigma}$ (c) Head-Ries, $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = (\hat{\bar{\tau}}_{ji})^{1-\sigma}$



Note: Real wage changes caused by changes in bilateral trade shifters between 2003 and 2012 among 30 countries in the European single market. In Panel (a), we set $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = \hat{\bar{r}}_{ij}/\hat{\bar{r}}_{ii}$ for $i \neq j$ with $\hat{\bar{r}}_{ij}$ computed using Proposition 3. In Panel (b), we set $\hat{\bar{\tau}}_{ij}$ to be the change in tariffs between EU members, as in Caliendo et al. (2017). In Panel (c), for $i \neq j$, we use $(\hat{\bar{\tau}}_{ij})^{1-\sigma} = (\hat{\bar{\tau}}_{ji})^{1-\sigma}$ given by the generalization of the approach in Head and Ries (2001) in expression (OA.56). For each shock, we report minus the real wage change obtained with the baseline semiparametric estimates in Figure 3.

Supplementary Material for

Aggregate Implications of Firm Heterogeneity: A Nonparametric Analysis of Monopolistic Competition Trade Models

Not Intended for Publication

by Rodrigo Adao, Costas Arkolakis, and Sharat Ganapati

A Supplementary Theory Appendix: Extensions

This appendix presents five extensions of our baseline framework. In Section A.1, we extend our model to include heterogeneous firms in multiple sectors whose production function uses multiple factors and sector-specific inputs. In Section A.2, we relax the assumption of full support in the distribution of entry potentials to allow for zero trade flows between countries. In Section A.3, we incorporate import tariffs into trade costs and government revenue. Section A.4 extends our baseline framework to allow firms to produce multiple products. Finally, Section A.5 relaxes the CES assumption in our framework by allowing for a general class of demand functions with a single aggregator.

A.1 Multi-Sector, Multi-Factor Heterogeneous Firm Model with Input-Output Links

In this section, we extend our baseline framework to allow for firm heterogeneity in a model with multiple sectors, multiple factors of production, and input-output linkages. Our specification of the model can be seen as a generalziation of the formulation in Costinot and Rodriguez-Clare (2013).

A.1.1 Environment

The world economy is constituted of countries with multiple sectors indexed by s. Each country has a representative household that inelastically supplies $\bar{L}_{i,f}$ units of multiple factors of production indexed by f.

Preferences. The representative household in country j has CES preferences over the composite good of multiple sectors, $s = 1, ...S$:

$$U_j = \left[\sum_s \gamma_j^s \left(Q_j^k\right)^{\frac{\lambda_j - 1}{\lambda_j}} \right]^{\frac{\lambda_j - 1}{\lambda_j}}.$$

Given the price of the sectoral composite goods, the share of spending on sector s is

$$c_j^s = \gamma_j^s \left(\frac{P_j^s}{P_j}\right)^{1-\lambda_j} \tag{A.1}$$

where the consumption price index is

$$P_j = \left[\sum_k \gamma_j^s \left(P_j^k\right)^{1-\lambda_j} \right]^{\frac{1}{1-\lambda_j}}. \tag{A.2}$$

Sectoral final composite good. In each sector s of country j, there is a perfectly competitive market for a non-tradable final good whose production uses different varieties of the tradable input good in sector s:

$$Q_j^s = \left(\sum_i \int_{\Omega_{ij}^s} \left(\bar{b}_{ij}^s b_{ij}^s(\omega)\right)^{\frac{1}{\sigma^s}} \left(q_{ij}^s(\omega)\right)^{\frac{\sigma^s - 1}{\sigma^s}} \, d\omega \right)^{\frac{\sigma^s}{\sigma^s - 1}}$$

where $\sigma_j^s > 1$ and Ω_{ij}^s is the set of sector s's varieties of intermediate goods produced in country i available in country j.

The demand of country j by variety ω of sector s in country i is

$$q_{ij}^{s}(\omega) = \left(\bar{b}_{ij}^{s} b_{ij}^{s}(\omega)\right) \left(\frac{p_{ij}^{s}(\omega)}{P_{j}^{s}}\right)^{-\sigma^{s}} \frac{E_{j}^{s}}{P_{j}^{s}}$$

where E_j^s is the total spending of country j in sector s.

Because the market for the composite sectoral good is competitive, the price is the CES price index of intermediate inputs:

$$\left(P_{j}^{s}\right)^{1-\sigma^{s}} = \sum_{i} \int_{\Omega_{ij}^{s}} \left(\bar{b}_{ij}^{s} b_{ij}^{s}(\omega)\right) \left(p_{ij}^{s}(\omega)\right)^{1-\sigma^{s}} \, d\omega. \tag{A.3}$$

Sectoral intermediate good. In sector s of country i, there is a representative competitive firm that produces a non-traded sectoral intermediate good using different factors and the non-traded composite final good of different sectors. The production function is

$$q_{i}^{s} = \left[\alpha_{i}^{s} \left(L_{i}^{s}\right)^{\frac{\mu_{i}^{s}-1}{\mu_{i}^{s}}} + \left(1-\alpha_{i}^{s}\right) \left(M_{i}^{s}\right)^{\frac{\mu_{i}^{s}-1}{\mu_{i}^{s}}}\right]^{\frac{\mu_{i}^{s}}{\mu_{i}^{s}-1}},$$

where

$$L_{i}^{s} = \left[\sum_{f} \beta_{i}^{s,f} \left(L_{i}^{s,f}\right)^{\frac{\eta_{i}^{s}-1}{\eta_{i}^{s}}}\right]^{\frac{\eta_{i}^{s}}{\eta_{i}^{s}-1}} \quad \text{and} \quad M_{i}^{s} = \left[\sum_{k} \theta_{i}^{ks} \left(Q_{i}^{k}\right)^{\frac{\kappa_{i}^{s}-1}{\kappa_{i}^{s}}}\right]^{\frac{\kappa_{i}^{s}}{\kappa_{i}^{s}-1}}.$$

Zero profit implies that the price of the sectoral intermediate good is

$$\bar{p}_{i}^{s} = \left[\alpha_{i}^{s} \left(W_{i}^{s}\right)^{1-\mu_{i}^{s}} + \left(1-\alpha_{i}^{s}\right) \left(V_{i}^{s}\right)^{1-\mu_{i}^{s}}\right]^{\frac{1}{1-\mu_{i}^{s}}}, \tag{A.4}$$

where

$$W_{i}^{s} = \left[\sum_{f} \beta_{i}^{s,f} \left(w_{i}^{f}\right)^{1-\eta_{i}^{s}}\right]^{\frac{1}{1-\eta_{i}^{s}}} \quad \text{and} \quad V_{i}^{s} = \left[\sum_{k} \theta_{i}^{ks} \left(P_{i}^{k}\right)^{1-\kappa_{i}^{s}}\right]^{\frac{1}{1-\kappa_{i}^{s}}}. \tag{A.5}$$

The share of total production cost in sector s spent on factor f and input k are given by

$$l_{i}^{s,f} = \beta_{i}^{s,f} \left(\frac{w_{i}^{f}}{W_{i}^{s}}\right)^{1-\eta_{i}^{s}} \alpha_{i}^{s} \left(\frac{W_{i}^{s}}{\bar{p}_{i}^{s}}\right)^{1-\mu_{i}^{s}} \quad \text{and} \quad m_{i}^{ks} = \theta_{i}^{ks} \left(\frac{P_{i}^{k}}{V_{i}^{s}}\right)^{1-\kappa_{i}^{s}} \left(1-\alpha_{i}^{s}\right) \left(\frac{V_{i}^{s}}{\bar{p}_{i}^{s}}\right)^{1-\mu_{i}^{s}}. \tag{A.6}$$

Production of traded intermediate varieties ω. Assume that sector s has a continuum of monopolistic firms that produce output using only a non-tradable input q_i^s. In order to sell q in market j, variety ω of country i faces a cost function given by

$$C_{ij}(\omega, q) = \bar{p}_{i}^{s} \frac{\tau_{ij}^{s}(\omega)}{a_{i}^{s}(\omega)} \frac{\bar{\tau}_{ij}^{s}}{\bar{a}_{i}^{s}} q + \bar{p}_{i}^{s} \bar{f}_{ij}^{s} f_{ij}^{s}(\omega)$$

where $\bar{p}_i^s$ is the price of the non-tradable input q_i^s in country i.

Given this production technology, the optimal price is $p_{ij}^{s}(\omega) = \frac{\sigma_{j}^{s}}{\sigma_{j}^{s}-1} \frac{\tau_{ij}^{s}(\omega)}{a_{i}^{s}(\omega)} \frac{\bar{\tau}_{ij}^{s}}{\bar{a}_{i}^{s}} \bar{p}_{i}^{s}$ and the associated revenue is

$$R_{ij}^{s}(\omega) = r_{ij}^{s}(\omega) \bar{r}_{ij}^{s} \left[\left(\frac{\bar{p}_{i}^{s}}{P_{j}^{s}}\right)^{1-\sigma^{s}} E_{j}^{s}\right] \tag{A.7}$$

where

$$r_{ij}^{s}(\omega) \equiv b_{ij}^{s}(\omega)\left(\frac{\tau_{ij}^{s}(\omega)}{a_{i}^{s}(\omega)}\right)^{1-\sigma^{s}} \quad \text{and} \quad \bar{r}_{ij} \equiv \bar{b}_{ij}^{s}\left(\frac{\sigma^{s}}{\sigma^{s}-1}\frac{\bar{\tau}_{ij}^{s}}{\bar{a}_{i}^{s}}\right)^{1-\sigma^{s}}. \tag{A.8}$$

Firm ω of country i chooses to enter a foreign market j if, and only if, $\pi_{ij}^{s}(\omega) = (1/\sigma_{j}^{s})R_{ij}^{s}(\omega) - \bar{p}_{i}^{s}\bar{f}_{ij}^{s}f_{ij}^{s}(\omega) \geq 0$. This condition determines the set of firms from country i that operate in sector s of country j:

$$\omega \in \Omega_{ij}^{s} \quad \Leftrightarrow e_{ij}^{s}(\omega) \geq \sigma^{s}\frac{\bar{f}_{ij}^{s}}{\bar{r}_{ij}^{s}}\left[\left(\frac{\bar{p}_{i}^{s}}{P_{j}^{s}}\right)^{\sigma^{s}}\frac{P_{j}^{s}}{E_{j}^{s}}\right], \tag{A.9}$$

where

$$e_{ij}^{s}(\omega) \equiv \frac{r_{ij}^{s}(\omega)}{f_{ij}^{s}(\omega)}. \tag{A.10}$$

Entry of traded intermediate varieties ω. Firms in sector s of country i can create a new variety by spending $\bar{F}_{i}^{s}$ units of the non-tradable sectoral input q_{i}^{s}. In this case, they take a draw of the variety characteristics from an arbitrary distribution:

$$v_{i}(\omega) \equiv \left\{a_{i}^{s}(\omega), b_{ij}^{s}(\omega), \tau_{ij}^{s}(\omega), f_{ij}^{s}(\omega)\right\}_{j} \sim G_{i}^{s}(v). \tag{A.11}$$

In equilibrium, free entry implies that N_{i}^{s} firms pay the fixed cost of entry in exchange for an ex-ante expected profit of zero,

$$\sum_{j} E\left[\max\left\{\pi_{ij}^{s}(\omega);\ 0\right\}\right] = \bar{p}_{i}^{s}\bar{F}_{i}^{s}. \tag{A.12}$$

Market clearing. We follow Dekle et al. (2008) by allowing for a set of exogenous transfers. Thus, the spending on goods of sector s by country i is

$$E_{i}^{s} = c_{j}^{s}\left(w_{i}L_{i} + T_{i}\right) + \sum_{k} m_{i}^{sk}\left(\bar{p}_{i}^{k}q_{i}^{k}\right). \tag{A.13}$$

The market clearing conditions for factor f in country i is

$$w_{i}^{f}\bar{L}_{i}^{f} = \sum_{s} l_{i}^{s,f}\left(\bar{p}_{i}^{s}q_{i}^{s}\right). \tag{A.14}$$

Since all the revenue of the sectoral intermediate good comes from sales to the firms producing the varieties ω, we have that

$$\frac{\bar{p}_{i}^{s}q_{i}^{s}}{N_{i}^{s}} = \underbrace{\sum_{j}\left(1-\frac{1}{\sigma^{s}}\right)Pr[\omega \in \Omega_{ij}^{s}]E[R_{ij}^{s}(\omega)|\omega \in \Omega_{ij}^{s}]}_{\text{final good production}} + \underbrace{\sum_{j}\bar{p}_{i}^{s}\bar{f}_{ij}^{s}Pr[\omega \in \Omega_{ij}^{s}]E[f_{ij}^{s}(\omega)|\omega \in \Omega_{ij}^{s}]}_{\text{fixed-cost of entering markets}} + \underbrace{\bar{p}_{i}^{s}\bar{F}_{i}^{s}}_{\text{fixed cost of entry}}.$$

The free entry condition in (A.12) implies that

$$\bar{p}_{i}^{s}\bar{F}_{i}^{s} = \sum_{j} E\left[\max\left\{\pi_{ij}^{s}(\omega);\ 0\right\}\right] = \sum_{j} Pr[\omega \in \Omega_{ij}^{s}]\left(\frac{1}{\sigma^{s}}E[R_{ij}^{s}(\omega)|\omega \in \Omega_{ij}^{s}] - \bar{p}_{i}^{s}\bar{f}_{ij}^{s}E[f_{ij}^{s}(\omega)|\omega \in \Omega_{ij}^{s}]\right).$$

Thus, $\bar{p}_i^s q_i^s = \sum_j N_i^s Pr[\omega \in \Omega_{ij}^s] E[R_{ij}^s(\omega) | \omega \in \Omega_{ij}^s]$ and, therefore,

$$\bar{p}_i^s q_i^s = \sum_j \bar{r}_{ij}^s \left[\left(\frac{\bar{p}_i^s}{P_j^s} \right)^{1-\sigma^s} E_j^s \right] \left[\int_{\omega \in \Omega_{ij}^s} r_{ij}^s(\omega)\, d\omega \right]. \tag{A.15}$$

Equilibrium. Given the distribution in (A.11), the equilibrium is P_i, $\{\Omega_{ij}^s\}_{j,s}$, $\{P_i^s, N_i^s, \bar{p}_i^s, W_i^s, V_i^s, \bar{p}_i^s q_i^s, E_i^s, c_i^s\}_s$, $\{m_i^{sk}\}_{k,s}$, $\{l_i^{s,f}\}_{f,s}$ and $\{w_i^f\}_f$ for all i that satisfy equations (A.2), (A.9). (A.3), , (A.12), (A.4), (A.5), (A.6), (A.13), (A.1), (A.15), (A.14).

A.1.2 Extensive and Intensive margin of Firm-level Export

We now turn to the characterization of the bilateral levels of entry and sales in each sector. As before, we consider the marginal distribution of $\left(r_{ij}^s(\omega), e_{ij}^s(\omega)\right)$ implied by G_i^s, which can be decomposed without loss of generality as

$$r_{ij}^s(\omega) \sim H_{ij}^{r,s}(r|e), \qquad \text{and} \qquad e_{ij}^s(\omega) \sim H_{ij}^{e,s}(e), \tag{A.16}$$

where $H_{ij}^{e,s}$ has full support in $\mathbb{R}_+$.

Extensive margin of firm-level exports. The share of firms in sector s of country i serving market j is $n_{ij}^s = Pr\left[\omega \in \Omega_{ij}^s\right]$. We define $\epsilon_{ij}^s(n) \equiv \left(H_{ij}^{e,s}\right)^{-1}(1-n)$ such that

$$\ln \epsilon_{ij}^s(n_{ij}^s) = \ln\left(\sigma^s \bar{f}_{ij}^s / \bar{r}_{ij}^s\right) + \ln\left(\bar{p}_i^s\right)^{\sigma^s} - \ln E_j^s \left(P_j^s\right)^{\sigma^s - 1}. \tag{A.17}$$

Thus, we obtain a sector-specific version of the relationship between the function of the share of firms from i selling in j and the linear combination of exogenous bilateral trade shifters and endogenous outcomes in the origin and destination markets.

Intensive margin of firm-level exports. The average revenue of firms from country i in country j is $\bar{x}_{ij}^s \equiv E\left[R_{ij}^s(\omega) | \omega \in \Omega_{ij}^s\right]$. Define the average revenue potential of exporters when n_{ij}^s% of $i's$ firms in sector s export to j as $\rho_{ij}^s\left(n_{ij}^s\right) \equiv \frac{1}{n_{ij}^s} \int_0^{n_{ij}^s} \tilde{\rho}_{ij}^s(n)\, dn$ where $\tilde{\rho}_{ij}^s(n) \equiv E[r | e = \epsilon_{ij}^s(n)]$ is the average revenue potential in quantile n of the entry potential distribution. Using the transformation $n = 1 - H_{ij}^{e,s}(e)$ such that $e = \epsilon_{ij}^s(n)$ and $dH_{ij}^{e,s}(e) = -dn$, we can follow the same steps as in the baseline model to show that

$$\ln \bar{x}_{ij}^s - \ln \rho_{ij}^s(n_{ij}^s) = \ln\left(\bar{r}_{ij}^s\right) + \ln\left(\bar{p}_i^s\right)^{1-\sigma^s} + \ln E_j^s \left(P_j^s\right)^{\sigma^s - 1}. \tag{A.18}$$

Thus, we obtain a sector-specific version of the relationship between the composition-adjusted per-firm sales and a linear combination of exogenous bilateral trade shifters and endogenous outcomes in the origin and destination markets.

A.1.3 General Equilibrium

We now write the equilibrium conditions in terms of $\rho_{ij}^s(n)$ and $\epsilon_{ij}^s(n)$. We start by writing the price index P_j^s in (A.3) in terms of $\rho_{ij}^s(n)$. Using the expression for $p_{ij}^s(\omega)$ and (A.3), we have that $(P_j^s)^{1-\sigma^s} = \sum_i \bar{r}_{ij}^s \left(\bar{p}_i^s\right)^{1-\sigma^s} \int_{\Omega_{ij}^s} r_{ij}^s(\omega)\; d\omega$. Since $\int_{\Omega_{ij}^s} r_{ij}^s(\omega)\; d\omega = N_i^s Pr[\omega \in \Omega_{ij}^s] E[r | \omega \in \Omega_{ij}^s] = N_i^s n_{ij}^s \rho_{ij}^s(n_{ij}^s)$, we can write P_j^s as

$$\left(P_j^s\right)^{1-\sigma^s}=\sum_i \bar{r}_{ij}^s\left(\bar{p}_i^s\right)^{1-\sigma^s}\rho_{ij}^s(n_{ij}^s)n_{ij}^s N_i^s. \tag{A.19}$$

We then turn to the free entry condition in (A.12). Following the same steps as in Appendix A.1, it is straight forward to show that

$$\mathbb{E}\left[\max\left\{\pi_{ij}^s(\omega);\ 0\right\}\right]=\frac{1}{\sigma^s}n_{ij}^s\bar{x}_{ij}^s-\bar{p}_i^s\bar{f}_{ij}^s\int_0^{n_{ij}}\frac{\tilde{\rho}_{ij}^s(n)}{\epsilon_{ij}^s(n)}\,dn,$$

which implies that the free entry condition in (A.12) is equivalent to

$$\sigma^s\bar{p}_i^s F_i^s=\sum_j n_{ij}^s\bar{x}_{ij}^s-\sum_j\left(\sigma^s\bar{p}_i^s\bar{f}_{ij}^s\right)\int_0^{n_{ij}}\frac{\tilde{\rho}_{ij}^s(n)}{\epsilon_{ij}^s(n)}\,dn.$$

Notice that the summation of (A.17) and (A.18) implies that $\ln\left(\sigma^s\bar{p}_i^s\bar{f}_{ij}^s\right)=\ln\bar{x}_{ij}^s-\ln\bar{\rho}_{ij}^s(n_{ij}^s)+\ln\bar{\epsilon}_{ij}^s(n_{ij}^s)$. Thus, following again the same steps as in Appendix A.1, it is straight forward to show that

$$\sigma^s\bar{p}_i^s\bar{F}_i^s=\sum_j n_{ij}^s\bar{x}_{ij}^s\left(1-\frac{\int_0^{n_{ij}^s}\frac{\tilde{\rho}_{ij}^s(n)}{\epsilon_{ij}^s(n)}\,dn}{\int_0^{n_{ij}^s}\frac{\tilde{\rho}_{ij}^s(n)}{\epsilon_{ij}^s(n_{ij}^s)}\,dn}\right). \tag{A.20}$$

Finally, we established above that $\bar{p}_i^s q_i^s=\sum_j N_i^s Pr[\omega\in\Omega_{ij}^s]E[R_{ij}^s(\omega)|\omega\in\Omega_{ij}^s]$. Since $n_{ij}^s\equiv Pr[\omega\in\Omega_{ij}^s]$ and $\bar{x}_{ij}^s\equiv E[R_{ij}^s(\omega)|\omega\in\Omega_{ij}^s]$, then

$$\bar{p}_i^s q_i^s=\sum_j N_i^s n_{ij}^s\bar{x}_{ij}^s. \tag{A.21}$$

This implies that, given $\left\{\{L_i^f\}_f,\{F_i^s,\alpha_i^s,\gamma_i^s,\eta_i^s,\mu_i^s,\kappa_i^s\}_s,\{\bar{r}_{ij}^s,\bar{f}_{ij}^s\}_{j,s},\{\beta_i^{s,f}\}_{f,s},\{\theta_i^{sk}\}_{k,s},\lambda_i\right\}_i$, an equilibrium vector $\left\{P_i,\{n_{ij}^s,\bar{x}_{ij}^s\}_{j,s},\{P_i^s,N_i^s,\bar{p}_i^s,W_i^s,V_i^s\bar{p}_i^s q_i^s,E_i^s,c_i^s\}_s,\{m_i^{sk}\}_{k,s},\{l_i^{s,f}\}_{f,s},\{w_i^f\}_f\right\}_i$ satisfies the following conditions.

1. The extensive and intensive margins of firm-level sales, n_{ij}^s and $\bar{x}_{ij}^s$, satisfy (A.17) and (A.18) for all s, i and j.
2. The price of the final sectoral good P_j^s is given by (A.19). The final consumption good price P_i is given by (A.2).
3. The number of entrants in sector s of country i N_i^s satisfies the free entry condition in (A.20).
4. The price of the the intermediate sector good $\bar{p}_i^s$ is given by (A.4) where W_i^s and V_i^s are given by (A.5).
5. The total revenue of the intermediate sectoral good $\bar{p}_i^s q_i^s$ is given by (A.21).
6. Spending on the final sectoral good E_i^s is (A.13) with final consumption spending share c_i^s given by (A.1) and the intermediate consumption spending share m_i^{sk} given by (A.6).
7. Factor price w_i^s implies that the factor market clearing in (A.14) holds with l_i^f given by (A.6).

A.1.4 Nonparametric Counterfactual Predictions

We now use the equilibrium characterization above to compute counterfactual changes in aggregate outcomes using the functions $\epsilon_{ij}^s(n)$ and $\rho_{ij}^s(n)$. As in our baseline model, this implies that we do not need any parametric restrictions in the distribution of firm heterogeneity G_i. The implications of firm heterogeneity for the model's aggregate counterfactual predictions are summarized by $\epsilon_{ij}^s(n)$ and $\rho_{ij}^s(n)$.

The extensive and intensive margins of firm-level sales in (A.17) and (A.18) imply that

$$\ln \frac{\epsilon_{ij}^{s}(n_{ij}^{s}\hat{n}_{ij}^{s})}{\epsilon_{ij}^{s}(n_{ij}^{s})} = \ln\left(\hat{\bar{f}}_{ij}^{s}/\hat{\bar{r}}_{ij}^{s}\right) + \ln\left(\hat{\bar{p}}_{i}^{s}\right)^{\sigma^{s}} - \ln \hat{E}_{j}^{s}\left(\hat{P}_{j}^{s}\right)^{\sigma^{s}-1}. \tag{A.22}$$

$$\ln \hat{\bar{x}}_{ij}^{s} - \ln \frac{\rho_{ij}^{s}(n_{ij}^{s}\hat{n}_{ij}^{s})}{\rho_{ij}^{s}(n_{ij}^{s})} = \ln\left(\hat{\bar{r}}_{ij}^{s}\right) + \ln\left(\hat{\bar{p}}_{i}^{s}\right)^{1-\sigma^{s}} + \ln \hat{E}_{j}^{s}\left(\hat{P}_{j}^{s}\right)^{\sigma^{s}-1}. \tag{A.23}$$

The price of the final sectoral good P_j^s in (A.19) implies that

$$(\hat{P}_{j}^{s})^{1-\sigma^{s}} = \sum_{i} \frac{\bar{x}_{ij}^{s} n_{ij}^{s} N_{i}^{s}}{E_{j}^{s}}\left(\hat{\bar{r}}_{ij}^{s}\left(\hat{\bar{p}}_{i}^{s}\right)^{1-\sigma^{s}} \frac{\rho_{ij}^{s}(n_{ij}^{s}\hat{n}_{ij}^{s})}{\rho_{ij}^{s}(n_{ij}^{s})}\hat{n}_{ij}^{s}\hat{N}_{i}^{s}\right).$$

Let $x_{ij}^{s} \equiv \bar{x}_{ij}^{s} n_{ij}^{s} N_{i}^{s}/E_{j}^{s} = X_{ij}^{s}/\left(\sum_{o} X_{oj}^{s}\right)$ be the spending share of country j on country i. Thus,

$$(\hat{P}_{j}^{s})^{1-\sigma^{s}} = \sum_{i} x_{ij}^{s}\left(\hat{\bar{r}}_{ij}^{s}\left(\hat{\bar{p}}_{i}^{s}\right)^{1-\sigma^{s}} \frac{\rho_{ij}^{s}(n_{ij}^{s}\hat{n}_{ij}^{s})}{\rho_{ij}^{s}(n_{ij}^{s})}\hat{n}_{ij}^{s}\hat{N}_{i}^{s}\right). \tag{A.24}$$

The final consumption good price P_i in (A.2) implies that

$$\hat{P}_{j}^{1-\lambda_{j}} = \sum_{k} c_{j}^{s}\left(\hat{P}_{j}^{k}\right)^{1-\lambda_{j}}. \tag{A.25}$$

The free entry condition in (A.20) implies that

$$\hat{\bar{p}}_{i}^{s}\hat{\bar{F}}_{i}^{s}\sum_{j} n_{ij}^{s}\bar{x}_{ij}^{s}\left(1-\frac{\int_{0}^{n_{ij}^{s}}\frac{\tilde{\rho}_{ij}^{s}(n)}{\epsilon_{ij}^{s}(n)}\,dn}{\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}^{s}(n)}{\epsilon_{ij}^{s}(n_{ij})}\,dn}\right) = \sum_{j} n_{ij}^{s}\bar{x}_{ij}^{s}(\hat{n}_{ij}^{s}\hat{\bar{x}}_{ij}^{s})\left(1-\frac{\int_{0}^{n_{ij}^{s}\hat{n}_{ij}^{s}}\frac{\tilde{\rho}_{ij}^{s}(n)}{\epsilon_{ij}^{s}(n)}\,dn}{\int_{0}^{n_{ij}^{s}\hat{n}_{ij}^{s}}\frac{\tilde{\rho}_{ij}^{s}(n)}{\epsilon_{ij}^{s}(n_{ij}^{s}\hat{n}_{ij}^{s})}\,dn}\right). \tag{A.26}$$

The price of the the intermediate sector good $\bar{p}_i^s$ in (A.4) implies that

$$\hat{\bar{p}}_{i}^{s} = \left[\tilde{\alpha}_{i}^{s}\left(\hat{W}_{i}^{s}\right)^{1-\mu_{i}^{s}} + (1-\tilde{\alpha}_{i}^{s})\left(\hat{V}_{i}^{s}\right)^{1-\mu_{i}^{s}}\right]^{\frac{1}{1-\mu_{i}^{s}}}, \tag{A.27}$$

where $\tilde{\alpha}_i^s$ is the share of labor in total cost of sector s in country i.

From (A.5), $\hat{W}_i^s$ and $\hat{V}_i^s$ are given by

$$\hat{W}_{i}^{s} = \left[\sum_{f} l_{i}^{s,f}\left(\hat{w}_{i}^{f}\right)^{1-\eta_{i}^{s}}\right]^{\frac{1}{1-\eta_{i}^{s}}} \quad \text{and} \quad \hat{V}_{i}^{s} = \left[\sum_{k} m_{i}^{ks}\left(\hat{P}_{i}^{k}\right)^{1-\kappa_{i}^{s}}\right]^{\frac{1}{1-\kappa_{i}^{s}}}. \tag{A.28}$$

The total revenue of the intermediate sectoral good $\bar{p}_i^s q_i^s$ in (A.21) implies

$$\widehat{\bar{p}_{i}^{s}q_{i}^{s}} = \sum_{j} x_{ij}^{s}\hat{N}_{i}^{s}\hat{n}_{ij}^{s}\hat{\bar{x}}_{ij}^{s}.$$

Let $\iota_i \equiv w_i L_i/(w_i L_i + T_i)$. Spending on the final sectoral good E_i^s in (A.13) implies that

$$\hat{E}_{i}^{s} = \hat{c}_{j}^{s}\frac{w_{i}L_{i}+T_{i}}{E_{j}}\left(\iota_{i}\hat{w}_{i} + (1-\iota_{i})\hat{T}_{i}\right) + \sum_{k}\frac{\tilde{M}_{i}^{ks}}{E_{i}^{s}}\widehat{\bar{p}_{i}^{s}q_{i}^{s}}.$$

where $\tilde{M}_i^{ks}$ is the value of intermediate sales of sector k to s in country i.

The final consumption spending share c_i^s in (A.1) implies that

$$\hat{c}_j^s = \left(\frac{\hat{P}_j^s}{\hat{P}_j}\right)^{1-\lambda_j}. \tag{A.29}$$

The intermediate consumption spending share m_i^{sk} in (A.6) implies that

$$\hat{m}_i^{ks} = \left(\frac{\hat{P}_i^k}{\hat{V}_i^s}\right)^{1-\kappa_i^s} \left(\frac{\hat{V}_i^s}{\hat{\bar{p}}_i^s}\right)^{1-\mu_i^s}. \tag{A.30}$$

The labor spending share l_i^f in (A.6).

$$\hat{l}_i^{s,f} = \left(\frac{\hat{w}_i^f}{\hat{W}_i^s}\right)^{1-\eta_i^s} \left(\frac{\hat{W}_i^s}{\hat{\bar{p}}_i^s}\right)^{1-\mu_i^s}. \tag{A.31}$$

The factor market clearing in (A.14) implies that

$$\hat{w}_i^f = \sum_s \zeta_i^{f,s} \hat{l}_i^{s,f} \left(\widehat{\bar{p}_i^s q_i^s}\right) \tag{A.32}$$

where $\zeta_i^{f,s}$ is the share of factor f income coming from sector s in country i.

Thus, the system (A.22)–(A.31) determines the counterfactual predictions in the model.

A.2 Allowing for zero bilateral trade

In this section, we extend our baseline framework to allow zero trade flows between two countries.

A.2.1 Environment

Consider the same environment described in Section 2.1.

A.2.2 Extensive and Intensive Margin of Firm Export

As in our baseline, we consider the distribution of $(r_{ij}(\omega), e_{ij}(\omega))$ implied by $G_i(.)$:

$$r_{ij}(\omega) \sim H_{ij}^r(r|e), \qquad \text{and} \qquad e_{ij}(\omega) \sim H_{ij}^e(e). \tag{A.33}$$

To allow for zero trade flows, we follow Helpman et al. (2008) by allows the support of the entry potential distribution to be bounded. Specifically, assume that $H_{ij}(e)$ has full support over $[0, \bar{e}_{ij}]$.

Extensive margin of firm-level exports. Recall that $n_{ij} \equiv Pr[\omega \in \Omega_{ij}]$ where Ω_{ij} is given by (5). It implies that

$$n_{ij} = \begin{cases} 1 - H_{ij}^e(e_{ij}^*) & \text{if} \quad e_{ij}^* \leq \bar{e}_{ij} \\ 0 & \text{if} \quad e_{ij}^* > \bar{e}_{ij} \end{cases} \qquad \text{where} \quad e_{ij}^* \equiv \sigma \frac{\bar{f}_{ij}}{\bar{r}_{ij}} \left[\left(\frac{w_i}{P_j}\right)^\sigma \frac{P_j}{E_j}\right].$$

Let us now define the extensive margin function as

$$\tilde{\epsilon}_{ij}(n_{ij}) \equiv \begin{cases} \left(H_{ij}^{e}\right)^{-1}(1-n) & \text{if} \quad n>0 \\ \bar{e}_{ij} & \text{if} \quad n=0 \end{cases}.$$

Using this definition and the expression for n_{ij} above, we get that

$$\tilde{\epsilon}_{ij}(n_{ij}) = \min\left\{e_{ij}^{*}, \bar{e}_{ij}\right\}.$$

We then define $\epsilon_{ij}(n) \equiv \tilde{\epsilon}_{ij}(n_{ij})/\bar{e}_{ij}$ and $\tilde{f}_{ij} \equiv \bar{f}_{ij}/\bar{e}_{ij}$. Then,

$$\ln \epsilon_{ij}(n_{ij}) = \min\left\{\ln\left(\sigma \tilde{f}_{ij}\bar{r}_{ij}\right) + \ln\left(w_i^{\sigma}\right) - \ln\left(E_j P_j^{\sigma-1}\right), 0\right\}. \tag{A.34}$$

Intensive margin of firm-level exports. Conditional on $n_{ij} > 0$, we now compute the average revenue in j:

$$\bar{x}_{ij} = \bar{r}_{ij}\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma} E_j\right]\frac{1}{n_{ij}}\int_{e_{ij}^{*}}^{\underline{e}_{ij}} E[r|e]\ dH_{ij}^{e}(e).$$

We consider the transformation $n = 1 - H_{ij}^{e}(e)$ such that $e = \epsilon_{ij}(n)$ and $dH_{ij}(e) = -dn$. By defining $\tilde{\rho}_{ij}(n) \equiv E[r|e = \epsilon_{ij}(n)]$ and $\rho_{ij}(0) = 0$,

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln\left(\bar{r}_{ij}\right) + \ln\left(w_i^{1-\sigma}\right) + \ln\left(E_j P_j^{\sigma-1}\right). \tag{A.35}$$

A.2.3 General Equilibrium

We now write the equilibrium conditions in terms of $\rho_{ij}(.)$ and $\epsilon_{ij}(.)$. Since $x_{ij} = \bar{x}_{ij}n_{ij}N_i/E_j$ and $\sum_i x_{ij} = 1$, the expression above immediately implies that

$$P_j^{1-\sigma} = \sum_{i:n_{ij}>0} \bar{r}_{ij}\left(w_i\right)^{1-\sigma}\bar{\rho}_{ij}(n_{ij})\left(n_{ij}N_i\right) \tag{A.36}$$

We then turn to the free entry condition in (7). Following the same steps as in Appendix A.1, it is straight forward to show that

$$\sigma w_i F_i = \sum_j n_{ij}\bar{x}_{ij} - \sum_j \left(\sigma w_i \bar{f}_{ij}\right)\int_0^{n_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

For $n_{ij} > 0$, the ratio of (A.34) and (A.35) implies that $\sigma \bar{f}_{ij}w_i = \bar{x}_{ij}\epsilon_{ij}(n_{ij})/\rho_{ij}(n_{ij})$. Thus,

$$\sigma w_i F_i = \sum_{j:n_{ij}>0} n_{ij}\bar{x}_{ij}\left(1 - \frac{\int_0^{n_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn}{\int_0^{n_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij})}\ dn}\right). \tag{A.37}$$

Following the same steps as in Appendix A.1, it is straight forward to show that

$$w_i L_i = \sum_{j:n_{ij}>0} N_i n_{ij}\bar{x}_{ij}. \tag{A.38}$$

Thus, given $\left\{\bar{L}_i, \bar{F}_i, \{\bar{r}_{ij}, \tilde{f}_{ij}\}_j\right\}_i$, an equilibrium vector $\{\{n_{ij}, \bar{x}_{ij}\}_j, P_i, N_i, E_i, w_i\}_i$ satisfies the following conditions.

1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, satisfy (A.34) and (A.35) for all i and j.
2. For all i, the price index is given by (A.36).
3. For all i, free entry is given by (A.37).
4. For all i, total spending, E_i, satisfies (8).
5. For all i, the labor market clearing is given by (A.38).

A.2.4 Nonparametric Counterfactual Predictions

We now use the equilibrium characterization above to compute counterfactual changes in aggregate outcomes using the functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. We further assume that, in bilateral pairs for which initially $n_{ij} = 0$, we still have that $n'_{ij} = 0$. Thus, (A.34) implies that

$$\frac{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}{\epsilon_{ij}(n_{ij})} = \frac{\hat{\bar{f}}_{ij}}{\hat{\bar{r}}_{ij}} \left[\left(\frac{\hat{w}_i}{\hat{P}_j} \right)^{\sigma} \frac{\hat{P}_j}{\hat{E}_j} \right] \quad \text{for} \quad n_{ij} > 0 \tag{A.39}$$

$$n'_{ij} = 0 \quad \text{for} \quad n_{ij} = 0. \tag{A.40}$$

The intensive margin equation remains the same:

$$\hat{\bar{x}}_{ij} = \hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \left[\left(\frac{\hat{w}_i}{\hat{P}_j} \right)^{1-\sigma} \hat{E}_j \right] \quad \text{for} \quad n_{ij} > 0. \tag{A.41}$$

The price index equation in (A.36) implies that

$$\hat{P}_j^{1-\sigma} = \sum_{i:n_{ij}>0} x_{ij} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-\sigma} \hat{n}_{ij} \hat{N}_i \right). \tag{A.42}$$

The spending equation in (8) implies that

$$\hat{E}_i = \iota_i \left(\hat{w}_i \hat{\bar{L}}_i \right) + (1 - \iota_i) \hat{\bar{T}}_i, \tag{A.43}$$

The labor market clearing condition in (A.38) implies

$$\hat{w}_i \hat{\bar{L}}_i = \sum_{j:n_{ij}>0} y_{ij} \left(\hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij} \right). \tag{A.44}$$

The free entry condition in (A.37) implies that

$$\hat{w}_i \sum_{j:n_{ij}>0} n_{ij} \bar{x}_{ij} \left(1 - \frac{\int_0^{n_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij})}\, dn} \right) = \sum_{j:n_{ij}>0} n_{ij} \bar{x}_{ij} (\hat{n}_{ij} \hat{\bar{x}}_{ij}) \left(1 - \frac{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\, dn} \right). \tag{A.45}$$

Thus, the system (A.39)–(A.45) determines the counterfactual predictions in the model.

A.3 Model with Import Tariffs

In this section, we follow Costinot and Rodriguez-Clare (2013) to extend our baseline framework to allow for import tariffs.

A.3.1 Environment

We assume that country j charges an ad-valorem tariff of t_{ij} such that the total trade costs between country i and j is $\bar{\tau}_{ij}(1+t_{ij})$. We consider a monopolistic competitive environment in which firms maximize profits given the demand in (1). For firm ω of country i, the optimal price in market j is $p_{ij}(\omega) = \frac{\sigma}{\sigma-1}\frac{\bar{\tau}_{ij}(1+t_{ij})w_i}{\bar{a}_i}\frac{\tau_{ij}(\omega)}{a_i(\omega)}$ with an associated revenue of

$$R_{ij}(\omega) = \bar{r}_{ij} r_{ij}(\omega) \left[\left(\frac{w_i}{P_j} \right)^{1-\sigma} E_j \right], \tag{A.46}$$

where

$$r_{ij}(\omega) \equiv b_{ij}(\omega) \left(\frac{\tau_{ij}(\omega)}{a_i(\omega)} \right)^{1-\sigma} \quad \text{and} \quad \bar{r}_{ij} \equiv \bar{b}_{ij} \left(\frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij}(1+t_{ij})}{\bar{a}_i} \right)^{1-\sigma}. \tag{A.47}$$

The firm's entry decision depends on the profit generated by the revenue in (A.46), $\sigma^{-1}(1+t_{ij})^{-1}R_{ij}(\omega)$, and the fixed-cost of entry, $w_i \bar{f}_{ij} f_{ij}(\omega)$. Specifically, firm ω of i enters j if, and only if, $\pi_{ij}(\omega) = \sigma^{-1}(1+t_{ij})^{-1}R_{ij}(\omega) - w_i \bar{f}_{ij} f_{ij}(\omega) \geq 0$. This yields the set of firms from i selling in j:

$$\omega \in \Omega_{ij} \quad \Leftrightarrow e_{ij}(\omega) \geq \sigma(1+t_{ij}) \frac{\bar{f}_{ij}}{\bar{r}_{ij}} \left[\left(\frac{w_i}{P_j} \right)^{\sigma} \frac{P_j}{E_j} \right], \tag{A.48}$$

where

$$e_{ij}(\omega) \equiv \frac{r_{ij}(\omega)}{f_{ij}(\omega)}. \tag{A.49}$$

The aggregate trade flows (including tariff) is still given by

$$X_{ij} = \int_{\omega \in \Omega_{ij}} R_{ij}(\omega)\, d\omega. \tag{A.50}$$

As before, free entry implies that N_i satisfies

$$\sum_j E\left[\max\left\{ \pi_{ij}(\omega);\ 0 \right\} \right] = w_i \bar{F}_i. \tag{A.51}$$

Market clearing. As shown by Costinot and Rodriguez-Clare (2013), the country's spending now also includes the tariff revenue:

$$E_i = w_i \bar{L}_i + \bar{T}_i + \sum_i \frac{t_{ij}}{1+t_{ij}} X_{ij}. \tag{A.52}$$

Now a fraction $t_{ij}/(1+t_{ij})$ of total revenue goes to the government of country j. So, labor in country i only receive a fraction $1/(1+t_{ij})$ of the sales revenue. Thus, $w_i L_i = (1+t_{ij})^{-1} \int_{\omega \in \Omega_{ij}} R_{ij}(\omega)\, d\omega$ and, by (A.47),

$$w_i \bar{L}_i = \frac{\bar{r}_{ij}}{1+t_{ij}} \left(\frac{w_i}{P_j} \right)^{1-\sigma} E_j \left[\int_{\omega \in \Omega_{ij}} r_{ij}(\omega)\, d\omega \right]. \tag{A.53}$$

A.3.2 Extensive and Intensive Margin of Firm Export

Using the same definitions of the baseline model, expression (A.48) yields

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left(\sigma(1+t_{ij})\bar{f}_{ij}/\bar{r}_{ij}\right) + \ln\left(w_i^{\sigma}\right) - \ln\left(E_j P_j^{\sigma-1}\right). \tag{A.54}$$

Again, following the same steps of the baseline model, equation (A.46) implies the same intensive margin equation:

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln\left(\bar{r}_{ij}\right) + \ln\left(w_i^{1-\sigma}\right) + \ln\left(E_j P_j^{\sigma-1}\right). \tag{A.55}$$

A.3.3 General Equilibrium

Part 1. To derive the labor market clearing condition notice that there are three sources of demand for labor: production of goods, fixed-cost of entering a market and fixed-cost of creating a variety. Thus,

$$w_i \bar{L}_i = \sum_j N_i Pr[\omega \in \Omega_{ij}]\left(1-\frac{1}{\sigma}\right)\frac{1}{1+t_{ij}} E\left[R_{ij}(\omega)\,|\omega \in \Omega_{ij}\right] + \sum_j N_i Pr[\omega \in \Omega_{ij}] w_i \bar{f}_{ij} E\left[f_{ij}(\omega)\,|\omega \in \Omega_{ij}\right] + N_i w_i \bar{F}_i$$

From the free entry condition, we know that

$$w_i \bar{F}_i = \sum_j \mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] = \sum_j Pr[\omega \in \Omega_{ij}]\left(\frac{1}{\sigma}\frac{1}{1+t_{ij}} E\left[R_{ij}(\omega)\,|\omega \in \Omega_{ij}\right] - w_i \bar{f}_{ij} E\left[f_{ij}(\omega)\,|\omega \in \Omega_{ij}\right]\right),$$

which implies that

$$w_i \bar{L}_i = \sum_j N_i Pr[\omega \in \Omega_{ij}] E\left[R_{ij}(\omega)\,|\omega \in \Omega_{ij}\right]\frac{1}{1+t_{ij}}.$$

Thus, since $\bar{x}_{ij} \equiv E\left[R_{ij}(\omega)\,|\omega \in \Omega_{ij}\right]$ and $n_{ij} = Pr[\omega \in \Omega_{ij}]$, this immediately implies that

$$w_i \bar{L}_i = \sum_j \frac{N_i n_{ij} \bar{x}_{ij}}{1+t_{ij}}. \tag{A.56}$$

Part 2. Since $p_{ij}(\omega) = \frac{\sigma}{\sigma-1}\frac{\bar{\tau}_{ij} w_i}{\bar{a}_i}\frac{\tau_{ij}(\omega)}{a_i(\omega)}$, the expression for $P_j^{1-\sigma}$ in (2) implies that

$$P_j^{1-\sigma} = \sum_i \left[\bar{b}_{ij}\left(\frac{\sigma}{\sigma-1}\frac{(1+t_{ij})\bar{\tau}_{ij}}{\bar{a}_i}\right)^{1-\sigma}\right]\left(w_i^{1-\sigma}\right)\int_{\Omega_{ij}}\left(b_{ij}(\omega)\right)\left(\frac{\tau_{ij}(\omega)}{a_i(\omega)}\right)^{1-\sigma} d\omega$$

Using the definitions in (4), we can write this expression as

$$P_j^{1-\sigma} = \sum_i \bar{r}_{ij}\left(w_i^{1-\sigma}\right)\int_{\Omega_{ij}} r_{ij}(\omega)\ d\omega$$

Notice that $\int_{\Omega_{ij}} r_{ij}(\omega)\ d\omega = N_i Pr[\omega \in \Omega_{ij}]E[r|\omega \in \Omega_{ij}] = N_i n_{ij}\rho_{ij}(n_{ij})$. This immediately yields

$$P_j^{1-\sigma} = \sum_i \bar{r}_{ij} w_i^{1-\sigma}\rho_{ij}(n_{ij}) n_{ij} N_i. \tag{A.57}$$

Part 3. We start by writing

$$\begin{aligned}\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] &= Pr[\omega\in\Omega_{ij}]E\left[\pi_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]+Pr[\omega\notin\Omega_{ij}]0\\ &= Pr[\omega\in\Omega_{ij}]\left(\tfrac{1}{\sigma}\tfrac{1}{1+t_{ij}}E\left[R_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]-w_i\bar{f}_{ij}E\left[f_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]\right)\\ &= n_{ij}\left(\tfrac{1}{\sigma}\tfrac{1}{1+t_{ij}}\bar{x}_{ij}-w_i\bar{f}_{ij}E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]\right)\end{aligned}$$

where the second equality follows from the expression for $\pi_{ij}(\omega)=\frac{1}{\sigma}\frac{1}{1+t_{ij}}R_{ij}(\omega)-w_i\bar{f}_{ij}f_{ij}\left(\omega\right)$, and the third equality follows from the definitions of $\bar{x}_{ij}\equiv E\left[R_{ij}\left(\omega\right)|\omega\in\Omega_{ij}\right]$ and $e_{ij}(\omega)\equiv r_{ij}(\omega)/f_{ij}(\omega)$.

By defining $e^{*}_{ij}\equiv\sigma(1+t_{ij})\frac{\bar{f}_{ij}}{\bar{r}_{ij}}\left[\left(\frac{w_i}{P_j}\right)^{\sigma}\frac{P_j}{E_j}\right]$, we can write

$$E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]=\int_{e^{*}_{ij}}^{\infty}\frac{1}{e}\left[\int_{0}^{\infty}rdH^{r}_{ij}\left(r|e\right)\right]\frac{dH^{e}(e)}{1-H^{e}(e^{*}_{ij})}$$

Consider the transformation $n=1-H_{ij}(e)$ such that $e=\bar{\epsilon}_{ij}(n)$. In this case, $dH_{ij}(e)=-dn$ and $n_{ij}=1-H_{ij}(e^{*}_{ij})$, which implies that

$$E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega\in\Omega_{ij}\right]=\frac{1}{n_{ij}}\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}dn.$$

Thus,

$$\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right]=\frac{1}{\sigma}\frac{1}{1+t_{ij}}n_{ij}\bar{x}_{ij}-w_i\bar{f}_{ij}\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Thus, the free entry condition is

$$\sigma w_i\bar{F}_i=\sum_j\frac{n_{ij}\bar{x}_{ij}}{1+t_{ij}}-\sum_j\left(\sigma w_i\bar{f}_{ij}\right)\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn. \tag{A.58}$$

Notice that the summation of (A.54) and (A.55) implies that

$$\ln\left(\sigma(1+t_{ij})w_i\bar{f}_{ij}\right)=\ln\bar{x}_{ij}-\ln\rho_{ij}(n_{ij})+\ln\epsilon_{ij}(n_{ij})$$

which yields

$$\sigma w_i\bar{F}_i=\sum_j\frac{n_{ij}\bar{x}_{ij}}{1+t_{ij}}-\sum_j\frac{n_{ij}\bar{x}_{ij}}{1+t_{ij}}\frac{\epsilon_{ij}(n_{ij})}{\rho_{ij}(n_{ij})}\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Using the market clearing condition in (OA.1), we have that

$$\frac{1}{N_i}=\sigma\frac{\bar{F}_i}{\bar{L}_i}+\sum_j\frac{n_{ij}\bar{x}_{ij}}{(1+t_{ij})w_i\bar{L}_i}\frac{\epsilon_{ij}(n_{ij})}{\rho_{ij}(n_{ij})}\int_{0}^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\ dn. \tag{A.59}$$

which immediately yields equation (OA.7).

Part 4. Equation (A.52) implies that

$$E_i=w_i\bar{L}_i+\bar{T}_i+\sum_j\frac{t_{ji}}{1+t_{ji}}(N_jn_{ji}\bar{x}_{ji}). \tag{A.60}$$

Thus, given $\left\{\bar{L}_i,\bar{F}_i,\{t_{ij},\bar{r}_{ij},\bar{f}_{ij}\}_j\right\}_i$, an equilibrium vector $\left\{\{n_{ij},\bar{x}_{ij}\}_j,P_i,N_i,E_i,w_i\right\}_i$ satisfies the following conditions.

1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, satisfy (A.54) and (A.55) for all i and j.
2. For all i, the price index is given by (A.57).
3. For all i, free entry is given by (A.59).
4. For all i, total spending, E_i, satisfies (A.60).
5. For all i, the labor market clearing is given by (A.56).

A.3.4 Nonparametric Counterfactual Predictions

We now use the equilibrium characterization above to compute counterfactual changes in aggregate outcomes using the functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$.

From (A.54) and (A.55),

$$\frac{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}{\epsilon_{ij}(n_{ij})} = \widehat{(1+t_{ij})}\frac{\hat{\bar{f}}_{ij}}{\hat{\bar{r}}_{ij}}\frac{\hat{w}_i^{\sigma}}{\hat{E}_j\hat{P}_j^{\sigma-1}}. \tag{A.61}$$

$$\hat{\bar{x}}_{ij} = \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})}\hat{\bar{r}}_{ij}\frac{\hat{E}_j\hat{P}_j^{\sigma-1}}{\hat{w}_i^{\sigma-1}} \tag{A.62}$$

From (A.57),

$$\hat{P}_j^{1-\sigma} = \sum_i x_{ij}\hat{\bar{r}}_{ij}\hat{w}_i^{1-\sigma}\frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})}\hat{n}_{ij}\hat{N}_i. \tag{A.63}$$

From (A.59),

$$N_i\hat{N}_i = \left[\sigma\frac{\bar{F}_i}{\bar{L}_i}\frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{(1+t_{ij})w_i\bar{L}_i}\frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\widehat{(1+t_{ij})}\hat{w}_i\hat{\bar{L}}_i}\frac{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\,dn}\right]^{-1}$$

Using (A.59) to substitute for $\sigma\frac{\bar{F}_i}{\bar{L}_i}$,

$$\hat{N}_i = \left[\left(1-\sum_j y_{ij}\frac{\int_0^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij})}\,dn}\right)\frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j y_{ij}\frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\widehat{(1+t_{ij})}\hat{w}_i\hat{\bar{L}}_i}\frac{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\tilde{\rho}_{ij}(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\,dn}\right]^{-1}. \tag{A.64}$$

where $y_{ij} = \frac{X_{ij}}{(1+t_{ij})w_i\bar{L}_i}$ is the share of income in i from sales to j.

From (A.60),

$$\hat{E}_i = \iota_i\hat{w}_i\hat{\bar{L}}_i + \vartheta_i\hat{\bar{T}}_i + \sum_j\left(\frac{t_{ji}}{1+t_{ji}}\frac{\hat{t}_{ji}}{\widehat{(1+t_{ji})}}\right)\frac{X_{ji}}{E_i}(\hat{N}_j\hat{n}_{ji}\hat{\bar{x}}_{ji}). \tag{A.65}$$

where $\iota_i \equiv Y_i/E_i$ and $\vartheta_i \equiv \bar{T}_i/E_i$.

Thus, the system (A.61)–(A.65) determines the counterfactual predictions in the model.

A.4 Multi-product Firms

In this section, we extend our framework to incorporate multi-product firms.

A.4.1 Environment

Demand. We maintain the assumption that each country j has a representative household that inelastically supplies $\bar{L}_j$ units of labor. The demand for variety ω from country i is

$$q_{ij}(\omega) = \bar{b}_{ij}\left(\frac{p_{ij}(\omega)}{P_j}\right)^{-\sigma}\frac{E_j}{P_j}, \tag{A.66}$$

where, in market j, E_j is the total spending, $p_{ij}(\omega)$ is the price of variety ω of country i, and P_j is the CES price index,

$$P_j^{1-\sigma} = \sum_i \int_{\Omega_{ij}^v} \bar{b}_{ij}\left(p_{ij}(\omega)\right)^{1-\sigma}\ d\omega, \tag{A.67}$$

and Ω_{ij}^v is the set of varieties produced in country i that are sold in country j.

Production. We consider a monopolistic competitive environment. Each firm η can choose how many varieties to sell in each market. In order to operate in market j, the firm must pay a fixed entry cost $w_i\bar{f}_{ij}f_{ij}(\eta)$. Conditional on entry, selling N varieties entails a labor cost of $w_i\frac{1}{1+1/\alpha}N^{1+1/\alpha}$. For every variety, the firm then has a unit production cost of $w_i\frac{\tau_{ij}(\eta)}{a_i(\eta)}\frac{\bar{\tau}_{ij}}{\bar{a}_i}$.

For each variety ω of firm η from country i, the optimal price in market j is $p_{ij}(\omega) = \frac{\sigma}{\sigma-1}\frac{\bar{\tau}_{ij}w_i}{\bar{a}_i}\frac{\tau_{ij}(\eta)}{a_i(\eta)}$ with an associated revenue of

$$R_{ij}^N(\eta) = \bar{r}_{ij}^N r_{ij}^N(\eta)\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma}E_j\right], \tag{A.68}$$

where

$$r_{ij}^N(\eta) \equiv \left(\frac{\tau_{ij}(\eta)}{a_i(\eta)}\right)^{1-\sigma} \quad \text{and} \quad \bar{r}_{ij}^N \equiv \bar{b}_{ij}\left(\frac{\sigma}{\sigma-1}\frac{\bar{\tau}_{ij}}{\bar{a}_i}\right)^{1-\sigma}. \tag{A.69}$$

The firm then decides how many varieties to sell by solving the following problem:

$$\max_N \frac{1}{\sigma}R_{ij}^N(\eta)N - w_i\frac{1}{1+1/\alpha}N^{1+1/\alpha},$$

which implies that

$$N_{ij}(\eta) = \left(\frac{1}{\sigma}\frac{R_{ij}^N(\eta)}{w_i}\right)^{\alpha}. \tag{A.70}$$

Thus, firm sales are

$$R_{ij}(\eta) = N_{ij}(\eta)R_{ij}^N(\eta) = \frac{1}{\sigma^\alpha w_i^\alpha}\left(\bar{r}_{ij}^N r_{ij}^N(\eta)\right)^{1+\alpha}\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma}E_j\right]^{1+\alpha}.$$

To simplify the notation, conditional on entering market j, the sales of firm η can be written as

$$R_{ij}(\eta) = \bar{r}_{ij}r_{ij}(\eta)\,w_i^{1-(1+\alpha)\sigma}\left(E_jP_j^{\sigma-1}\right)^{1+\alpha} \tag{A.71}$$

$$r_{ij}(\eta) \equiv \left(r_{ij}^N(\eta)\right)^{1+\alpha} \quad \text{and} \quad \bar{r}_{ij} \equiv \frac{1}{\sigma^\alpha}\left(\bar{r}_{ij}^N\right)^{1+\alpha}. \tag{A.72}$$

Conditional on entering market j, the firm's profit in that market is

$$
\begin{aligned}
\pi_{ij}(\eta) \quad = \quad & N_{ij}(\eta)\tfrac{1}{\sigma}R^N_{ij}(\eta) - w_i\tfrac{1}{1+1/\alpha}N_{ij}(\eta)^{1+1/\alpha} - w_i\bar{f}_{ij}f_{ij}(\eta) \\
& \left(\tfrac{1}{\sigma}\tfrac{R^N_{ij}(\eta)}{w_i}\right)^{\alpha}\tfrac{1}{\sigma}R^N_{ij}(\eta) - w_i\tfrac{1}{1+1/\alpha}\left(\tfrac{1}{\sigma}\tfrac{R^N_{ij}(\eta)}{w_i}\right)^{1+\alpha} - w_i\bar{f}_{ij}f_{ij}(\eta) \\
& \tfrac{1}{(1+\alpha)\sigma}\tfrac{1}{\sigma^{\alpha}w_i^{\alpha}}\left(R^N_{ij}(\eta)\right)^{1+\alpha} - w_i\bar{f}_{ij}f_{ij}(\eta)
\end{aligned}
$$

and, therefore,

$$\pi_{ij}(\eta) = \frac{1}{(1+\alpha)\sigma}\bar{r}_{ij}r_{ij}(\eta)\,w_i^{1-(1+\alpha)\sigma}\left(E_jP_j^{\sigma-1}\right)^{1+\alpha} - w_i\bar{f}_{ij}f_{ij}(\eta). \tag{A.73}$$

The cost of variety creation is $C^V_{ij}(\eta) = w_i\frac{1}{1+1/\alpha}\left(\frac{1}{\sigma}\frac{R^N_{ij}(\eta)}{w_i}\right)^{1+\alpha}$, which can be written as

$$C^V_{ij}(\eta) = \frac{\alpha}{(1+\alpha)\sigma}\bar{r}_{ij}r_{ij}(\eta)\,w_i^{1-(1+\alpha)\sigma}\left(E_jP_j^{\sigma-1}\right)^{1+\alpha}. \tag{A.74}$$

Firm η of i sells in j if, and only if profits are positive, $\pi_{ij}(\eta) \geq 0$. This yields the set of firms of country i operating in j:

$$\eta \in \Omega_{ij} \quad \Leftrightarrow e_{ij}(\eta) \geq (1+\alpha)\sigma\frac{\bar{f}_{ij}}{\bar{r}_{ij}}\frac{w_i^{(1+\alpha)\sigma}}{\left(E_jP_j^{\sigma-1}\right)^{1+\alpha}} \tag{A.75}$$

where

$$e_{ij}(\eta) \equiv \frac{r_{ij}(\eta)}{f_{ij}(\eta)}. \tag{A.76}$$

Entry. An entrant firm pays a fixed labor cost $\bar{F}_i$ to draw its type from an arbitrary distribution:

$$v_i(\eta) \equiv \left\{a_i(\eta), \tau_{ij}(\eta), f_{ij}(\eta)\right\}_j \sim G^v_i(v), \tag{A.77}$$

In equilibrium, N_i firms pay the fixed cost of entry in exchange for an ex-ante expected profit of zero. The free entry implies that

$$\sum_j E\left[\max\left\{\pi_{ij}(\eta);\ 0\right\}\right] = w_i\bar{F}_i. \tag{A.78}$$

Market clearing. We follow Dekle et al. (2008) by introducing exogenous international transfers, so that spending is

$$E_i = w_i\bar{L}_i + \bar{T}_i, \quad \sum_i \bar{T}_i = 0. \tag{A.79}$$

Since labor is the only factor of production, labor income in i equals the total revenue of firms from i: $w_iL_i = \sum_j \int R_{ij}(\eta)d\eta$. Given the expression in (A.71),

$$w_i\bar{L}_i = \sum_j \bar{r}_{ij}w_i^{1-(1+\alpha)\sigma}\left(E_jP_j^{\sigma-1}\right)^{1+\alpha}\int_{\eta\in\Omega_{ij}} r_{ij}(\eta)\,d\eta. \tag{A.80}$$

Equilibrium. Given the arbitrary distribution in (A.77), the equilibrium is defined as the vector $\{P_i, \{\Omega_{ij}\}_j, N_i, E_i, w_i\}_i$ satisfying equations (A.67), (A.75), (A.78), (A.79), (A.80) for all i.

A.4.2 Extensive and Intensive Margin of Firm Exports

The extensive and intensive margin of firm exports follows the single-product case. We now use the definitions of entry and revenue potentials to characterize firm-level entry and sales in different markets in general equilibrium. We consider the CDF of $(r_{ij}(\eta), e_{ij}(\eta))$ implied by G_i. We assume that

$$r_{ij}(\eta) \sim H_{ij}^{r}(r|e), \qquad \text{and} \qquad e_{ij}(\eta) \sim H_{ij}^{e}(e). \tag{A.81}$$

Extensive margin of firm-level exports. The share of firms of country i serving market j is $n_{ij} = Pr[\eta \in \Omega_{ij}]$. Defining $\epsilon_{ij}(n) \equiv \left(H_{ij}^{e}\right)^{-1}(1-n)$, equation (A.75) yields

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left((1+\alpha)\sigma\frac{\bar{f}_{ij}}{\bar{r}_{ij}}\right) + \ln w_i^{(1+\alpha)\sigma} - \ln\left(E_j P_j^{\sigma-1}\right)^{1+\alpha}. \tag{A.82}$$

Intensive margin of firm-level exports. The average revenue of firms from country i in country j is $\bar{x}_{ij} \equiv E\left[R_{ij}(\eta)\,|\eta \in \Omega_{ij}\right]$ where $R_{ij}(\eta)$ is given by (A.71). Define the average revenue potential of exporters when n_{ij}% of $i's$ firms in sector s export to j as $\rho_{ij}(n_{ij}) \equiv \frac{1}{n_{ij}}\int_0^{n_{ij}} \rho_{ij}^{m}(n)\,dn$ where $\rho_{ij}^{m}(n) \equiv E[r|e = \epsilon_{ij}(n)]$ is the average revenue potential in quantile n of the entry potential distribution. Using the transformation $n = 1 - H_{ij}^{e}(e)$ such that $e = \epsilon_{ij}(n)$ and $dH_{ij}^{e}(e) = -dn$, we can follow the same steps as in the baseline model to show that

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln \bar{r}_{ij} + \ln w_i^{1-(1+\alpha)\sigma} + \ln\left(E_j P_j^{\sigma-1}\right)^{1+\alpha}. \tag{A.83}$$

Extensive margin of products per firm. The average number of products per firm of i operating in market j is $N_{ij}^{v} = E\left[N_{ij}(\eta)\,|\eta \in \Omega_{ij}\right]$. The expression for $N_{ij}(\eta)$ in (A.70) implies that

$$N_{ij}^{v} = \frac{1}{\sigma^{\alpha}}\left(\bar{r}_{ij}^{N}\right)^{\alpha} w_i^{-\alpha\sigma}\left(E_j P_j^{\sigma-1}\right)^{\alpha} E\left[\left(r_{ij}^{N}(\eta)\right)^{\alpha}|\eta \in \Omega_{ij}\right]$$

and, since $\bar{r}_{ij} \equiv \frac{1}{\sigma^{\alpha}}\left(\bar{r}_{ij}^{N}\right)^{1+\alpha}$,

$$N_{ij}^{v} = \sigma^{-\frac{\alpha}{1+\alpha}}\bar{r}_{ij}^{\frac{\alpha}{1+\alpha}} w_i^{-\alpha\sigma}\left(E_j P_j^{\sigma-1}\right)^{\alpha} E\left[\left(r_{ij}(\eta)\right)^{\frac{\alpha}{1+\alpha}}|\eta \in \Omega_{ij}\right].$$

To obtain the gravity equation, we consider a similar transformation as the one above. Define $\rho_{ij}^{v}(n_{ij}) \equiv \frac{1}{n_{ij}}\int_0^{n_{ij}} \tilde{\rho}_{ij}^{v}(n)\,dn$ where $\tilde{\rho}_{ij}^{v}(n) \equiv E[r^{\frac{\alpha}{1+\alpha}}|e = \epsilon_{ij}(n)]$. Using the transformation $n = 1 - H_{ij}^{e}(e)$ such that $e = \epsilon_{ij}(n)$ and $dH_{ij}^{e}(e) = -dn$, we can follow the same steps as in the baseline model to show that

$$\ln N_{ij}^{v} - \ln \rho_{ij}^{v}(n_{ij}) = \ln \sigma^{-\frac{\alpha}{1+\alpha}} + \frac{\alpha}{1+\alpha}\ln \bar{r}_{ij} + \ln w_i^{-\alpha\sigma} + \ln\left(E_j P_j^{\sigma-1}\right)^{\alpha}. \tag{A.84}$$

The elasticity of the average number of varieties per firm with respect to changes in bilateral revenue shifters is $\alpha/(1+\alpha)$, conditional on the composition control function, $\rho_{ij}^{v}(n_{ij})$, and the origin and destination fixed-effects. Thus, this semiparametric gravity equation can be used to estimate the parameter α. We show below that this parameter is necessary for counterfactual analysis.

A.4.3 General Equilibrium

Part 1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, satisfy (A.82) and (A.83) for all i and j. Together with N_i, they determine bilateral trade flows, $X_{ij} = N_i n_{ij}\bar{x}_{ij}$.

Part 2. For all i, total spending, E_i, satisfies (A.79).

Part 3. To derive the labor market clearing condition notice that there are four sources of demand for labor: production of goods, cost of producing new varieties, fixed-cost of entering a market, and fixed-cost of creating a firm. Thus,

$$\begin{array}{rcl} w_i\bar{L}_i & = & \sum_j N_i Pr[\eta\in\Omega_{ij}]\left(1-\frac{1}{\sigma}\right)E\left[R_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right] \\ & + & \sum_j N_i Pr[\eta\in\Omega_{ij}]E\left[C_{ij}^V(\eta)|\eta\in\Omega_{ij}\right] \\ & + & \sum_j N_i Pr[\eta\in\Omega_{ij}]w_i\bar{f}_{ij}E\left[f_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right] \\ & + & N_i w_i\bar{F}_i \end{array}$$

From the free entry condition, we know that $w_i\bar{F}_i=\sum_j\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right]$. Thus,

$$w_i\bar{F}_i=\sum_j Pr[\eta\in\Omega_{ij}]\left(\frac{1}{\sigma}E\left[R_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right]-E\left[C_{ij}^V(\eta)|\eta\in\Omega_{ij}\right]-w_i\bar{f}_{ij}E\left[f_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right]\right),$$

which implies that

$$w_i\bar{L}_i=\sum_j N_i Pr[\eta\in\Omega_{ij}]E\left[R_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right].$$

Thus, since $\bar{x}_{ij}\equiv E\left[R_{ij}\left(\eta\right)|\eta\in\Omega_{ij}\right]$ and $n_{ij}\equiv Pr[\eta\in\Omega_{ij}]$, this immediately implies

$$w_i\bar{L}_i=\sum_j N_i n_{ij}\bar{x}_{ij}. \tag{A.85}$$

Part 4. Since $p_{ij}\left(\eta\right)=\frac{\sigma}{\sigma-1}\frac{\bar{\tau}_{ij}w_i}{\bar{a}_i}\frac{\tau_{ij}(\eta)}{a_i(\eta)}$ for every variety of firm η, the expression for $P_j^{1-\sigma}$ in (A.67) implies that

$$P_j^{1-\sigma}=\sum_i\bar{b}_{ij}\int_{\Omega_{ij}}N_{ij}(\eta)\left(p_{ij}\left(\eta\right)\right)^{1-\sigma}\ d\eta.$$

Using the expression for $N_{ij}(\eta)$ in (A.70) and the definitions in (A.72), this expression can be written as

$$P_j^{1-\sigma}=\sum_i w_i^{1-\sigma}\frac{\bar{r}_{ij}}{w_i^{\alpha}}\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma}\frac{E_j}{w_i}\right]^{\alpha}\int_{\Omega_{ij}}r_{ij}\left(\eta\right)\ d\eta.$$

Notice that $\int_{\Omega_{ij}}r_{ij}\left(\eta\right)\ d\eta=N_i Pr[\eta\in\Omega_{ij}]E[r|\eta\in\Omega_{ij}]=N_i n_{ij}\rho_{ij}(n_{ij})$. This immediately yields

$$P_j^{1-\sigma}=\sum_i\bar{r}_{ij}w_i^{1-\sigma}\left[\left(\frac{w_i}{P_j}\right)^{1-\sigma}\frac{E_j}{w_i}\right]^{\alpha}\rho_{ij}(n_{ij})n_{ij}N_i,$$

and, therefore,

$$P_j^{1-\sigma}=\sum_i\bar{r}_{ij}w_i^{1-(1+\alpha)\sigma}\left(E_jP_j^{\sigma-1}\right)^{\alpha}\rho_{ij}(n_{ij})n_{ij}N_i. \tag{A.86}$$

Part 5. From (A.71) and (A.73),

$$\begin{array}{rcl} \mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] & = & Pr[\eta\in\Omega_{ij}]E\left[\frac{1}{(1+\alpha)\sigma}R_{ij}(\eta)-w_i\bar{f}_{ij}f_{ij}(\eta)|\eta\in\Omega_{ij}\right] \\ & & n_{ij}\left(\frac{1}{(1+\alpha)\sigma}\bar{x}_{ij}-w_i\bar{f}_{ij}E\left[r_{ij}(\eta)/e_{ij}(\eta)|\eta\in\Omega_{ij}\right]\right). \end{array}$$

By defining $e^*_{ij} \equiv (1+\alpha)\sigma \frac{\bar{f}_{ij}}{\bar{r}_{ij}} \frac{w_i^{(1+\alpha)\sigma}}{\left(E_j P_j^{\sigma-1}\right)^{1+\alpha}}$, we can write

$$E\left[r_{ij}(\eta)/e_{ij}(\eta)|\eta \in \Omega_{ij}\right] = \int_{e^*_{ij}}^{\infty} \frac{1}{e}\left[\int_0^{\infty} r dH^r_{ij}\left(r|e\right)\right] \frac{dH^e(e)}{1-H^e(e^*_{ij})}.$$

Consider the transformation $n = 1 - H_{ij}(e)$ such that $e = \bar{\epsilon}_{ij}(n)$. In this case, $dH_{ij}(e) = -dn$ and $n_{ij} = 1 - H_{ij}(e^*_{ij})$, which implies that

$$E\left[r_{ij}(\omega)/e_{ij}(\omega)|\omega \in \Omega_{ij}\right] = \frac{1}{n_{ij}} \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)} dn.$$

Thus,

$$\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] = \frac{1}{(1+\alpha)\sigma} n_{ij}\bar{x}_{ij} - w_i \bar{f}_{ij} \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Thus, the free entry condition is

$$(1+\alpha)\sigma w_i \bar{F}_i = \sum_j n_{ij}\bar{x}_{ij} - \sum_j \left((1+\alpha)\sigma w_i \bar{f}_{ij}\right) \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Notice that the summation of (A.82) and (A.83) implies that

$$\ln\left((1+\alpha)\sigma w_i \bar{f}_{ij}\right) = \ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) + \ln \epsilon_{ij}(n_{ij}),$$

which yields

$$(1+\alpha)\sigma w_i \bar{F}_i = \sum_j n_{ij}\bar{x}_{ij} - \sum_j \bar{x}_{ij} \frac{\epsilon_{ij}(n_{ij})}{\rho_{ij}(n_{ij})} \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

By substituting the definition of $\rho_{ij}(n)$, we can write the free entry condition as

$$(1+\alpha)\sigma w_i \bar{F}_i = \sum_j n_{ij}\bar{x}_{ij} - \sum_j n_{ij}\bar{x}_{ij} \frac{\epsilon_{ij}(n_{ij})}{\int_0^{n_{ij}} \rho^m_{ij}(n)\ dn} \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn.$$

Using the market clearing condition in (A.80), we have that

$$\frac{1}{N_i} = (1+\alpha)\sigma \frac{\bar{F}_i}{\bar{L}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\epsilon_{ij}(n_{ij})}{\int_0^{n_{ij}} \rho^m_{ij}(n)\ dn} \int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn,$$

which immediately yields

$$N_i = \left[(1+\alpha)\sigma \frac{\bar{F}_i}{\bar{L}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\ dn}{\int_0^{n_{ij}} \frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n_{ij})}\ dn}\right]^{-1}. \tag{A.87}$$

Part 6. The equilibrium vector $\{n_{ij}, \bar{x}_{ij}, E_i, w_i, P_i, N_i\}_{i,j}$ is determined by equations (A.82), (A.83), (A.79), (A.85), (A.86), and (A.87).

A.4.4 Nonparametric Counterfactual Predictions

We now use the equilibrium characterization above to compute counterfactual changes in aggregate outcomes using the functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$, as well as σ and α.

1. The extensive and intensive margins of firm-level sales, n_{ij} and $\bar{x}_{ij}$, in (A.82) and (A.83) imply

$$\frac{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}{\epsilon_{ij}(n_{ij})} = \frac{\hat{\bar{f}}_{ij}}{\hat{\bar{r}}_{ij}} \frac{\hat{w}_i^{(1+\alpha)\sigma}}{\left(\hat{E}_j \hat{P}_j^{\sigma-1}\right)^{1+\alpha}}, \tag{A.88}$$

$$\hat{\bar{x}}_{ij} = \hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \left[\hat{w}_i^{1-(1+\alpha)\sigma} \left(\hat{E}_j \hat{P}_j^{\sigma-1}\right)^{1+\alpha}\right]. \tag{A.89}$$

2. Let $\iota_i \equiv w_i L_i / E_i = \left(\sum_d X_{id}\right) / \left(\sum_o X_{oi}\right)$ be the output-spending ratio in country i in the initial equilibrium. The spending equation in (A.79) implies

$$\hat{E}_i = \iota_i \left(\hat{w}_i \hat{\bar{L}}_i\right) + (1-\iota_i)\hat{\bar{T}}_i, \tag{A.90}$$

3. Let $y_{ij} \equiv \left(N_i n_{ij} \bar{x}_{ij}\right) / \left(w_i L_i\right) = X_{ij} / \left(\sum_{j'} X_{ij'}\right)$ be the share of i's revenue from sales to j. The labor market clearing condition in (A.85) implies

$$\hat{w}_i \hat{\bar{L}}_i = \sum_j y_{ij} \left(\hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij}\right). \tag{A.91}$$

4. The price index (A.86) implies

$$\hat{P}_j^{1-\sigma} = \sum_i \frac{\bar{r}_{ij} w_i^{1-(1+\alpha)\sigma} \left(E_j P_j^{\sigma-1}\right)^{\alpha} \rho_{ij}(n_{ij}) n_{ij} N_i}{P_j^{1-\sigma}} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-(1+\alpha)\sigma} \left(\hat{E}_j \hat{P}_j^{\sigma-1}\right)^{\alpha} \hat{n}_{ij} \hat{N}_i\right).$$

Since $\bar{x}_{ij} = \rho_{ij}(n_{ij}) \bar{r}_{ij} w_i^{1-(1+\alpha)\sigma} \left(E_j P_j^{\sigma-1}\right)^{1+\alpha}$ and $x_{ij} = X_{ij}/E_j = \bar{x}_{ij} n_{ij} N_i / E_j$,

$$\begin{aligned} \hat{P}_j^{1-\sigma} &= \sum_i \frac{\bar{x}_{ij} n_{ij} N_i}{E_j} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-(1+\alpha)\sigma} \left(\hat{E}_j \hat{P}_j^{\sigma-1}\right)^{\alpha} \hat{n}_{ij} \hat{N}_i\right) \\ &= \sum_i x_{ij} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-(1+\alpha)\sigma} \left(\hat{E}_j \hat{P}_j^{\sigma-1}\right)^{\alpha} \hat{n}_{ij} \hat{N}_i\right). \end{aligned}$$

Rearranging this expression, we get that

$$\hat{P}_j^{(1-\sigma)(1+\alpha)} = \sum_i x_{ij} \left(\hat{\bar{r}}_{ij} \frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})} \hat{w}_i^{1-(1+\alpha)\sigma} \hat{E}_j^{\alpha} \hat{n}_{ij} \hat{N}_i\right). \tag{A.92}$$

5. The free entry condition in (A.87) implies

$$N_i \hat{N}_i = \left[(1+\alpha)\sigma \frac{\bar{F}_i}{\bar{L}_i} \frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\hat{w}_i \hat{\bar{L}}_i} \frac{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}\hat{n}_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\, dn}\right]^{-1}.$$

Using (A.87) to substitute for $(1+\alpha)\sigma\frac{\bar{F}_i}{\bar{L}_i}$,

$$\hat{N}_i = \left[\left(1-\sum_j y_{ij}\frac{\int_0^{n_{ij}}\frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}}\frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\,d}\right)\frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i}+\sum_j y_{ij}\frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\hat{w}_i\hat{\bar{L}}_i}\frac{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\,dn}\right]^{-1}. \tag{A.93}$$

Thus, the system (A.88)–(A.93) determines the counterfactual predictions in the model. Notice that the system only requires: (i) the elasticity functions $\{\rho_{ij}(n),\epsilon_{ij}(n)\}_{i,j}$, (ii) the elasticities σ and α, and (iii) the data in the initial equilibrium for trade flows and exporter firm shares, $\{X_{ij},n_{ij}\}_{i,j}$.

A.5 Non-CES Preferences

A.5.1 Environment

Demand. In country j with income y_j, we assume that the Marshallian demand function for product ω can be written as

$$q_j\left(p\left(\omega\right);\boldsymbol{p}_j,y_j\right)=q\left(p\left(\omega\right);P_j\left(\boldsymbol{p}_j,y_j\right),y_j\right) \tag{A.94}$$

where $P_j\left(\boldsymbol{p}_j,y_j\right)$ is a price aggregator and $\boldsymbol{p}_j$ is the vector of all prices in market j. This class of demand functions includes a number of homothetic and non-homothetic examples, as discussed in Arkolakis et al. (2019a) and Matsuyama and Ushchev (2017).

We make two assumptions following Arkolakis et al. (2019a). First, we assume that the demand function features a choke price or, in other words, for each $P_j(\boldsymbol{p},y)$ there exists $a\in\mathbb{R}$ such that $q_j\left(x;\boldsymbol{p}_j,y_j\right)=0$ for all $x\geq a$. This way we can abstract from the fixed cost of entry – that is, we assume that $\bar{f}_{ij}=0$ for all i and j. Second, we assume that the demand elasticity $\varepsilon_j\left(p\left(\omega\right);P,y\right)=-\partial\ln q\left(p\left(\omega\right);P,y\right)/\partial\ln p$ is decreasing in $p\left(\omega\right)$. For exposition, we suppress the dependence of the demand function and its elasticity on on P and y.

Production. We assume that the production function is

$$C_{ij}(\omega,q)=w_i\frac{\tau_{ij}(\omega)}{a_i(\omega)}\frac{\bar{\tau}_{ij}}{\bar{a}_i}q.$$

Notice that, relative to the baseline, we abstract from the fixed cost of entry. In this case, the extensive margin of firm exports arises from the chock price in demand. We define the firm-specific cost shifter as

$$c_{ij}\left(\omega\right)\equiv\frac{\tau_{ij}\left(\omega\right)}{a_i\left(\omega\right)}$$

Since the production function is constant returns to scale, the quantity for each firm ω can be defined for each pair of markets separately. Thus, given aggregates P and y, the profit maximization problem of firm ω from i when selling in j is

$$\pi_{ij}\left(\omega\right)=\max_{p(\omega)}\left\{\left(p\left(\omega\right)-w_i\frac{\bar{\tau}_{ij}}{\bar{a}_i}c_{ij}\left(\omega\right)\right)q_{ij}\left(p\left(\omega\right)\right)\right\}$$

The associated first order condition is given by

$$\left(1-w_i\frac{\bar{\tau}_{ij}}{\bar{a}_i}\frac{c_{ij}\left(\omega\right)}{p_{ij}\left(\omega\right)}\right)=-1/\left(\partial\ln q_j\left(p_{ij}\left(\omega\right)\right)/\partial\ln p\right)=1/\varepsilon_j\left(p_{ij}\left(\omega\right)\right).$$

Thus, markups are inversely related to the elasticity of demand:

$$m_{ij}(\omega) \equiv \frac{p_{ij}(\omega)}{w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)} = \frac{\varepsilon_j(p_{ij}(\omega))}{\varepsilon_j(p_{ij}(\omega)) - 1}.$$

Furthermore, our second assumption guarantees that the markup is strictly decreasing on the marginal cost, that the price is strictly increasing on marginal cost, and that quantities and sales are strictly decreasing on the marginal cost (see related arguments in Arkolakis et al. (2019a)). This implies that we can perform a change of variables to express all variables in terms of the marginal cost of production:

$$\pi\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)\right) = \left(\frac{m\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)\right) - 1}{m\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)\right)}\right) R\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)\right). \tag{A.95}$$

Since a higher marginal cost lowers markups and sales, the profit function is strictly decreasing on $w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega)$. Therefore, given every P and y, there exists a unique revenue potential threshold that determines entry into a market:

$$\omega \in \Omega_{ij} \Leftrightarrow w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega) \le c_j^*(P_j, y_j) \quad \text{such that} \quad \pi\left(c_j^*(P_j, y_j); P_j, y_j\right) = 0. \tag{A.96}$$

Conditional on entering, the revenues and profits are

$$R_{ij}(\omega) = R_{ij}\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega); P_j, y_j\right) \quad \text{and} \quad \pi_{ij}(\omega) = \pi\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega); P_j, y_j\right). \tag{A.97}$$

Entry. Let us assume that firms pay a fixed entry cost $\bar{F}_i$ in domestic labor to get a draw of variety characteristics from a distribution:

$$v_i(\omega) = \{a_i(\omega), \tau_{ij}(\omega)\}_j \sim G_i(v) \tag{A.98}$$

In expectation, firms only pay the fixed entry cost if ex-ante profits exceed entry them:

$$\sum_j E\left[\max\{\pi_{ij}(\omega);\ 0\}\right] = w_i \bar{F}_i, \tag{A.99}$$

Market clearing. We follow Dekle et al. (2008) by introducing exogenous international transfers, so that spending is

$$E_i = w_i \bar{L}_i + \bar{T}_i, \quad \sum_i \bar{T}_i = 0. \tag{A.100}$$

Since labor is the only factor of production, labor income in i equals the total revenue of firms from i:

$$y_i = w_i \bar{L}_i = \int_{\omega \in \Omega_{ij}} R_{ij}\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c_{ij}(\omega); P, y\right). \tag{A.101}$$

Equilibrium. Given the distribution in (A.98) , the equilibrium is the vector $\{P_i, y_i, \{\Omega_{ij}\}_j, N_i, E_i, w_i\}_i$ satisfying (A.94), (A.96), (A.99), (A.100), (A.101) for all i.

A.5.2 Extensive and Intensive margin of Firm-level Export

We now turn to the characterization of the semiparametric gravity equations. We consider the distribution of firm-specific shifts of marginal costs implies by G_i:

$$c_{ij}(\omega) \sim H_{ij}(c), \tag{A.102}$$

where H_{ij} has full support in $\mathbb{R}_+$.

Extensive margin of firm-level exports. The share of firms of country i serving market j is $n_{ij} = Pr\left[\omega \in \Omega_{ij}\right] = Pr\left[c_{ij}(\omega) \leq \frac{\bar{a}_i}{\bar{\tau}_{ij} w_i} c_j^*\left(P_j, y_j\right)\right]$.

$$n_{ij} = H_{ij}\left(\frac{\bar{a}_i}{\bar{\tau}_{ij} w_i} c_j^*\left(P_j, y_j\right)\right)$$

We define $\epsilon_{ij}(n) \equiv \left(H_{ij}\right)^{-1}(n)$. Notice that it is now strictly increasing in n. Thus,

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left(\bar{a}_i / \bar{\tau}_{ij}\right) - \ln w_i + \ln c_j^*\left(P_j, y_j\right). \tag{A.103}$$

This is the semiparametric gravity equation for the extensive margin of firm exports.

Intensive margin of firm-level exports. As before, the average revenue of firms from country i in country j is $\bar{x}_{ij} \equiv E\left[R_{ij}\left(\omega\right)|\omega \in \Omega_{ij}\right]$. Thus,

$$\bar{x}_{ij} = \int_0^{\frac{\bar{a}_i}{\bar{\tau}_{ij} w_i} c_j^*(P_j, y_j)} R_{ij}\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} c; P_j, y_j\right) \frac{dH_{ij}(c)}{H_{ij}\left(\frac{\bar{a}_i}{\bar{\tau}_{ij} w_i} c_j^*\left(P_j, y_j\right)\right)}$$

We can then use the transformation, $n = H_{ij}(c)$ such that $dn = dH_{ij}(c)$, $c = \epsilon_{ij}(n)$, and $n_{ij} = H_{ij}\left(\frac{\bar{a}_i}{\bar{\tau}_{ij} w_i} c_j^*\left(P_j, y_j\right)\right)$. Thus,

$$\bar{x}_{ij} = \frac{1}{n_{ij}} \int_0^{n_{ij}} R_{ij}\left(w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} \epsilon_{ij}(n); P_j, y_j\right) dn.$$

Using (A.103),

$$\bar{x}_{ij} = \frac{1}{n_{ij}} \int_0^{n_{ij}} R_{ij}\left(c_j^*\left(P_j, y_j\right) \frac{\epsilon_{ij}(n)}{\epsilon_{ij}(n_{ij})}; P_j, y_j\right) dn. \tag{A.104}$$

In this case, we can derive an expression for average sales as function of $\epsilon_{ij}(n)$. So, although we do not have a gravity equation for average firm exports, this is entirely determined by the function governing the semiparametric gravity equation for the extensive margin of firm exports.

For this extension, we do not derive the model's general equilibrium predictions because it requires specifying the price index aggregator of demand, $P_j\left(\boldsymbol{p}_j, y_j\right)$. Conditional on knowing the price index function, we can follow the same steps as in our baseline model to characterize the price index in terms of the distribution of firm firm-specific marginal cost shifters, $H_{ij}(c)$, and, therefore, in terms of the extensive margin elasticity function, $\epsilon_{ij}(n) \equiv \left(H_{ij}\right)^{-1}(n)$.

B Additional Results: Multiple Country Groups

Our baseline estimates impose that the elasticity functions are identical for all exporter-destination pairs ($G = 1$). In this section, we estimate alternative specifications where we allow the elasticity function to vary across groups of countries.

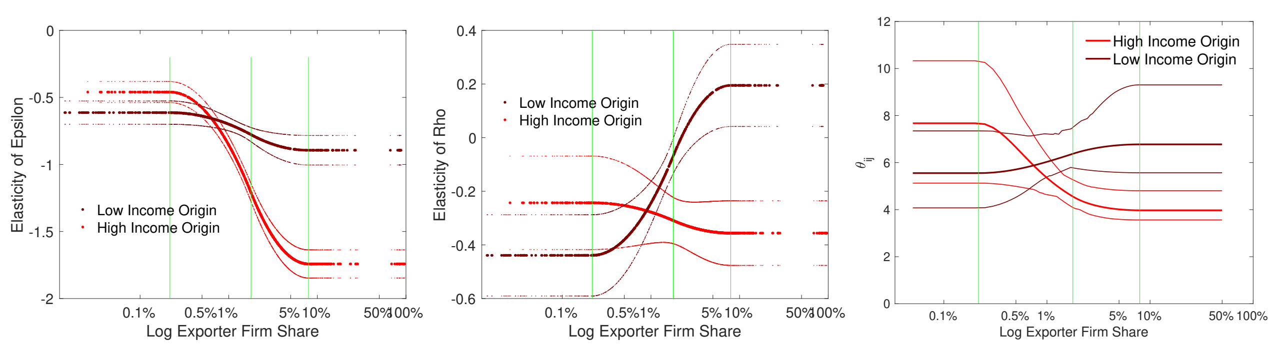
B.1 Heterogeneity with respect to per capita income

We first implement our estimation procedure with country groups defined defined in terms of per capita GDP. This type of heterogeneity in trade elasticity has been explored by Adao et al. (2017). We use a cutoff of \$9,000 of per capita GDP in 2002 to divide our sample into developed and developing nations. Column (6) of Table OA.1 in Appendix B.1 shows the list of developed and developing countries in our sample.

Figure SM.1 reports estimates for two groups defined in terms of development of the origin country. Panel (a) indicates that the extensive margin elasticity varies less with the exporter firm share in developing origin countries. In addition, Panel (b) indicates that, in developing countries, selection forces are weaker for high levels of n_{ij}. This translates into a roughly constant trade elasticity of six for developing countries. For developed countries, our estimates are similar to the baseline results in Figure 3. This reflects the fact that developed origin countries constitute the majority of our sample.

We also implement our estimation procedure for four groups defined in terms of per capita income of both the origin and destination countries. Figure SM.2 shows that the per capita income of the destination country does not have a large impact on the estimates reported in Figure SM.1.

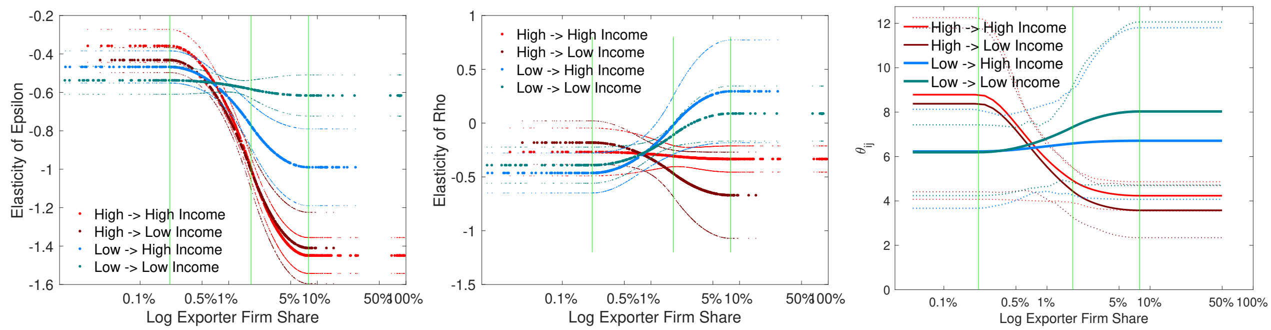
Figure SM.1: Semiparametric gravity estimation – Country groups defined in terms of per capita income of origin country



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a two groups ($G = 2$). Groups defined in terms of origin country per capita income – see column (6) of Table OA.1 of Appendix B.1. Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 1,000 bootstrap draws for $\theta(n)$.

Figure SM.2: Semiparametric gravity estimation – Country groups defined in terms of per capita income of origin and destination countries



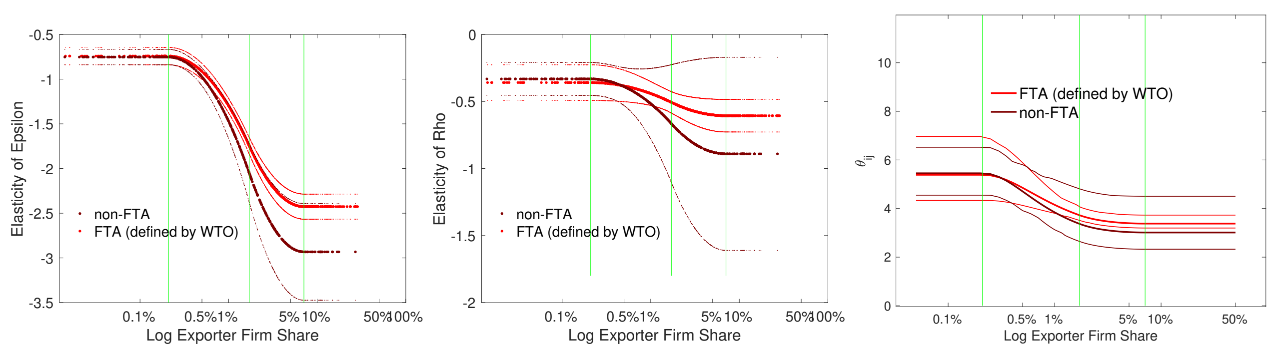
(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33). Estimates obtained with a cubic spline over three intervals ($K = 3$) for a four groups ($G = 4$). Groups defined in terms of per capita income of origin and destination countries – see column (6) of Table OA.1 of Appendix B.1. Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

B.2 Heterogeneity with respect to bilateral attributes - Free Trade Areas

We implement our estimation procedure for two groups of exporter-importer pairs defined as country pairs inside and outside a common Free Trade Areas (FTA) (using the CEPII bilateral gravity dataset). A large body of literature has documented that membership in free trade areas reduces trade costs – see Head and Mayer (2014). We investigate here if it also affects the different elasticity margins of trade flows. Figure SM.3 indicates that there are only small differences in the estimates for countries inside and outside common free trade areas.

Figure SM.3: Semiparametric gravity estimation – Country groups defined in terms of membership in free trade areas



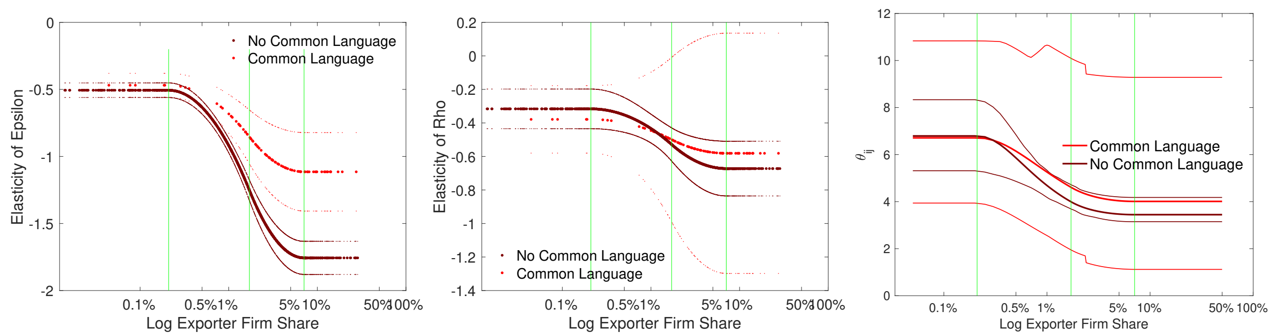
(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,443 origin-destination pairs described in Table OA.1 of Appendix B.1. Estimates obtained with a cubic spline over three intervals ($K = 3$) for a two groups ($G = 2$). Groups defined as country pairs inside and outside a common Free Trade Areas (FTA). Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 1,000 bootstrap draws for $\theta(n)$.

B.3 Heterogeneity with respect to bilateral attributes - Common Languages and Currency

Figures SM.4 and SM.5 investigate whether the elasticity functions differ across country pairs that share a common language or currency. The literature has documented that these two characteristics are associated with higher bilateral trade flows – see Head and Mayer (2014). We again find no evidence that such characteristics affect the elasticity functions.

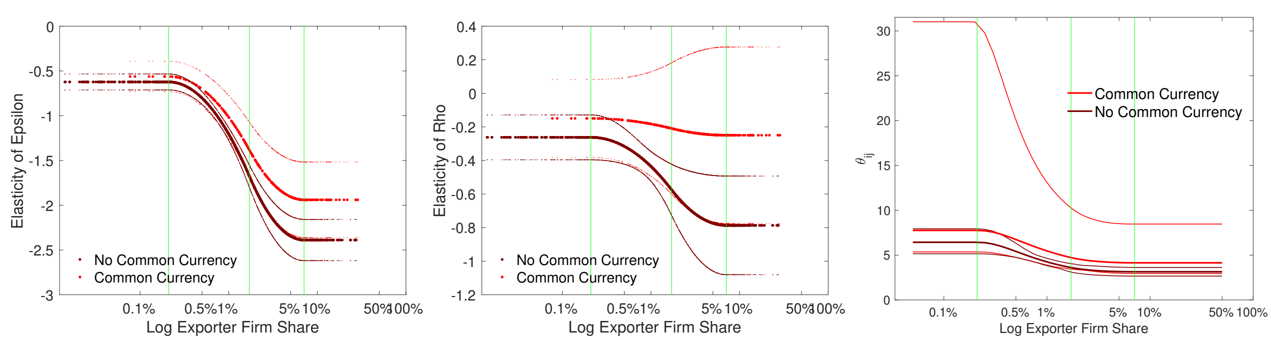
Figure SM.4: Semiparametric gravity estimation – Country groups based on having a common language



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,443 origin-destination pairs described in Table OA.1 of Appendix B.1. Estimates obtained with a cubic spline over three intervals ($K = 3$) for a two groups ($G = 2$). Country groups defined in terms of having at least 9% of a country speak the same language. Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 1,000 bootstrap draws for $\theta(n)$.

Figure SM.5: Semiparametric gravity estimation – Country groups based on having a common currency



(a) Elasticity of $\epsilon(n)$ (b) Elasticity of $\rho(n)$ (c) Implied $\theta(n)$

Note. Estimates obtained with GMM estimator in (33) in the 2012 sample of 1,443 origin-destination pairs described in Table OA.1 of Appendix B.1. Estimates obtained with a cubic spline over three intervals ($K = 3$) for a two groups ($G = 2$). Country groups defined in terms of sharing an official currency. Calibration of $\tilde{\sigma} = 2.9$. Thick lines are the point estimates and thin lines are the 95% confidence intervals computed with robust standard errors for $\epsilon(n)$ and $\rho(n)$ and 100 bootstrap draws for $\theta(n)$.

C Computation Algorithms

C.1 Hat Algebra

We now describe an algorithm to compute the changes in aggregate outcomes that solve the system in Appendix A.2 for an arbitrary trade cost change from $\{\bar{\tau}_{ij}\}_{ij}$ to $\{\bar{\tau}'_{ij}\}_{ij}$.

1. Compute the partition of the shock with length R: $d\ln\bar{\tau}_{ij} = \frac{1}{R}\left(\ln\bar{\tau}'_{ij} - \ln\bar{\tau}_{ij}\right)$. Consider the initial equilibrium with $\theta_{ij}(n^0_{ij})$ and $\{x^0_{ij}, y^0_{ij}, \iota^0_j\}$.

2. For each step r, we consider the initial conditions $\left(\varepsilon_{ij}(n^{r-1}_{ij}), \varrho_{ij}(n^{r-1}_{ij}), \theta_{ij}(n^{r-1}_{ij})\right)$ and $\{x^{r-1}_{ij}, y^{r-1}_{ij}, \iota^{r-1}_j\}$.

   (a) Compute $(d\ln\boldsymbol{\tau}^{w,r}, d\ln\boldsymbol{\tau}^{p,r})$ and $d\ln\boldsymbol{w}^r$ using (OA.38) (for a given numerarie with $d\ln w^r_m = 0$).

   (b) Solve $d\ln\boldsymbol{P}^r$ using (OA.37).

   (c) Solve $d\ln n^r_{ij}$ and $d\ln\bar{x}^r_{ij}$ using (OA.14) and (OA.15).

   (d) Solve for $d\ln N^r_i$ using (OA.24).

   (e) Compute the change in bilateral trade flows: $d\ln X^r_{ij} = d\ln\bar{x}^r_{ij} + d\ln n^r_{ij} + d\ln N^r_i$.

   (f) Compute the initial conditions for the next step: $X^r_{ij} = X^{r-1}_{ij}e^{d\ln X^r_{ij}}$ and $n^r_{ij} = n^{r-1}_{ij}e^{d\ln n^r_{ij}}$.

   (g) Compute $x^r_{ij} = X^r_{ij}/\sum_o X^r_{oj}$, $y^r_{ij} = X^r_{ij}/\sum_d X^r_{id}$, $\iota^r_i = \left(\sum_d X^r_{id}\right)/\left(\sum_o X^r_{oi}\right)$, and $\left(\varepsilon_{ij}(n^r_{ij}), \varrho_{ij}(n^r_{ij}), \theta_{ij}(n^r_{ij})\right)$.

3. Compute changes in aggregate variables as

$$\hat{\boldsymbol{w}}^{linear} = \exp\left(\sum_{r=1}^{R} d\ln\boldsymbol{w}^r\right), \quad \hat{\boldsymbol{P}}^{linear} = \exp\left(\sum_{r=1}^{R} d\ln\boldsymbol{P}^r\right), \quad \hat{\boldsymbol{N}}^{linear} = \exp\left(\sum_{r=1}^{R} d\ln\boldsymbol{N}^r\right).$$

4. Use $\{\hat{w}^{linear}_i, \hat{P}^{linear}_i\}$ as an initial guess is the solution of the hat algebra system. Set the same numerarie as above, $\hat{w}_m \equiv 1$.

   (a) Given guess of $(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}})$, compute

$$\frac{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}{\epsilon_{ij}(n_{ij})} = \left(\hat{\bar{\tau}}_{ij}\right)^{\sigma-1}\left[\left(\frac{\hat{w}_i}{\hat{P}_j}\right)^{\sigma}\frac{\hat{P}_j}{\iota_j\hat{w}_j}\right],$$

$$\hat{\bar{x}}_{ij} = \left(\hat{\bar{\tau}}_{ij}\right)^{1-\sigma}\frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})}\left[\left(\frac{\hat{w}_i}{\hat{P}_j}\right)^{1-\sigma}\iota_j\hat{w}_j\right].$$

   From (OA.13),

$$\hat{N}_i = \left[1 - \sum_j y_{ij}\left(\frac{\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n_{ij})}\,dn}\right) + \sum_j y_{ij}\frac{\hat{n}_{ij}\hat{\bar{x}}_{ij}}{\hat{w}_i}\left(\frac{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n)}\,dn}{\int_0^{n_{ij}\hat{n}_{ij}}\frac{\rho^m_{ij}(n)}{\epsilon_{ij}(n_{ij}\hat{n}_{ij})}\,dn}\right)\right]^{-1}.$$

   (b) Compute the functions:

$$F^P_j\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}}\right) \equiv \hat{P}^{1-\sigma}_j - \sum_i x_{ij}\left(\left(\hat{\bar{\tau}}_{ij}\right)^{1-\sigma}\frac{\rho_{ij}(n_{ij}\hat{n}_{ij})}{\rho_{ij}(n_{ij})}\hat{w}^{1-\sigma}_i\hat{n}_{ij}\hat{N}_i\right)$$

$$F_i^w\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}}\right) \equiv \hat{w}_i - \sum_j y_{ij}\left(\hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij}\right).$$

(c) Find $\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}}\right)$ that minimizes $\left\{|F_j^P\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}}\right)|, |F_i^w\left(\hat{\boldsymbol{w}}, \hat{\boldsymbol{P}}\right)|\right\}$.

C.2 Gains from trade

We now describe an algorithm to compute the gains from trade described in Appendix A.6.

1. Define the uni-dimensional function

$$F_i(\hat{n}_{ii}^A) = \sum_j y_{ij}\left(1 - \frac{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n)}\, dn}{\int_0^{n_{ij}} \frac{\rho_{ij}^m(n)}{\epsilon_{ij}(n_{ij})}\, dn}\right) - \frac{1}{\hat{N}_i^A(\hat{n}_{ii}^A)} \frac{y_{ii}}{x_{ii}}\left(1 - \frac{\int_0^{n_{ii}\hat{n}_{ii}^A} \frac{\rho_{ii}^m(n)}{\epsilon_{ii}(n)}\, dn}{\int_0^{n_{ii}\hat{n}_{ii}^A} \frac{\rho_{ii}^m(n)}{\epsilon_{ii}(n_{ii}\hat{n}_{ii}^A)}\, dn}\right)$$

where

$$\hat{N}_i^A(\hat{n}_{ii}^A) = \frac{\rho_{ii}\left(n_{ii}\right)}{\epsilon_{ii}\left(n_{ii}\right)} \frac{\epsilon_{ii}\left(n_{ii}\hat{n}_{ii}^A\right)}{\rho_{ii}\left(n_{ii}\hat{n}_{ii}^A\right)} \frac{1}{\hat{n}_{ii}^A} \frac{1}{x_{ii}}.$$

2. For each i, we find $\hat{n}_{ii}^A$ such that $F_i(\hat{n}_{ii}^A) = 0$. We then compute the gains from trade using equation (23).