NBER WORKING PAPER SERIES

FROM HETEROGENEOUS FIRMS TO HETEROGENEOUS TRADE ELASTICITIES: THE AGGREGATE IMPLICATIONS OF FIRM EXPORT DECISIONS

Rodrigo Adão
Costas Arkolakis
Sharat Ganapati

Working Paper 28081
http://www.nber.org/papers/w28081

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2020, Revised December 2024

We thank Xiaohong Chen, Tim Christensen, Xiang Ding, Phil Haile, Keith Head, Federico Huneeus, Giovanni Compiani, Chang-Tai Hsieh, Samuel Kortum, Stephen Redding, and Yacan Wang for valuable comments. We thank Takashi Onada, Yongjie Xiong, Mengly Xu, Ryan Thurz, 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

From Heterogeneous Firms to Heterogeneous Trade Elasticities: The Aggregate Implications of Firm Export Decisions
Rodrigo Adão, Costas Arkolakis, and Sharat Ganapati
NBER Working Paper No. 28081
November 2020, Revised December 2024
JEL No. C14, F12, F14, F15

ABSTRACT

We study the consequences of globalization in monopolistic competition models where heterogeneous firms select into export markets. We summarize firm heterogeneity at the extensive and intensive margins of firm exports with two nonparametric elasticity functions that depend only on the share of active firms in a market. Given changes in trade costs, these elasticity functions are sufficient to compute the model's counterfactual predictions, and their shape generates heterogeneity in welfare responses across countries. To estimate these functions, we use the model's semiparametric gravity equations of firm export margins, which yield trade elasticity estimates that vary with the number of exporters in a market and the country's level of development. Compared to constant-elasticity gravity models, our estimates imply gains from trade that are larger in developed countries but smaller in developing countries, with differences arising mainly due to the entry and selection of heterogeneous firms.

Rodrigo Adão
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
rodrigo.adao@chicagobooth.edu

Costas Arkolakis
Department of Economics
Yale University, 87 Trumbull Street
P.O. Box 208268
New Haven, CT 06520-8268
and NBER
costas.arkolakis@yale.edu

Sharat Ganapati
Georgetown University
Edward B. Bunn, S.J. Intercultural Center
37th and O Streets, NW
Washington, DC 20057
and CESifo
and also NBER
sg1390@georgetown.edu

A data appendix is available at http://www.nber.org/data-appendix/w28081

1 Introduction

International trade reflects the endogenous decision of heterogeneous firms to select which countries to export to. Larger, more productive firms are more likely to export and their exporting expansion induces resource reallocation that leads smaller firms to exit (Melitz, 2003; Bernard and Jensen, 2004; Melitz and Redding, 2014). However, the relationship between firms and trade likely varies across countries, as large, high-productivity firms are more prevalent in developed economies (Hsieh and Olken, 2014). How does the decision of heterogeneous firms to export shape the impact of globalization across different countries? Following trade cost changes, do differences in firm characteristics translate into differences in aggregate responses?

The literature, while explicitly acknowledging the role of firm heterogeneity for selection into export and new firm entry, often assigns it a secondary role when quantifying the gains from globalization. This is largely due to strong parametric restrictions on firm heterogeneity. Such restrictions are useful in practice: they easily link available data to counterfactual predictions. But this is not a free lunch: they also come at the cost of restricting how firm heterogeneity shapes responses to changes in trade costs. A canonical example is the assumption of a Pareto distribution for firm productivity, which implies that gains from trade are the same in neoclassical and heterogeneous-firm models and new firm entry is invariant to globalization (Chaney, 2008; Arkolakis, Costinot and Rodríguez-Clare, 2012). More generally, parametric distributional assumptions determine both the aggregate implications of firm heterogeneity and the set of moments used for identification (e.g., Eaton, Kortum and Kramarz, 2011; Melitz and Redding, 2015; Head, Mayer and Thoenig, 2014; Bas, Mayer and Thoenig, 2017; Fernandes, Klenow, Meleshchuk, Pierola and Rodríguez-Clare, 2023).

In this paper, we propose a new methodology to analyze the aggregate consequences of globalization in monopolistic competition models where heterogeneous firms make exporting decisions. Firm heterogeneity affects the extensive and intensive margins of firm exports through two nonparametric elasticity functions, which depend only on the share of active firms in a market. Following changes in trade costs, these elasticity functions summarize how the export decisions of heterogeneous firms determine the model's counterfactual predictions. In particular, their shape generates heterogeneity in welfare responses across countries through creation of new varieties (entry) and decisions to sell to different destinations (selection). To estimate these functions, we use our model's semiparametric gravity equations for the firm export margins, which yield trade elasticity estimates that vary with the number of exporters in a market and the country's level of development. Compared to constant-elasticity gravity models, our estimates imply gains from trade that are larger in developed countries but smaller in developing countries, with differences arising mainly from firm entry and selection.

We consider a multi-country extension of the monopolistic competition model in Melitz (2003) where heterogeneous firms in an origin decide whether to create a new variety and which destinations to sell it to. As in the extensive literature reviewed by Melitz and Redding (2014), we focus on the aggregate implications of these entry and selection decisions. Thus, in our baseline model, we abstract from markup heterogeneity across firms, and maintain the assumption of constant elasticity of substitution (CES) preferences in Melitz (2003). In contrast, we allow for an arbitrary distribution of firm fundamentals; namely, the distribution of destination-specific shifters of productivity, demand, and costs. Our setting generates rich patterns of heterogeneity in firm export decisions, both within and between destinations, such as those documented by Eaton et al. (2004, 2011) and Fernandes et al. (2023).

In this environment, we derive two elasticity functions that summarize all specified sources of firm heterogeneity, and determine firm exports through the extensive and intensive margins. These functions represent the elasticities of the two margins of firm exports to a destination with respect to bilateral trade costs in semiparametric gravity equations, where origin and destination fixed-effects absorb endogenous country-level outcomes. Our first main result is that the elasticity of trade flows to trade costs, which is the sum of these two margins, is not a parameter. Instead, it is a univariate function of the share of the firms from the origin that sell in a given destination (i.e., exporter firm share).[1] Our second main result is that, conditional on these elasticity functions, the economy's general equilibrium does not depend on any other heterogeneous characteristic of firms. These results directly generalize for models featuring multiple sectors, multiple factors, and multi-product firms. We also show how to extend our results to allow for heterogeneous markups with non-CES demand, as in Matsuyama and Ushchev (2017) and Arkolakis et al. (2019a).[2]

We use these theoretical insights to establish that the two elasticity functions summarize the role of firm heterogeneity in the economy's aggregate response to changes in trade costs. That is, given these elasticity functions, the counterfactual predictions of the model do not depend on its micro structure. We further show that the two elasticity functions determine entry and selection of heterogeneous firms. Most importantly, they regulate the sign and

[1] Parametric assumptions restrict the shape of these elasticity functions. For instance, the elasticities of all margins are invariant to the exporter firm share if firm productivity has the Pareto distribution, as in Chaney (2008), and are decreasing with Truncated Pareto, as in Melitz and Redding (2015), or Log-normal, as in Head et al. (2014) and Bas et al. (2017).

[2] In this case, the two elasticity functions determine firm export margins as well as (good and factor) market clearing. With homogeneous fixed cost of exporting, they are also sufficient to characterize the free entry condition. However, when firms are heterogeneous in their fixed costs, the characterization of profit margins across destinations also requires elasticity functions in gravity equations for the percentiles of the firm export distribution (instead of only that for average firm exports).

magnitude of components of welfare responses associated with entry and selection decisions. As a result, we leverage the CES demand in our model to establish that the impact of trade costs on welfare is the sum of neoclassical gains due to technology and terms of trade (see e.g. Dixit and Norman (1980) and Baqaee and Farhi (2019)), and firm gains due to entry and selection (see e.g. Hsieh et al. (2020) and Redding and Weinstein (2024)). Finally, we use our elasticity functions to derive a nonparametric extension of the sufficient statistics in Arkolakis et al. (2012) and Melitz and Redding (2015).

Our theoretical analysis suggests that the shape of the elasticity functions of firm export margins is central for the counterfactual predictions of monopolistic competition models. Parametric assumptions on the distribution of firm fundamentals are not innocuous. While their tractability facilitates estimation and simulations, they constrain how export decisions affect counterfactual predictions. In particular, the shape of the elasticity functions determines the extent to which selection into foreign markets leads to selection out of the domestic market, and whether globalization leads to an increase in the number of firms in a country. Thus, by providing reliable estimates of these elasticity functions, we can assess the aggregate implications of globalization when heterogeneous firms make entry and selection decisions.

Motivated by these theoretical insights, we turn to the estimation of the elasticity functions of firm export margins. In particular, we extend conventional gravity tools (see e.g. Head and Mayer (2014)) to estimate the model's semiparametric gravity equations for the extensive and intensive margins of firm exports. We find that the elasticity functions of both margins vary with the exporter firm share in an origin-destination as well as with the development level of the origin and destination. In developing countries, trade flows are less responsive to trade costs in markets with fewer active firms than in those with more active firms. In contrast, developed countries display the opposite pattern: they have a larger responsiveness to trade costs in markets where only the few, most efficient firms participate.[3] We show that these findings are inconsistent with existing estimates based on parametric restrictions on firm heterogeneity, but they are consistent with the evidence in Hsieh and Klenow (2009) and Hsieh and Olken (2014) that, relative to developed countries, developing countries have a fatter tail of low-productivity, small firms, which operate mainly in the domestic market.

We conclude with a counterfactual analysis that quantifies how firm heterogeneity affects aggregate responses to changes in trade costs. First, we consider a small uniform reduction in trade costs for all countries. Our heterogeneous elasticity estimates imply heterogeneous effects across developed and developing countries. Compared to a constant-elasticity benchmark, our estimates yield larger welfare gains in developed countries but smaller gains in developing

[3] We also show that these patterns are invariant to the sectoral composition of trade flows, and to other country characteristics (like their level of trade integration).

countries, with differences arising mainly from firm entry and selection. This pattern also holds for large shocks, such as a counterfactual move to autarky. Additionally, we show that similar patterns emerge in a counterfactual exercise inspired by the Generalized System of Preferences (GSP), where developed countries reduce barriers on imports from developing countries. In this case, however, the heterogeneous shock interacts with our elasticity estimates, leading to a substantial contribution of firm decisions to terms of trade.

Our paper is related to the extensive theoretical and empirical literature on firms in international trade (for reviews, see Bernard et al. (2007); Redding (2011); Melitz and Redding (2014)). From a theoretical perspective, we build on the insights in Chaney (2008), Arkolakis et al. (2012) and Melitz and Redding (2015) that the value of the trade elasticity summarizes the effects of heterogeneous firms on trade and welfare. Our main contribution is to generalize these insights without parametric restrictions on the distribution of firm fundamentals: we show that the elasticity functions of the intensive and extensive margins of firm exports summarize the general equilibrium implications of firm entry and selection.

We use these elasticity functions to characterize the properties of monopolistic competition models with CES demand, as well as the welfare gains from globalization. Our work extends the decomposition proposed by Atkeson and Burstein (2010) to an environment with multiple asymmetric countries, domestic selection, and arbitrary heterogeneity. For any given country, welfare gains from firm entry and selection are non zero to a first-order when countries are asymmetric, but they are indeed second-order when countries are symmetric.[4] We also derive a nonparametric extension of the sufficient statistics in Arkolakis et al. (2012) and Melitz and Redding (2015). It indicates that what matters for the gains from trade is not the (constant) "trade elasticity," but instead the (variable) "domestic elasticity" and the endogenous entry and selection decisions of domestic firms. Intuitively, as in Costinot and Rodríguez-Clare (2018), the gains from trade of a country correspond to the area below its import demand curve, which requires accounting for any variation in the domestic trade elasticity along the path from trade to autarky.

From an empirical perspective, our semiparametric approach builds a bridge between the literature that estimates constant-elasticity gravity trade models, reviewed by Head and Mayer (2014), and the one that makes use of granular data on firm outcomes to estimate the distribution of firm fundamentals driving their exporting decisions (e.g. Bas et al. (2017) and Egger et al. (2023) for productivity and Eaton et al. (2011) and Fernandes et al. (2023) also for demand and trade costs). We show how to extend existing tools to estimate variable

[4]With the appropriate weights, global welfare changes are entirely given by technological gains, as in Atkeson and Burstein (2010). We trace this result to the efficiency of the equilibrium in monopolistic competition models with CES demand, due to constant markups and efficient entry (see also discussion in Dhingra and Morrow (2012) for closed economies, and Egger and Huang (2023) for open economies).

trade elasticities in semiparametic gravity equations, while accounting for export decisions of heterogeneous firms. Our approach complements work that estimates variable elasticities using parametric models– e.g., Novy (2013), Fajgelbaum and Khandelwal (2016), Lind and Ramondo (2018), and Bas et al. (2017). We provide evidence that trade elasticities vary with the number of exporters and the country's development level. Finally, we contribute to a literature that measures gains from varieties, as in Broda and Weinstein (2006) and Feenstra and Weinstein (2010), by leveraging our variable elasticity estimates to measure the welfare implications of firm entry and selection.

Our empirical approach builds upon recent advancements on the nonparametric identification of models with self-selection of heterogeneous agents (Berry and Haile, 2014; Adão, 2015). We do so to address the well-known challenge that observed outcomes among active firms in a market (i.e. cross-sectional moments) are insufficient to nonparametrically identify the distribution of fundamentals for firms that are not active in that market (Heckman and Honore, 1990).[5] Instead, we exploit *cross-market* variation in firm export margins induced by trade costs in order to nonparametrically identify the elasticity functions that summarize the role of firm heterogeneity in general equilibrium.

Our work is closely related to recent papers conducting nonparametric counterfactual analysis in international trade models (Adao et al., 2017; Bartelme et al., 2019).[6] These flexible approaches require knowledge of multivariate functions whose nonparametric estimation is challenging in finite samples – for example, Adao et al. (2017) must estimate each country's demand function for all factors in the world economy. Compared to these papers, we consider a different class of models that feature monopolistic competition. Our methodology has the advantage of only requiring the estimation of univariate elasticity functions.

Finally, we study the aggregate implications of selection of heterogeneous firms into exporting markets, as in Melitz (2003). While we extend our nonparametric approach to incorporate firm heterogeneity in markups, sectors, inputs, and products, we abstract from other potentially important dimensions of firm heterogeneity, such as skill intensity (Burstein and Vogel, 2017), sourcing locations (Blaum et al., 2015; Antras et al., 2017), information sets (Dickstein and Morales, 2018), product quality (Kugler and Verhoogen, 2008), and innovation (Atkeson and Burstein, 2010; Bustos, 2011).

Our paper is organized as follows. Section 2 characterizes the two elasticity functions that

[5] In this sense, our estimation approach differs from those that measure firm heterogeneity by looking at the observed distribution of size and export decisions of active firms, such as Eaton et al. (2011), or other cross-sectional moments such as pass-through by firm size, as in Baqaee et al. (2024).

[6] Our paper is also complementary to the literature offering sufficient statistics in neoclassical trade models (Baqaee and Farhi, 2019) and gravity trade models (Allen, Arkolakis and Takahashi, 2014; Kleinman, Liu and Redding, 2020). We note that this alternative approach relies on parametric assumptions for the implementation of empirical and counterfactual analyses.

summarize the aggregate implications of the export decisions of heterogeneous firms. In Section 3, we show how to use these elasticity functions for computing counterfactual predictions and welfare responses to changes in trade cost. Section 4 outlines the methodology to estimate the elasticity functions of firm export margins using the model's semiparametric gravity equations. We report estimates of the elasticity functions in Section 5, and counterfactual exercises in Section 6. Section 7 concludes.

2 From Heterogeneous Firms to Heterogeneous Trade Elasticities

We consider an economy in which monopolistic competitive firms of each origin country exhibit heterogeneity in destination-specific shifters of productivity, demand, and (variable and fixed) trade costs. This section studies how the endogenous exporting decisions of these heterogeneous firms shapes aggregate variables in general equilibrium. We show that the aggregate impact of such decisions is summarized by two elasticity functions that govern the extensive and intensive margins of firm exports.

2.1 Environment

We start by describing consumers and then describe the firm's problem and decisions.

Preferences. Each country j has a representative household that inelastically supplies $\bar{L}_j$ labor units, and has Constant Elasticity of Substitution (CES) preferences over varieties ω. Demand is subject to a bilateral taste shifter $\bar{b}_{ij}$ that is common to all varieties from i sold in j, and an idiosyncratic shifter $b_{ij}(\omega)$ that is specific to variety ω. The quantity that j demands of variety ω from origin i is

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

where σ is the elasticity of substitution, E_j is j's total spending, $p_{ij}(\omega)$ is the price of variety ω of i sold in j, and P_j is j's CES price index implicitly determined by j's budget constraint,

$$\sum_i \int_{\omega \in \Omega_{ij}} p_{ij}(\omega) q_{ij}(\omega) d\omega = E_j, \tag{2}$$

with Ω_{ij} the set of varieties of origin i available in j. This environment allows for variety-specific demand shifters. As shown by Eaton et al. (2011), heterogeneous taste shifters help to rationalize the heterogeneous decisions of firms to export to different destinations. Recently, Redding and Weinstein (2020, 2024) pointed out that heterogeneity in variety-specific demand shifters also play an important role in determining variation in price indices and trade flows across countries and years.

Technology. Each variety is produced by a single firm, so we refer to a variety as a firm. Production is subject to variable and fixed labor costs that are heterogeneous across firms. The cost of firm ω from i of selling q units in destination j is

$$C_{ij}(\omega, q) = w_i \frac{\bar{\tau}_{ij}}{\bar{a}_i} \frac{\tau_{ij}(\omega)}{a_i(\omega)} q + w_i \bar{f}_{ij} f_{ij}(\omega), \tag{3}$$

where w_i is the wage in origin i. The variable cost of selling q units in j includes both firm-specific iceberg shipping costs, $\bar{\tau}_{ij}\tau_{ij}(\omega)$, and productivity, $\bar{a}_i a_i(\omega)$. The second term, $w_i \bar{f}_{ij} f_{ij}(\omega)$, is the fixed labor cost necessary for firm ω from i to access consumers in j. We follow Eaton et al. (2011) –and depart from Melitz (2003)– to allow firms to be different not only in their productivity, but also in their fixed costs of exporting. We further introduce heterogeneity in the variable cost of serving different destinations. This allows the model to flexibly replicate various patterns of firm-level exports across destinations.

Entry and revenue potentials. We now define the two variables that determine firm export decisions. These variables summarize the sources of firm heterogeneity into the extensive and intensive margins that are sufficient to characterize equilibrium outcomes.

Under monopolistic competition, the firm's profit maximization problem implies that its optimal price is $p_{ij}(\omega) = \frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij}}{\bar{a}_i} \frac{\tau_{ij}(\omega)}{a_i(\omega)} w_i$ with an associated revenue of

$$R_{ij}(\omega) = (w_i^{1-\sigma} P_j^{\sigma-1} E_j) \bar{r}_{ij} r_{ij}(\omega) \tag{4}$$

where

$$r_{ij}(\omega) \equiv b_{ij}(\omega) \left(\frac{\tau_{ij}(\omega)}{a_i(\omega)} \right)^{1-\sigma} \quad \text{and} \quad \bar{r}_{ij} \equiv \bar{b}_{ij} \left(\frac{\sigma}{\sigma - 1} \frac{\bar{\tau}_{ij}}{\bar{a}_i} \right)^{1-\sigma}. \tag{5}$$

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

Firm ω from i sells in j if its variable profit exceeds its fixed cost. Given the profits implied by CES demand, this is equivalent to

$$\pi_{ij}(\omega) = \frac{1}{\sigma} R_{ij}(\omega) - w_i \bar{f}_{ij} f_{ij}(\omega) \geq 0. \tag{6}$$

Based on this equation, we characterize the set of firms from i selling in j, Ω_{ij}:

$$\Omega_{ij} = \{\omega : e_{ij}(\omega) \geq e^*_{ij}\} \tag{7}$$

where

$$e_{ij}(\omega) \equiv \frac{r_{ij}(\omega)}{f_{ij}(\omega)} \quad \text{and} \quad e^*_{ij} \equiv \frac{\sigma \bar{f}_{ij}}{\bar{r}_{ij}} \left[\left(\frac{w_i}{P_j} \right)^{\sigma} \frac{P_j}{E_j} \right]. \tag{8}$$

We refer to $e_{ij}(\omega)$, the ratio between the firm's revenue potential and its fixed cost, as the *entry potential* of firm ω from i in j, and to e^*_{ij} as the entry cutoff of firms from i in j. Among firms with identical revenue potential, heterogeneity in the fixed entry cost generates heterogeneity in entry potentials and, therefore, in decisions to enter different destinations.

2.2 The Extensive and Intensive Margins of Bilateral Trade Flows

Our focus now turns to define the two elasticity functions that control the extensive and intensive margins of firm exports and thus bilateral trade flows. To do so, we first define the share of firms from i selling in j and their average sales,

$$n_{ij} \equiv Pr\left[\omega \in \Omega_{ij}\right] \quad \text{and} \quad \bar{x}_{ij} \equiv \mathbb{E}\left[R_{ij}(\omega) | \omega \in \Omega_{ij}\right]. \tag{9}$$

We refer to n_{ij} and $\bar{x}_{ij}$ as the exporter firm share and the average firm exports, respectively.

We consider the distribution of $(r_{ij}(\omega), e_{ij}(\omega))$ for each origin i and destination j generated by the underlying joint distribution of firm fundamentals, $\{a_i(\omega), b_{ij}(\omega), \tau_{ij}(\omega), f_{ij}(\omega)\}$. Without loss of generality, we rely on the following decomposition:

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

Our specification allows for any pattern of heterogeneity and correlation in revenue and entry potentials, $(r_{ij}(\omega), e_{ij}(\omega))$, both within and between origin-destination pairs. Accordingly, it departs from the standard in the literature of explicitly imposing functional form assumptions on the distribution of firm fundamentals. Our general formulation thus encompasses several distributional assumptions in the literature. For instance, in Melitz (2003), the only source of firm heterogeneity is productivity such that $r_{ij}(\omega) = e_{ij}(\omega) = (a_i(\omega))^{\sigma-1}$. In this special case, the distribution of $a_i(\omega)$ 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). The single source of firm heterogeneity implies a strict hierarchy of entry across destinations and a perfect cross-firm correlation between the intensive and extensive margins of exports. To relax these implications, multiple papers incorporate additional sources of heterogeneity across firms. For example, Eaton et al. (2011) impose that the distribution of $a_i(\omega)$ is Pareto and of $(b_{ij}(\omega), f_{ij}(\omega))$ is log-normal, while Fernandes et al. (2023) assume a multivariate log-normal distribution of bilateral cost shifters.

We impose the following regularity restriction on the distribution of firm fundamentals.

Assumption 1. *$H^e_{ij}(e)$ 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 changes in trade costs induce a positive mass of firms to switch entry decisions in any given origin-destination pair. This is central for the invertibility argument used below.[7]

Extensive and intensive margins of firm exports. Given the definition in (9) under Assumption 1, the entry decision in (7) implies that

$$H^e_{ij}\left(e^*_{ij}\right) = \Pr\left(e_{ij}\left(\omega\right) < e^*_{ij}\right) = 1 - n_{ij}.$$

We define the extensive margin elasticity function as $\epsilon_{ij}(n) \equiv (H^e_{ij})^{-1}(1-n)$ such that $\epsilon_{ij}(n)$ is strictly decreasing, $\epsilon_{ij}(1) = 0$, and $\lim_{n\to 0} \epsilon_{ij}(n) = \infty$. We thus obtain the following equilibrium relationship between the exporter firm share, bilateral exogenous variables, and country-level endogenous variables:

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

The elasticity of the exporter firm share with respect to bilateral revenue shifters –holding constant other endogenous variables–,

$$\frac{\partial \ln n_{ij}}{\partial \ln \bar{r}_{ij}} = -\frac{1}{\varepsilon_{ij}(n_{ij})}, \tag{12}$$

is determined by the extensive margin elasticity, $\varepsilon_{ij}(n_{ij}) \equiv \frac{\partial \ln \epsilon_{ij}(n)}{\partial \ln n}|_{n=n_{ij}} < 0$. Interestingly, the extensive margin elasticity is not a parameter, but instead a univariate function of the exporter firm share n_{ij}. Depending on the underlying distribution of entry potentials, the extensive margin elasticity can vary not only with the number of exporter firms in any given origin-destination, but also with the origin or destination countries.

We can now use the definition in (9), the revenue expression in (4) and the entry decision

[7] The assumption of no upper bound on the support of e simplifies our derivations, but it is not essential and we relax it in Section 2.4.

in (7) to express average firm exports as

$$\bar{x}_{ij} = (w_i^{1-\sigma} P_j^{\sigma-1} E_j)\bar{r}_{ij} \int_{e_{ij}^*}^{\infty} \mathbb{E}[r|e] \frac{dH_{ij}^e(e)}{1 - H_{ij}^e(e_{ij}^*)},$$

where the integral denotes the mean revenue potential of firms from i selling in j. We define the intensive margin elasticity function as $\rho_{ij}(n) \equiv \frac{1}{n}\int_0^n \mathbb{E}[r|e = \epsilon_{ij}(n')]\ dn'$, with $\mathbb{E}[r|e = \epsilon_{ij}(n)]$ the mean revenue potential of firms in quantile n of the entry potential distribution of ij. Using the change of variables $n = 1 - H_{ij}^e(e)$, we obtain the following equilibrium expression for average firm exports:

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln(\bar{r}_{ij}) + \ln\left(w_i^{1-\sigma}\right) + \ln\left(P_j^{\sigma-1} E_j\right). \tag{13}$$

This expression relates the composition-adjusted average firm exports, $\bar{x}_{ij}/\rho_{ij}(n_{ij})$, to a combination of exogenous bilateral shifters and endogenous outcomes in the origin and destination countries. In turn, the sensitivity of average firm exports to changes in bilateral revenue shifters –holding constant other endogenous variables–,

$$\frac{\partial \ln \bar{x}_{ij}}{\partial \ln \bar{r}_{ij}} = 1 - \frac{\varrho_{ij}(n_{ij})}{\varepsilon_{ij}(n_{ij})}, \tag{14}$$

is determined by the intensive margin elasticity, $\varrho_{ij}(n_{ij}) \equiv \frac{\partial \ln \rho_{ij}(n)}{\partial \ln n}|_{n=n_{ij}} > -1$. We can rewrite this elasticity as $\varrho_{ij}(n_{ij}) = \mathbb{E}[r|e = \epsilon_{ij}(n_{ij})]/\rho_{ij}(n_{ij}) - 1$, which implies that $\varrho_{ij}(n_{ij})$ –and thus how average firm exports respond to trade shocks– depends on the ratio between the mean revenue potential of marginal exporters, $\mathbb{E}[r|e = \epsilon_{ij}(n_{ij})]$, and the mean revenue potential of infra-marginal exporters, $\rho_{ij}(n_{ij})$. Thus, the correlation between revenue and entry potentials determines the sign and magnitude of $\varrho_{ij}(n_{ij})$. In contrast, $\varrho_{ij}(n_{ij}) < 0$ in the Melitz model with a single source of firm heterogeneity ($r_{ij}(\omega) = e_{ij}(\omega)$) as marginal exporters are always worse than infra-marginal exporters, $\frac{\partial \rho_{ij}(n)}{\partial n} = \frac{1}{n^2}\int_0^n (\epsilon_{ij}(n) - \epsilon_{ij}(n'))\, dn' < 0$.

Bilateral trade flows. Next, we construct bilateral trade flows between countries as $X_{ij} \equiv N_i n_{ij} \bar{x}_{ij}$, with N_i the mass of firms in origin i. The elasticity of trade flows with respect to bilateral revenue shifters, holding constant other endogenous variables, is

$$\theta_{ij}(n_{ij}) \equiv \frac{\partial \ln X_{ij}}{\partial \ln \bar{r}_{ij}} = \frac{\partial \ln n_{ij}}{\partial \ln \bar{r}_{ij}} + \frac{\partial \ln \bar{x}_{ij}}{\partial \ln \bar{r}_{ij}} = 1 - \frac{1 + \varrho_{ij}(n_{ij})}{\varepsilon_{ij}(n_{ij})}, \tag{15}$$

with $\theta_{ij}(n) > 1$ (as $\varepsilon_{ij}(n) < 0$ and $\varrho_{ij}(n) > -1$). This elasticity, in contrast to standard quantitative trade models, can vary across origins and destinations as well as with the exporter firm share for a given origin-destination. As discussed above, this is a consequence of how the distribution of firm fundamentals shapes the extensive and intensive margins of firm exports. The following proposition summarizes the results above.

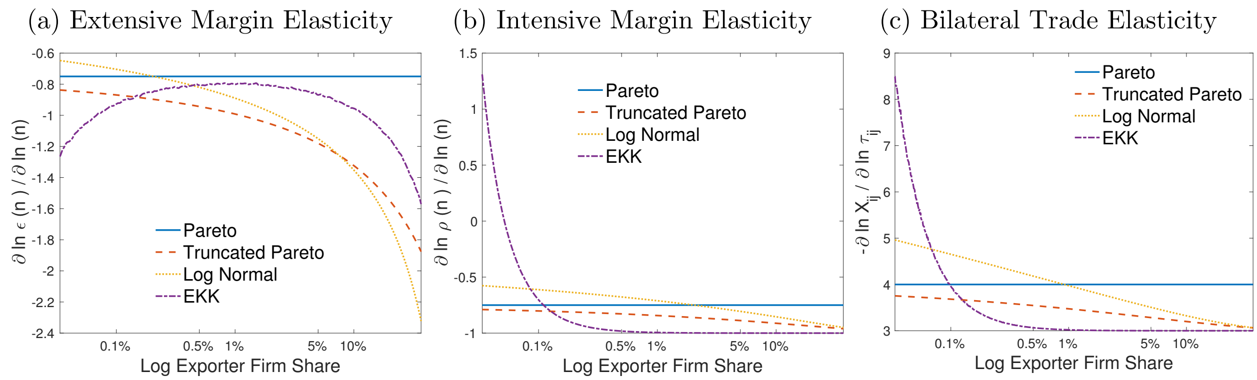
Proposition 1. *Consider the monopolistic competition model with CES demand of Section 2.1 under Assumption 1. Then:*
a. For each origin-destination, the exporter firm share, n_{ij}, and the average firm exports, $\bar{x}_{ij}$, are given by equations (11) and (13), which are separable on country-level endogenous variables, exogenous bilateral shifters, and two elasticity functions, $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$.
b. Given country-level endogenous variables, these elasticity functions summarize the role of firm heterogeneity in the elasticity of bilateral trade flows to bilateral revenue shifters.

The proposition has a direct analogy to the inversion argument used to identify demand systems in Berry (1994) and Berry and Haile (2014), self-selection models in Adão (2015), and perfectly competitive trade models in Adao et al. (2017). Here, we leverage the structure of the monopolistic competition model to invert the equilibrium equations for the exporter firm share n_{ij} and the average firm exports $\bar{x}_{ij}$. This inversion yields log-linear expressions which depend on endogenous country-level variables and exogenous bilateral shifters. These gravity-like expressions will be central for our strategy to estimate $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Note that the relevant elasticity functions only depend on the exporter firm share, n_{ij}. This feature significantly lowers the data requirements for estimation. In the settings cited above, inversion produces elasticity functions that depend on vectors with dimensions that match the number of choices (such as markets or products).[8]

Any parametric assumption on the distribution of firm fundamentals constrains the shape of the elasticity functions of firm export margins. To illustrate this point, Figure 1 plots the elasticity functions implied by the parametric assumptions in the literature discussed above. The blue solid lines indicate that a Pareto distribution of firm productivity yields constant elasticities of all margins. The other parameterizations yield a declining elasticity of $\epsilon_{ij}(n)$ for most of the support, which, by equation (12), implies that the exporter firm share tends to be more sensitive to trade costs when n_{ij} is low. Similarly, all other parameterizations yield a declining elasticity of $\rho_{ij}(n)$, indicating that the ratio between the average revenue potential of new entrants and incumbents is higher when n_{ij} is small. The third panel combines these

[8]For example, in Adao et al. (2017), the nonparametric gravity system depends on a destination-specific function whose dimension is equal to the number of factors in the world economy. We note that the univariate elasticity function emerges from the separability of export decisions across markets in our model. Such a separability does not hold if the firm's profitability in a destination depends on its decision to operate in other destinations, as in Tintelnot (2017) and Morales et al. (2019). In these cases, we have to invert the joint decision to export to all destinations, which increases the dimensionality of the elasticity functions.

Figure 1: Distributional Assumptions and the Elasticity of Firm Export Margins



Note. Panels (a) and (b) report respectively the elasticities of $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$, and panel (c) reports the elasticity of bilateral trade flows to bilateral trade costs. We report the elasticity functions obtained when the productivity distribution is Pareto with shape parameter of four (Chaney, 2008), truncated Pareto with cutoff parameter of $H = 2.85$ (Melitz and Redding, 2015), or log-normal (Head et al., 2014). The specification from EKK uses the baseline estimates reported in Eaton et al. (2011).

two margins to show that, although the elasticity of bilateral trade is decreasing on n_{ij} for all non-Pareto cases, the slope crucially depends on parametric choices.

2.3 Sufficient Statistics of Firm Heterogeneity and General Equilibrium

We now show that $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$ summarize how aggregate variables depend on the export decisions of heterogeneous firms. This requires specifying conditions for free entry, budget balance, and labor market clearing.

Firms incur a fixed labor cost, $\bar{F}_i$, to draw their fundamentals. With free entry, the equilibrium mass of firms in country i, N_i, expects to make zero profit:

$$\sum_j \mathbb{E}[\max\{\pi_{ij}(\omega), 0\}] = \sum_j n_{ij}(\bar{x}_{ij} - \bar{c}_{ij}) = w_i \bar{F}_i, \tag{16}$$

where $\bar{c}_{ij} \equiv \mathbb{E}[C_{ij}(\omega)|\omega \in \Omega_{ij}]$ is the sum of the mean variable and fixed costs of firms from i selling in j. With CES demand, $\bar{c}_{ij}$ is given by

$$\bar{c}_{ij} = (1 - 1/\sigma)\mathbb{E}[R_{ij}(\omega)|\omega \in \Omega_{ij}] + w_i \bar{f}_{ij} \mathbb{E}[f_{ij}(\omega)|\omega \in \Omega_{ij}], \tag{17}$$

and the free entry condition in Equation (16) can be written in terms of the elasticity functions,

$\rho_{ij}(n)$ and $\epsilon_{ij}(n)$:[9]

$$\frac{1}{\sigma}\sum_j n_{ij}\bar{x}_{ij} = w_i\bar{F}_i + w_i\sum_j \bar{f}_{ij}\int_0^{n_{ij}}\frac{\rho_{ij}(n)}{\epsilon_{ij}(n)}(1+\varrho_{ij}(n))dn. \tag{18}$$

As we argued above, $\bar{x}_{ij}$ and n_{ij} can also be written in terms of these elasticites.

To derive the budget constraint, we follow Dekle et al. (2008) and allow for exogenous international transfers $\{\bar{T}_i\}$ with $\sum_i \bar{T}_i = 0$. Thus, total spending equals labor income and transfers in each country j, so that the budget constraint in (2) is equivalent to

$$\sum_i N_i n_{ij}\bar{x}_{ij} = w_j\bar{L}_j + \bar{T}_j = E_j. \tag{19}$$

In addition, since labor is the only factor of production, labor income in i equals the total revenue of firms from i:

$$\sum_j N_i n_{ij}\bar{x}_{ij} = w_i\bar{L}_i. \tag{20}$$

Since (19)-(20) only depend $\bar{x}_{ij}$ and n_{ij}, they can also be written as a function of the elasticity functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$. We can then state the following proposition.

Proposition 2. *Consider the monopolistic competition model with CES demand of Section 2.1 under Assumption 1. Assume knowledge of the exogenous fundamentals $\{\bar{r}_{ij}, \bar{f}_{ij}, \bar{L}_i, \bar{T}_i, \bar{F}_i\}$, the elasticity of substitution σ, and the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Then:*
a. The equilibrium vector $\{P_i, N_i, w_i\}$ solves the system of equations (18)-(20) with n_{ij} and $\bar{x}_{ij}$ given by (11) and (13).
b. The equilibrium is Pareto efficient.

The main implication of Proposition 2 is that the distribution of firm fundamentals affects the economy's equilibrium only insofar it determines the shape of the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. In the next section, we further show that these elasticity functions yield sufficient statistics for the welfare consequences of changes in trade costs.

We prove the second part of the proposition in Appendix A.1.1. It generalizes the equilibrium efficiency results of Dhingra and Morrow (2012) and Zhelobodko et al. (2011) to multiple countries and multiple sources of firm heterogeneity. While intuitive, given the prior literature, our result is not trivial. Whereas due to constant markups, relative quantities are

[9]To see this, note that $\mathbb{E}[f_{ij}(\omega)|\omega \in \Omega_{ij}] = \mathbb{E}[r_{ij}(\omega)/e_{ij}(\omega)|e_{ij}(\omega) > e^*_{ij}] = (1/n_{ij})\int_{e^*_{ij}}^{\infty}(1/e)\mathbb{E}[r|e]dH^e_{ij}(e) = (1/n_{ij})\int_0^{n_{ij}}\mathbb{E}[r|e=\epsilon_{ij}(n)]/\epsilon_{ij}(n)dn$, where the last equality relies on the same change of variable used to derive (13) (i.e., $n = 1 - H^e_{ij}(e)$). From the definition of $\rho_{ij}(n)$, $\rho_{ij}(n)(1+\varrho_{ij}(n)) = \mathbb{E}[r|e=\epsilon_{ij}(n)]$ and thus $n_{ij}w_i\bar{f}_{ij}\mathbb{E}[f_{ij}(\omega)|\omega \in \Omega_{ij}] = w_i\bar{f}_{ij}\int_0^{n_{ij}}\rho_{ij}(n)(1+\varrho_{ij}(n))/\epsilon_{ij}(n)dn$.

efficient under CES preferences in an economy with heterogeneous firms, endogenous entry and selection decisions of firms could be potentially distorted due to cross-market variation in the firm-level distribution of profit margins. Nevertheless, we find that these decisions are also efficient since CES demand implies that profit shares for each firm are constant and invariant of trade costs.

2.4 Extensions

We next discuss extensions of Propositions 1 and 2.a. All derivations are in Appendix A.2.

Non-CES demand. Our first extension considers non-CES demand specifications that allow for variable markups. We maintain the assumption that the demand function is known, but we now specify a general single-aggregator demand, be it homothetic as in Matsuyama and Ushchev (2017), or non-homothetic as in Arkolakis et al. (2019a). We assume that

$$q_{ij}(\omega) = \frac{1}{b_{ij}(\omega)} q_j \left(\frac{1}{b_{ij}(\omega)} \frac{p_{ij}(\omega)}{D_j} \right), \tag{21}$$

where D_j is a demand aggregator that is implicitly defined by the budget constraint in (2). CES demand is the special case in which $q_j(x) = x^{-\sigma}$ and $(D_j)^\sigma = P_j^{\sigma-1} E_j$.

First, note that Proposition 1 continues to hold. The same arguments of Section 2.2 imply that the extensive and intensive margins of firm exports are given by equations similar to (11) and (13).

Second, in the absence of fixed cost heterogeneity, Proposition 2.a also holds. However, when firms differ in their fixed cost of exporting, we need to modify Proposition 2.a. While the elasticity functions of firm export margins still determine budget balance and market clearing, they are not sufficient to characterize profit margins in the free entry condition. To see why, recall that we leveraged two properties of CES demand to express the mean cost of exporters, $\bar{c}_{ij}$, in terms of the intensive and extensive margins. The constant markup allowed us to obtain the mean variable cost of exporters from their average sales, $\bar{x}_{ij}$. In addition, the log-additivity of CES demand allowed us to express entry potential as a multiplicative function of revenue potential, and thus compute the expected fixed cost of exporters directly from $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ (see footnote 9). For more general demand functions, in order to compute the mean cost of exporters in the free entry condition, one needs to know additional elasticity functions that govern how bilateral revenue shifters affect percentiles of the distribution of firm exports; each elasticity is a univariate function of the exporter firm share.

Multiple sectors, multiple factors, input-output links, and import tariffs. Our second extension includes features common to quantitative trade models such as multiple factors of production, input-output links between multiple sectors, and import tariffs. Specifically, we extend the multi-sector, multi-factor gravity model of Costinot and Rodriguez-Clare (2013) to allow firms in each sector to be heterogeneous with respect to shifters of productivity, preferences, and variable and fixed trade costs. We restrict all firms in a sector to have the same nested CES production technology that uses multiple factors and multiple sectoral composite goods.[10] In this setting, we derive sector-specific analogs of (11) and (13) that determine aggregate variables in general equilibrium when combined with knowledge of the components of the production function that are common to all firms in each sector.

Allowing for zero trade flows. Next, we extend our model to allow for zero bilateral trade flows, as in Helpman et al. (2008). To do so, we consider a weaker version of Assumption 1 in which the support of the entry potential distribution is bounded: $H^e_{ij}(e)$ has full support over $[0, \bar{e}_{ij}]$. This does not affect the intensive margin equation in (13), but it introduces a censoring structure into the extensive margin equation in (11).

Multi-product firms. We finally extend our model to allow heterogeneous firms to produce multiple products, as in Bernard et al. (2011). We assume that firms face a convex labor cost of increasing the number of varieties supplied in each destination (see e.g. Arkolakis et al. (2019b)). In this setting, the expressions for the extensive and intensive margins of firm exports are still given by (11) and (13), but Proposition 2.a also requires knowledge of the cost function of introducing new varieties by each firm. We show that this function governs how bilateral revenue shifters affect the average number of products per firm in each destination (through a gravity-like expression analogous to (13)).

3 Nonparametric Counterfactual Analysis: The Aggregate Implications of Firm Export Decisions

This section establishes that the elasticity functions of firm export margins, $\epsilon_{ij}(n)$, $\rho_{ij}(n)$ and $\theta_{ij}(n)$, summarize how the export decisions of heterogeneous firms affect aggregate responses

[10] To simplify exposition, our derivations rely on nested CES preferences and technology. Note however that it is straightforward to extend our results to a more general structure of separable preferences and technology over sectoral composite goods while maintaining the assumption of CES preferences across varieties within each sector. We can use the alternative environment of Appendix Section A.2.1 to further relax the assumption of CES preferences within each sector.

to changes in trade costs. We also provide expressions for welfare changes in terms of these elasticities to show how responses depend on the adjustment margins in our model.

3.1 Counterfactual Responses to Changes in Bilateral Trade Costs

We consider how the economy responds to counterfactual changes in revenue shifters $\{\bar{r}_{ij}\}$. For any variable y, we use y^0 to denote its value at the initial equilibrium, and $\hat{y} \equiv y'/y^0$ and $d\log y$ to denote respectively its ratio and first-order log-change between the initial and counterfactual equilibria. Appendix A.1.2 establishes the requirements for computing the counterfactual responses of aggregate outcomes.

Proposition 3. *Consider a counterfactual change in bilateral revenue shifters $\{\bar{r}_{ij}\}$ in the monopolistic competition model with CES demand of Section 2.1 under Assumption 1. Assume knowledge of the elasticity of substitution σ, and the bilateral trade matrix at the initial equilibrium $\{X_{ij}^0\}$. Then, we can compute counterfactual responses in aggregate outcomes $\{X_{ij}, P_i, N_i, w_i\}$ with knowledge of:*
a. for small shocks, the bilateral trade elasticity matrix at the initial equilibrium $\{\theta_{ij}(n_{ij}^0)\}$;
b. for large shocks, the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$, and the exporter firm share matrix at the initial equilibrium $\{n_{ij}^0\}$.

The first part of the proposition focuses on the (local) response of aggregate outcomes to small shocks in bilateral revenue shifters. It establishes that such responses are a function of the demand elasticity of substitution σ, as well as two matrices evaluated at the initial equilibrium, the matrix of bilateral trade flows $\{X_{ij}^0\}$ and its associated elasticity matrix $\{\theta_{ij}(n_{ij}^0)\}$. In this case, separate knowledge of the extensive and intensive margin elasticities –and thus the distribution of firm fundamentals– is not required conditional on knowing the initial elasticity matrix of bilateral trade flows. In other words, the export decisions of heterogeneous firms only affect aggregate responses to small shocks through the heterogeneous bilateral trade elasticities, $\theta_{ij}^0 = \theta_{ij}(n_{ij}^0)$.[11]

The result hinges on two key observations. First, by definition, the local response of bilateral trade flows is the sum of the local responses of n_{ij} and $\bar{x}_{ij}$, as measured by the trade elasticity θ_{ij}^0 at the initial equilibrium. Thus, what remains to show is that, in changes, equilibrium conditions can be written as a function of bilateral trade flows, or the aggregate outcomes that we solve for. Indeed, we argued in Section 2.3 that budget balance and

[11] In Appendix A.1.2, we show that the same requirements are sufficient to compute responses of aggregate outcomes to small changes in population $\bar{L}_i$ and transfers $\bar{T}_i$. However, to compute responses to small changes in the fixed costs of exporting $\bar{f}_{ij}$ and entry $\bar{F}_i$, we need to know also the initial share of the country's labor force employed to cover fixed costs of exporting, which can be recovered using the functions $(\epsilon_{ij}(n), \rho_{ij}(n))$ and the initial matrix of export firm shares $\{n_{ij}^0\}$.

labor market clearing, (19) and (20), can be expressed in terms of bilateral trade flows and aggregate variables. Additionally, as shown in Appendix A.1.2, the free entry condition in (18) yields the following relationship,

$$d\ln N_i = \sum_j y_{ij}^0 \theta_{ij}^0 \varepsilon_{ij}^0 d\ln n_{ij}, \tag{22}$$

with $y_{ij}^0 \equiv X_{ij}^0 / \sum_{j'} X_{ij'}^0$. The result follows from this expression and equation (11), which guarantees that $\varepsilon_{ij}^0 d\ln n_{ij} = d\ln \epsilon_{ij}(n_{ij})$ is a function of changes in aggregate variables.

Expression (22) reflects a key mechanism associated with firm selection into exporting. As more resources are spent on exporter entry, $d\ln n_{ij} > 0$ for $i \neq j$, the economy's resource constraint, equation (20), links the selection decisions of firms from i across all markets. Differentiating the resource constraint implies that domestic participation declines,

$$d\ln n_{ii} = -\sum_{j\neq i} (y_{ij}^0 \varepsilon_{ij}^0 / y_{ii}^0 \varepsilon_{ii}^0) d\ln n_{ij} < 0. \tag{23}$$

The effect on firm entry depends on the shape of the trade elasticity function, since it summarizes cross-market variation of profit margins in the free entry condition, $d\ln N_i = \sum_{j\neq i} (\theta_{ij}^0 - \theta_{ii}^0) y_{ij}^0 \varepsilon_{ij}^0 d\ln n_{ij}$. Consider the calibrations in the existing literature presented in Figure 1. Whenever $\theta_{ij}^0 > \theta_{ii}^0$ for all $j \neq i$ the total mass of firms in origin i will decrease with exporter entry, $d\ln N_i < 0$ (recall that $\varepsilon_{ij}(n) < 0$).[12] When $\theta_{ij}^0 = \theta_{ii}^0 = \bar{\theta}_i$ for all j, as in the class of constant-elasticity gravity trade models in Arkolakis et al. (2012), changes in resources used to export to different destinations mechanically compensate each other, shutting down this entry channel, $d\ln N_i = 0$.

The second part of the proposition turns to the impact of large changes in bilateral revenue shifters. In this case, the trade elasticity matrix $\{\theta_{ij}(n_{ij})\}$ may endogenously change as the economy moves away from the initial equilibrium, and responses are shaped by the fundamentals' distribution of a new set of marginal firms. One thus needs to track changes in $\theta_{ij}(n)$ induced by responses in n_{ij}, which requires the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$, and the initial matrix of exporter firm shares, $\{n_{ij}^0\}$. Nonetheless, we do not need to know further details about the micro structure of the model, including the joint distribution of firm fundamentals and the initial matrix of exogenous fundamentals (e.g., $\bar{r}_{ij}^0$ or $\bar{f}_{ij}^0$). This part of the proposition is an application of the "hat-algebra" toolkit developed by Dekle et al. (2008), and a generalization of the sufficient statistics result of Arkolakis et al. (2012)

[12]We note that this condition holds if the bilateral trade elasticity is decreasing on the exporter firm share and only a small fraction of domestic firms export, $n_{ij}^0 < n_{ii}^0$.

(Proposition 2) beyond the class of constant-elasticity gravity models.[13]

We can build more intuition for the connection between the two parts of the proposition using a constant-elasticity benchmark,

$$\varepsilon_{ij}(n) = \bar{\varepsilon}_{ij} \quad \text{and} \quad \varrho_{ij}(n) = \bar{\varrho}_{ij}. \tag{24}$$

This special case is a flexible extension of the Pareto variant of Melitz (2003) in Chaney (2008).[14] In this case, the constant-elasticity distribution of fundamentals implies that the trade elasticity $\theta_{ij}(n) = \bar{\theta}_{ij} = 1 - (1 + \bar{\varrho}_{ij})/\bar{\varepsilon}_{ij}$ does not vary with the exporter firm share. We can then use the first part of the proposition to compute aggregate responses to large shocks based solely on knowledge of the bilateral trade elasticities $\{\bar{\theta}_{ij}\}$, by integrating local responses without tracking changes in n_{ij}. Intuitively, when these elasticity functions are constant, the dispersion of firm entry and revenue potentials is also constant across the entire support. This leads to responses of firm export margins that are invariant to the initial conditions.

3.2 The Margins of Welfare Responses to Changes in Bilateral Trade Costs

We now leverage the CES demand preferences in our model to characterize real wage responses to changes in trade costs. This is equivalent to welfare changes under trade balance (i.e., $\bar{T}_i = 0$). We use the intensive margin equation (13) to write the change in the real wage of country j in terms of changes in exogenous and endogenous variables in any origin i:

$$\ln \frac{\hat{w}_j}{\hat{P}_j} = \frac{1}{\sigma - 1} \left(\ln \hat{\bar{r}}_{ij} + \ln \hat{N}_i - \ln \hat{x}_{ij} + \ln \hat{n}_{ij} \frac{\rho_{ij}(n^0_{ij}\hat{n}_{ij})}{\rho_{ij}(n^0_{ij})} \right) + \ln \frac{\hat{w}_j}{\hat{w}_i}, \tag{25}$$

with $x_{ij} \equiv X_{ij}/E_j$ the share of origin i in the expenditures of destination j.

To obtain a decomposition, we take the average of this expression weighted by initial trade shares, x^0_{ij}:

$$\begin{aligned} \ln \tfrac{\hat{w}_j}{\hat{P}_j} \;=\; & \underbrace{\sum_i \frac{x^0_{ij}}{\sigma - 1} \ln \hat{\bar{r}}_{ij}}_{\text{Technology}} + \underbrace{\sum_i x^0_{ij} \ln \frac{\hat{w}_j}{\hat{w}_i}}_{\text{Terms of trade}} - \underbrace{\frac{1}{\sigma - 1} \sum_i x^0_{ij} \ln \hat{x}_{ij}}_{\text{Demand substitution}} \\ + \; & \underbrace{\sum_i \frac{x^0_{ij}}{\sigma - 1} \ln \hat{N}_i}_{\text{Firm entry}} + \underbrace{\sum_i \frac{x^0_{ij}}{\sigma - 1} \ln \hat{n}_{ij} \frac{\rho_{ij}(n^0_{ij}\hat{n}_{ij})}{\rho_{ij}(n^0_{ij})}}_{\text{Firm selection}} \end{aligned} \tag{26}$$

[13] As noted by Costinot and Rodriguez-Clare (2013), the "hat algebra" system for heterogeneous firm models also depends on the elasticity of substitution σ if the entry cost depends on the origin's wage.

[14] In Chaney (2008), $r_{ij}(\omega) = e_{ij}(\omega) = (a_i(\omega))^{\sigma-1}$ and $a_i(\omega) \sim 1 - a^{-\theta}$, leading to $\bar{\varrho}_{ij} = \bar{\varepsilon}_{ij} = -(\sigma - 1)/\theta$.

The first row measures the components of welfare responses that are present in neoclassical trade models. While the "technology" term captures the shock to the exogenous component of the cost of imported goods, the "terms of trade" term measures changes in the endogenous labor cost in origin i (relative to that of j).[15] These two channels are the "traditional" gains from trade in Hsieh et al. (2020), and capture the first-order impact of changes in trade costs on welfare in neoclassical models, such as the economy without wedges in Baqaee and Farhi (2019) (see Theorem 2). In addition, the optimal adjustment of the representative household's consumption bundle creates an offsetting "demand substitution" effect. This component is approximately zero for small shocks (i.e., $\sum_i x_{ij}^0 \ln \hat{x}_{ij} \approx \sum_i dx_{ij} = 0$), but it can be substantial for large shocks.[16]

The second row arises from changes in the available varieties in monopolistic competition models. They capture welfare changes due to endogenous firm entry in each origin i, as well as endogenous selection of heterogeneous firms from each origin i into destination j. Both terms are weighted by the parameter governing love for variety with CES demand, $1/(\sigma-1)$, and the initial share of origin i in j's expenditures, x_{ij}^0. In addition, the magnitude of the "firm selection" gains depends on the mean revenue potential of marginal firms relative to that of infra-marginal firms, $\ln \hat{n}_{ij} \frac{\rho_{ij}(n_{ij}^0 \hat{n}_{ij})}{\rho_{ij}(n_{ij}^0)} \approx \frac{\mathbb{E}[r|e=\epsilon_{ij}(n_{ij}^0)]}{\rho_{ij}(n_{ij}^0)} d\ln n_{ij}$ (which uses the expression for ϱ_{ij} in Section 2.2). Intuitively, the higher the revenue potential of entrants compared to incumbents, the larger are the welfare gains created by the endogenous selection of these marginal firms into destination j. These two channels are equivalent to the indirect effect in Atkeson and Burstein (2010) and the "new" gains from trade in Hsieh et al. (2020), but written in terms of changes in firm selection and the intensive margin elasticities instead of changes in average firm productivity.

It is worth noting that the technology term only depends on exogenous shocks and initial spending shares and is invariant to the shape of the distribution of firm fundamentals. All other terms depend on the distribution of firm fundamentals. The firm terms, in particular, are generally nonzero when countries are asymmetric, since they directly affect welfare in each country j through variety availability.[17] However, Proposition 2.b shows that the equilibrium

[15] The technology term is scaled by $1/(\sigma-1)$ because $\ln \hat{\bar{r}}_{ij}$ measures the demand shift associated with a cost shift (see the definition in (5)). To see this, consider shocks to iceberg trade costs for which $\ln \hat{\bar{r}}_{ij} = -(\sigma-1) \ln \hat{\bar{\tau}}_{ij}$, and the technology term is $-\sum_i x_{ij}^0 \ln \hat{\bar{\tau}}_{ij}$.

[16] The decomposition in Hsieh et al. (2020) does not have the demand substitution term, as their choice of Sato-Vartia weights cancels out substitution across origins.

[17] One exception is the special case of symmetric countries studied in Atkeson and Burstein (2010) where the first-order impact of changes in trade costs on welfare is given only by the technology term. Appendix A.1.3 shows that, when countries are symmetric, the link between firm entry and selection in (22) is identical across countries, which implies that the firm components of welfare exactly offset each other. In contrast, we also show that firm selection has a first-order impact on welfare when countries are asymmetric for the constant-elasticity benchmark with $\bar{\varepsilon}_{ij} = \bar{\varepsilon}$ and $\bar{\varrho}_{ij} = \bar{\varrho}$ for all ij (Chaney, 2008).

is efficient so that, up to a first-order, these terms are not important "on average" across countries. Indeed, only the technology term has a first-order impact on the global average real wage under trade balance:

$$\sum_j \frac{E_j^0}{E^0} d\ln\frac{w_j}{P_j} = \underbrace{\sum_j\sum_i \frac{X_{ij}^0}{E^0}\frac{d\ln\bar{r}_{ij}}{\sigma-1}}_{\text{Global Technology Effect}}, \tag{27}$$

with $E^0=\sum_j E_j^0=\sum_{i,j}X_{ij}^0$ denoting the world output. This formula for global welfare gains is closely related to the one derived by Atkeson and Burstein (2010). Up to a first-order approximation, the difference between welfare gains for country j in (26) and for the world in (27) can be interpreted as arising from between-country reallocation effects induced by responses in terms of trade, firm entry, and firm selection. At the global-level, equation (27) indicates that these reallocation effects cancel each other when we use Negishi (1960) weights.[18]

3.3 The Gains From Trade

We now turn to a preeminent counterfactual exercise: the gains from trade defined as the impact on welfare of moving to autarky. In Appendix A.1.4, we characterize the gains from trade as a corollary of Proposition 3.

Corollary 1. *Consider a counterfactual change in trade costs that moves country j from the trade equilibrium to the autarky equilibrium: $\hat{\bar{r}}_{ij}\to 0$ for all $i\neq j$. Then,*

$$\ln\frac{\hat{w}_j}{P_j} = \frac{1}{\sigma-1}\ln x_{jj}^0 + \frac{1}{\sigma-1}\ln\hat{N}_j + \frac{1}{\sigma-1}\ln\hat{n}_{jj}\frac{\rho_{jj}(n_{jj}^0\hat{n}_{jj})}{\rho_{jj}(n_{jj}^0)} \tag{28}$$

where $\hat{n}_{jj}$ and $\hat{N}_j$ are given by

$$\frac{\epsilon_{jj}(n_{jj}^0\hat{n}_{jj})}{\epsilon_{jj}(n_{jj}^0)} = x_{jj}^0\hat{N}_j\hat{n}_{jj}\frac{\rho_{jj}(n_{jj}^0\hat{n}_{jj})}{\rho_{jj}(n_{jj}^0)} \tag{29}$$

$$\hat{N}_j = \frac{1-\gamma_{jj}(n_{jj}^0\hat{n}_{jj})}{1-\sum_d y_{jd}^0\gamma_{jd}(n_{jd}^0)}, \tag{30}$$

[18]With any other set of weights it is easy to show that these terms would affect global welfare through the impact of shock-induced transfers across countries. The presence of inefficiencies can also lead to additional reallocation terms in welfare. This has been discussed in a context of growth externalities by Perla et al. (2021), firm size wedges by Bai et al. (2024), variable markups by Arkolakis et al. (2019a), or tariffs and exogenous markup wedges by Baqaee and Farhi (2019).

with $\gamma_{ij}(n)$ the share of labor employed to cover the fixed costs of firms from i selling in j, as defined in (OA.21).

Equation (28) follows immediately from the expression in (25) for $i = j$, since the expression for $i \neq j$ is not well-defined for this counterfactual exercise. Accordingly, the first term measures substitution towards domestic goods with $\hat{x}_{jj} = 1/x^0_{jj}$, and is no longer second-order. Conditional on domestic substitution, the two additional terms in (28) arise from the entry and selection decisions of domestic firms, which are given by (29)-(30). The discussion in Section 3.1 indicates that these channels may reduce the gains from trade. When moving to autarky, selection out of exporting ($d\ln n_{ij} < 0$ for $i \neq j$) leads to higher domestic firm survival ($d\ln n_{jj} > 0$) and, thus, higher welfare, $\ln \hat{n}_{jj}\rho_{jj}(n^0_{jj}\hat{n}_{jj})/\rho_{jj}(n^0_{jj}) = \int_{\ln n^0_{jj}}^{\ln n^0_{jj}\hat{n}_{jj}}(1+\varrho_{jj}(u))du > 0$ (since $\varrho_{jj}(n) > -1$). Furthermore, as discussed in Section 3.1, the welfare contribution of domestic firm entry may be positive or negative, depending on whether the bilateral trade elasticity is increasing or decreasing on n.

The expression in (28) is related to the sufficient statistic for the gains from trade in Arkolakis et al. (2012). To see this, we show in Appendix A.1.4 that the combination of the extensive and intensive margins of domestic sales in (29) implies that, locally, $\theta^0_{jj}\varepsilon^0_{jj} d\ln n_{jj} = -\left(d\ln x_{jj} - d\ln N_j\right)$, and

$$d\ln\frac{w_j}{P_j} = -\frac{1}{\tilde{\theta}^0_{jj}}(d\ln x_{jj} - d\ln N_j), \tag{31}$$

where $\tilde{\theta}^0_{jj} \equiv (\sigma-1)\theta^0_{jj}$ is the elasticity of domestic spending to domestic cost at the initial equilibrium. The expression in (31) indicates that what matters for the welfare gains from trade is the domestic trade elasticity, and not a generic "trade elasticity" as pointed by Arkolakis et al. (2012). In fact, the negative contribution to welfare of domestic firm selection is embedded into the domestic trade elasticity, which attenuates the welfare impact of any given change in domestic spending relative to the contribution of demand substitution in (28), as $\tilde{\theta}^0_{jj} > \sigma - 1$. Intuitively, a higher domestic elasticity means that it is easier to substitute foreign varieties for domestic varieties (through both extensive and intensive margins), which attenuates the welfare consequences of having to spend more on domestic varieties. Of course, when all elasticities are constant, one recovers the result in Arkolakis et al. (2012).

We view the above discussion as a synthesis of the results in Melitz and Redding (2015), who stress the importance of varying trade elasticities and firm heterogeneity, with the results of Arkolakis et al. (2012), who stress the sufficient role of the trade elasticity for aggregate responses to trade shock, but in a world with a constant trade elasticity. Relative to them, our characterization indicates that what matters are the (potentially variable) elasticities of trade flows, which summarize the entry and selection decisions of heterogeneous firms. We now turn to the estimation of these elasticities.

4 Estimation Strategy: Semiparametric Gravity

We develop next a semiparametric approach to estimate the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$ using the gravity-like equations for the margins of firm exports in (11) and (13).

We use data on the share of firms from i selling in j, n_{ij}, and their average sales, $\bar{x}_{ij}$, across a set of origin-destination pairs ij, which henceforth we refer to as markets. To leverage cross-market variation for estimation, we assume that the distribution of firm fundamentals has the same shape in all markets belonging to the same group:

Assumption 2. *Origin-destination pairs are divided into G groups, $\mathcal{G}_g$, such that*

$$H_{ij}(r,e) = H_g(r/\bar{\eta}^r_{ij}, e/\bar{\eta}^e_{ij}) \quad \textit{for all} \quad ij \in \mathcal{G}_g. \tag{32}$$

This assumption imposes that, for any two markets belonging to the same group g, the distribution of entry and revenue potentials only differs with respect to the (unobserved) scalars $\bar{\eta}^r_{ij}$ and $\bar{\eta}^e_{ij}$.[19] Importantly, for each group of markets g, we do not impose any parametric restriction on the shape of the distribution of firm fundamentals. The main implication of Assumption 2 is to restrict the model's elasticity functions to be identical across all markets in the same group:

$$\ln \epsilon_{ij}(n) = \ln \epsilon_g(n) + \ln \bar{\eta}^e_{ij} \quad \text{and} \quad \ln \rho_{ij}(n) = \ln \rho_g(n) + \ln \bar{\eta}^r_{ij} \quad \text{for all} \quad ij \in \mathcal{G}_g.$$

Assumption 2 has a long tradition in the estimation of endogenous selection models such as ours. Without this assumption, the results in Heckman and Honore (1990) imply that cross-section outcomes for firms in a single market cannot nonparametrically identify the distribution of firm fundamentals. Intuitively, if we only observe the set of active firms in a single market, parametric assumptions are necessary for extrapolating from the outcomes observed for this set of active firms to recover the distribution of fundamentals of firms that are not active in that market. In contrast, as in Heckman and Honore (1990), Assumption 2 allows the use of *cross-market* variation in firm export margins induced by trade costs for the nonparametric identification of the elasticity functions. In our empirical application, we provide estimates for groups defined in terms of per-capita income, market integration, and other shared characteristics.

[19] Our notation allows groups to be defined as destination-origin pairs over different years. In this case, our strategy could exploit variation over time for the same market while allowing the shape of the distribution to vary across all markets as long as it is constant over time for the same market.

We also impose restrictions on the data generating process of the bilateral shifters of revenue and entry.

Assumption 3. *We observe a vector of bilateral variables, $z_{ij} = \{z_{ij,k}\} \in \mathbb{R}^K$, such that*

$$\begin{aligned} \ln \bar{\eta}^r_{ij}\bar{r}_{ij} &= z_{ij}\kappa^r + \bar{\delta}^r_i + \bar{\zeta}^r_j + \eta^r_{ij}, \qquad \mathbb{E}[\eta^r_{ij}|z_{ij}, D] = 0 \\ \ln \bar{f}_{ij}/\bar{\eta}^e_{ij}\bar{r}_{ij} &= z_{ij}\kappa^e + \bar{\delta}^e_i + \bar{\zeta}^e_j + \eta^e_{ij}, \qquad \mathbb{E}[\eta^e_{ij}|z_{ij}, D] = 0 \end{aligned} \tag{33}$$

where D is the matrix of origin and destination dummies, and (κ^r, κ^e) are real vectors of length K with known first entries (κ^r_1, κ^e_1).

Assumption 3 plays the central role of specifying observable variables z_{ij} whose variation across market allows us to trace out the elasticities of the extensive and intensive margins of firm exports. This assumption has three parts, which we now discuss separately.

The first part of Assumption 3 is the separability of (33). Given origin and destination fixed-effects, bilateral shifters of revenue and entry are the sum of two components: the impact of the observed vectors, $z_{ij}\kappa^r$ and $z_{ij}\kappa^e$, and the unobserved shifters, η^r_{ij} and η^e_{ij}. Together with the equilibrium conditions for entry and sales in (11) and (13) under Assumption 2, equation (33) yields our semiparametric gravity equations:

$$\ln \bar{x}_{ij} - \ln \rho_g(n_{ij}) = z_{ij}\kappa^r + \delta^r_i + \zeta^r_j + \eta^r_{ij} \tag{34}$$

$$\ln \epsilon_g(n_{ij}) = z_{ij}\kappa^e + \delta^e_i + \zeta^e_j + \eta^e_{ij} \tag{35}$$

where $\delta^e_i \equiv \ln(w^\sigma_i) + \bar{\delta}^e_i$, $\zeta^e_j \equiv \bar{\zeta}^e_j - \ln(E_j P^{\sigma-1}_j)$, $\delta^r_i \equiv \ln(w^{1-\sigma}_i) + \bar{\delta}^r_i$, and $\zeta^r_j \equiv \ln\left(P^{\sigma-1}_j E_j\right) + \bar{\zeta}^r_j$. Intuitively, holding all else constant, the comparison of n_{ij} and $\bar{x}_{ij}$ across markets with different observed shifters, $z_{ij}\kappa^r$ and $z_{ij}\kappa^e$, identifies the elasticity functions $\epsilon_g(n)$ and $\rho_g(n)$. Note that the origin and destination fixed effects include endogenous outcomes (like wages and prices). For this reason, we need to maintain the assumption that the bilateral shifters are separable in the effect of the observable vector z_{ij}.[20]

The second part of Assumption 3 is the orthogonality between the observed and unobserved components, $\mathbb{E}[(\eta^r_{ij}, \eta^e_{ij})|z_{ij}, D] = 0$. This is the formal notion of "all else constant" that allows us to trace out $\epsilon_g(n)$ and $\rho_g(n)$ from the responses of n_{ij} and $\bar{x}_{ij}$ to z_{ij}, and is the typical exogeneity assumption in the estimation of gravity equations for trade flows, as reviewed by Head and Mayer (2014). We use the implied moment conditions for the estimation of (34)-(35): for any function $\mathcal{Z}_g(.)$, $\mathbb{E}[\mathcal{Z}_g(z_{ij})(\eta^r_{ij}, \eta^e_{ij})|D] = 0$. In our empirical application, z_{ij} includes trade cost shifters that are commonly used in the literature estimating gravity trade

[20] However, as in Berry and Haile (2014), we could consider arbitrary functions of z_{ij}, $\kappa^r_g(z_{ij})$ and $\kappa^e_g(z_{ij})$, instead of the linear functions, $z_{ij}\kappa^r$ and $z_{ij}\kappa^e$.

models. We follow Chen et al. (2024) by using splines to specify $\mathcal{Z}_g(.)$.

The last part of Assumption 3 is that we know the pass-through from one element of z_{ij} to the bilateral shifters, which we specify to be the first without loss. This assumption is analogous to that imposed by Heckman and Honore (1990) and Berry and Haile (2014). It is necessary in order to separate the impact of the bilateral shifters on n_{ij} and $\bar{x}_{ij}$ and the impact of z_{ij} on bilateral shifters. Intuitively, it specifies the scale of the bilateral shifters in terms of one component of z_{ij}. Such an assumption is implicit whenever observed shifters of trade costs are used for the estimation of the trade elasticity in gravity trade models. In our application, we follow the approach in Caliendo and Parro (2014) and Boehm et al. (2023) by imposing that bilateral variable trade costs are proportional to the cost of ad-valorem import tariffs.

Finally, we impose a basis for the elasticity functions $\rho_g(n)$ and $\epsilon_g(n)$.

Assumption 4. *The elasticity functions $\rho_g(n)$ and $\epsilon_g(n)$ are spanned by restricted cubic splines, $f_m(\ln n)$, over knots $m = 1, ..., M$,*

$$\begin{bmatrix} \ln \rho_g(n) \\ \ln \epsilon_g(n) \end{bmatrix} = \sum_{m=1}^{M} \begin{bmatrix} \gamma^{\rho}_{g,m} f_m(\ln n) \\ \gamma^{\epsilon}_{g,m} f_m(\ln n) \end{bmatrix}. \tag{36}$$

We approximate the shape of $\rho_g(n)$ and $\epsilon_g(n)$ with a cubic spline function over each interval $[\bar{n}_m, \bar{n}_{m+1}]$ of the support $[0, 1]$, as in Ryan (2012). To improve precision, we restrict the bottom and upper intervals to have a linear slope. Our estimates below are based on three intervals ($M = 3$).

Under Assumption 4, we then recover the residuals as a function of parameters, $\Theta \equiv (\kappa^e, \kappa^r, \{\gamma^{\rho}_{g,m}, \gamma^{\epsilon}_{g,m}\}_{g,m=1}^{G,M}, \{\delta^r_i, \delta^e_i, \zeta^r_j, \zeta^e_j\}_{i,j=1}^{N,N})$:

$$\begin{bmatrix} \eta^r_{ij} \\ \eta^e_{ij} \end{bmatrix} = \begin{bmatrix} u^r_{ij}(\Theta) \\ u^e_{ij}(\Theta) \end{bmatrix} \equiv \begin{bmatrix} \ln \bar{x}_{ij} - z_{ij}\kappa^r \\ -z_{ij}\kappa^e \end{bmatrix} + \sum_{m=1}^{M} \begin{bmatrix} -\gamma^{\rho}_{g,m} f_m(\ln n) \\ \gamma^{\epsilon}_{g,m} f_m(\ln n) \end{bmatrix} - \begin{bmatrix} \delta^r_i + \zeta^r_j \\ \delta^e_i + \zeta^e_j \end{bmatrix}.$$

We use the recovered residuals to construct the following Generalized Method of Moments (GMM) estimator for Θ:

$$\min_{\Theta} v(\Theta)' \hat{\Omega} v(\Theta), \quad \text{where} \quad v(\Theta) \equiv \begin{bmatrix} \sum_{ij} \left(u^r_{ij}(\Theta)\mathcal{Z}_g(z_{ij}), u^r_{ij}(\Theta) D_{ij}\right)' \\ \sum_{ij} \left(u^e_{ij}(\Theta)\mathcal{Z}_g(z_{ij}), u^e_{ij}(\Theta) D_{ij}\right)' \end{bmatrix}, \tag{37}$$

and $\hat{\Omega}$ is the two-step optimal matrix of moment weights.

There are two ways to interpret our strategy to estimate $\rho_g(n)$ 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 (Newey and McFadden, 1994). Alternatively, 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).[21] Chen et al. (2024) derive confidence intervals for sieve NPIV estimators, which we report in our robustness analysis below.

5 Estimation Results

In this section, we 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 vary with exporter firm shares and market characteristics.

5.1 Data

Our estimation sample contains 87 origin countries, and their firms' exports to 209 destination countries in 2012.[22] We measure the average firm revenue as $\bar{x}_{ij} \equiv X_{ij}/N_{ij}$ with N_{ij} and X_{ij} denoting the number and sales of firms from i in j, respectively. We obtain N_{ij} and X_{ij} from the OECD Trade by Enterprise Characteristics (TEC) for a set of developed origins, the World Bank Exporter Dynamics Database (EDD) for a set of developing origins, and from administrative customs data for Australia and China. We consider the sales of the origins in our sample to all destinations in the dataset. Appendix Table OA.1 lists all origin countries in our sample, and the associated source for each variable used in estimation.

Turning to the exporter firm share, we note that n_{ij} is defined as the ratio between the number of firms from i selling in j, N_{ij}, and the number of entrants in i, N_i. The challenge to measure n_{ij} is that N_i is not easily available in national statistics, since it includes also entrants that decide to never produce. We circumvent this issue by noting that, although we consider a static model to simplify exposition, our equilibrium is isomorphic to the stationary equilibrium of the dynamic setting in Melitz (2003), where the mass of successful entrants in market ij at any period, $n_{ij}N_i$, exactly replaces the mass of incumbents in ij exiting

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

[22] The choice of the year was determined by data availability, with the goal of maximizing coverage. Our sample accounted for 58% of world trade in 2012. We show that results are similar when we implement estimation in the period of 2010-2014 over which coverage is similar.

exogenously, δN_{ij} with δ denoting the exogenous death rate. Thus, we can measure the exporter firm share as $n_{ij} = n_{ii}N_{ij}/N_{ii}$ where, for origin i at any given period, N_{ii} is the number of active domestic firms and n_{ii} is the survival probability of new domestic entrants.[23] We measure n_{ii} as the one-year survival rate of tradable firms from the OECD Demographic Business Statistics (SDBS), and N_{ii} as the number of active tradable firms from the OECD Demographic Business Statistics (SDBS), the OECD Structural Statistics for Industry and Services (SSIS), and the World Bank Enterprise Surveys.[24]

We build on the gravity literature reviewed by Head and Mayer (2014) to include in z_{ij} the following variables: import tariffs, geographic distance, as well as dummies for trade agreements, shared language, shared currency, and colonial ties.[25] As stated above, we impose the common assumption in the literature that iceberg trade costs are proportional to import tariff costs – for a discussion, see Section 4 of Head and Mayer (2014). Thus, we set the pass-through parameters for tariffs to $\kappa_1^r = -\kappa_1^e = 1 - \sigma$. Throughout our analysis, we use estimates in the literature for the elasticity of substitution σ, and, in particular, set $\sigma = 3.2$ to match the mean estimate of the cross-firm elasticity in Redding and Weinstein (2024).[26]

The availability of data on $\bar{x}_{ij}$, n_{ij}, and z_{ij} determines our sample for the implementation of the estimator in (37). In Appendix B.1, we report descriptive statistics of our sample. For for each origin i, Table OA.2 reports the number of destinations with positive trade, along with the average and the standard deviation of the exporter firm share and average firm exports across these destinations. In addition, Figure 2 summarizes the distribution of $\ln(n_{ij})$ for all markets in our sample. Since n_{ij} is the only input of the elasticity functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$, we are only able to precisely estimate these functions in the part of the support in which we observe values of n_{ij}.

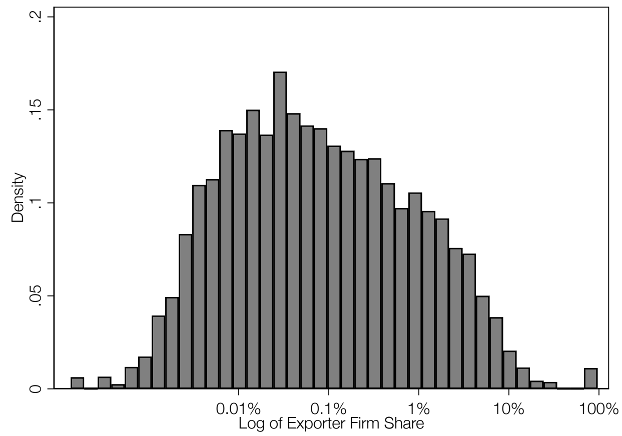
[23]Intuitively, our approach implies that a low survival rate represents a large pool of entrants that pay the sunk entry cost but fail to be productive enough to survive. A high survival rate reflects instead that most firms paying the entry cost are successful in production. Notice that our approach is more general than that in prior research imposing that $N_i = N_{ii}$ and $n_{ii} = 1$ (e.g., Fernandes et al., 2023), which shuts down changes in domestic firm composition that were empirically documented by Pavcnik (2002) and Trefler (2004), and theoretically characterized by Melitz (2003).

[24]We have survival rates for 27 origins 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 below that our results are robust to excluding from the sample countries without data on survival rates. We also show that our results are similar when we use survival rates over longer periods.

[25]The Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) is the main source of bilateral variables in z_{ij}. The only exception is the import tariff cost, which we define as the log of one plus the simple average of the ad-valorem equivalent tariff that j applies to imports from i across all 6-digit HS goods reported in the World Bank WITS. We show below that our estimates are similar but less precise if we use an instrumental variable for tariffs based on the strategy in Boehm et al. (2023) leveraging changes in MFN tariffs between 2002 and 2012.

[26]We show below that our main conclusions are robust to alternative assumptions about the pass-through from import tariffs to revenue and entry shifters. Alternatively, one can design a strategy to estimate σ using firm-level microdata on sales and prices for at least one market.

Figure 2: Empirical Distribution of Exporter Firm Shares, 2012



Note. Empirical distribution of $\ln(n_{ij})$ in the cross-section of origin-destination pairs in 2012.

5.2 Estimates of Semiparametric Gravity of Firm Exports

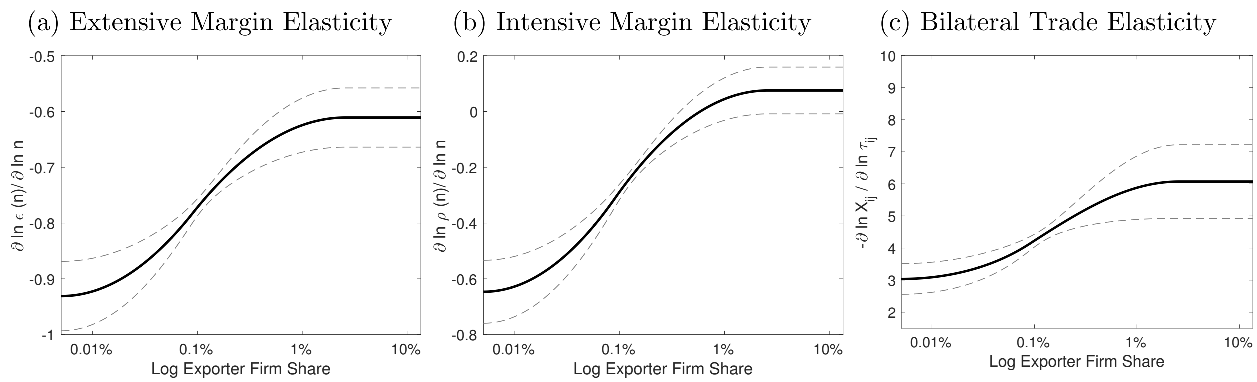
5.2.1 Elasticity Heterogeneity with Respect to Exporter Firm Share

We start by implementing the estimation of (37) for a single group pooling all markets. With Assumption 2, this restricts the shape of the distribution of entry and revenue potentials to be identical across all origin-destination pairs. Such an assumption is implicit in models imposing that all countries have the same gravity trade elasticity or the same shape parameter of the distribution of firm fundamentals. In our model, it yields common elasticity functions for all markets, which restricts the adjustment margins of firm exports to only vary across markets due to variation in the initial exporter firm share.

Figure 3 presents our semiparametric estimates. Panels (a) and (b) report the extensive and intensive margin elasticity functions respectively, and Panel (c) reports the elasticity of bilateral trade flows with respect to bilateral trade costs. The solid lines are the baseline estimates obtained from (37), and the dashed lines are the associated 90% confidence intervals implied by robust standard errors.

Panel (a) shows that the extensive margin elasticity, $\varepsilon(n) \equiv d\ln\epsilon(n)/d\ln n$, is increasing on the exporter firm share. We estimate that $\varepsilon(n_{ij}) = -0.9$ for n_{ij} below 0.01%, but it increases to $\varepsilon(n_{ij}) = -0.6$ for n_{ij} above 10%. This means that the entry potential distribution is more dispersed at the upper end of the support. Thus, the impact of trade shocks on firm entry is weaker in markets with few exporters that are mainly populated by firms with high entry potential. To understand magnitudes, Appendix Figure OA.1 shows that our estimates imply that a 1 log-point increase in trade costs causes an estimated decline in firm entry of 2.4 log-points in markets with low n_{ij}, but the decline is 3.6 log points for those with high n_{ij}.

Figure 3: Semiparametric Gravity of Firm Exports – Single Group



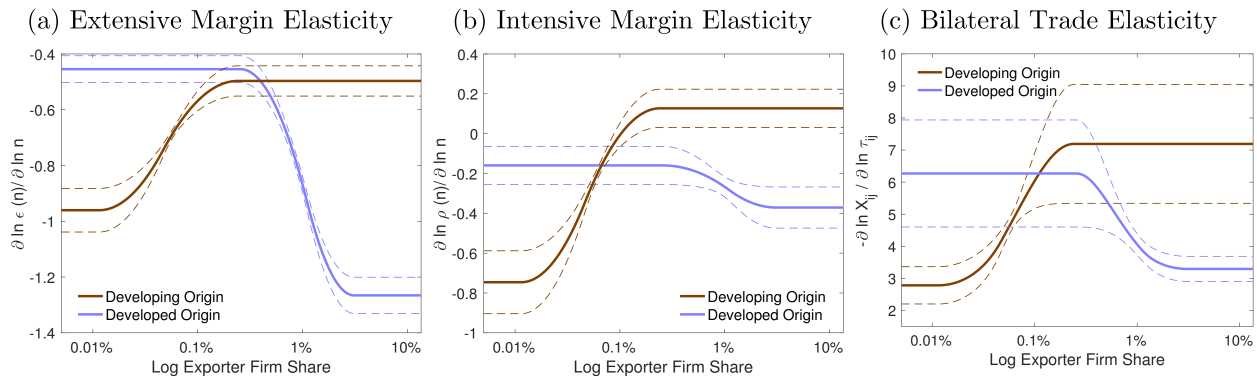
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Turning to panel (b), we find that the intensive margin elasticity, $\varrho(n) \equiv d\ln\rho(n)/d\ln n$, also increases with the exporter firm share. For markets with low n_{ij}, our estimate of $\varrho(n_{ij}) = -0.6$ implies that the average revenue potential of marginal entrants, $\mathbb{E}[r|e = \epsilon(n_{ij})]$, is 40% of that of incumbents, $\rho(n_{ij})$ (recall that $1 + \varrho(n) = \mathbb{E}[r|e = \epsilon(n)]/\rho(n)$). In contrast, markets with high n_{ij} have $\varrho(n_{ij})$ of roughly zero, implying that they have similar marginal and infra-marginal firms. Accordingly, revenue potential differences are larger among firms with high entry potential that operate in markets with a low n_{ij}. These markets exhibit stronger composition effects that attenuate the response of average firm exports to changes in trade costs. Appendix Figure OA.1 reports that a 1 log-point increase in trade costs causes a decline in average firm exports of 0.6 log-points in markets with low n_{ij}, but the decline is 2.5 log points in markets with high n_{ij}.

Panel (c) then reports the trade elasticity that combines our estimates of the extensive and intensive margin elasticities. Since both margins are increasing on the firm exporter share, so is the elasticity of bilateral trade flows. We estimate an elasticity of roughly three in markets with n_{ij} below 0.01%, which increases to six in markets with n_{ij} above 10%. Our estimates are within the range in the literature using cross-section variation to estimate how bilateral trade responds to trade costs (Head et al., 2014), but are slightly higher than the long-run estimates of Boehm et al. (2023) using exogenous tariff shocks. However, our trade elasticity estimates vary with the exporter firm share, which indicates that firm heterogeneity affects the response of aggregate outcomes to changes in trade costs.

We find substantial heterogeneity across markets in the elasticity of trade flows to trade costs. This contrasts with the benchmark in the quantitative gravity literature of constant-elasticity specifications. Indeed, Appendix Figure OA.2 shows that we reject the

Figure 4: Semiparametric Gravity of Firm Exports – Developed and Developing Origins



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination. We assume that there are two groups of markets ($G = 2$) defined by whether the origin country is developed (light purple) or developing (dark brown), as defined in Table OA.2. Panels (a) and (b) report respectively the elasticity of $\epsilon_g(n)$ and $\rho_g(n)$ with respect to the exporter firm share n, $\varepsilon_g(n)$ and $\varrho_g(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

constant-elasticity benchmark obtained from the estimation of (37) under (24) (instead of Assumption 4). Our semiparametric estimates imply that heterogeneity in exporter firm shares gives rise to substantial heterogeneity in trade elasticities, as illustrated by the empirical distribution of trade elasticities reported in Appendix Figure OA.3.

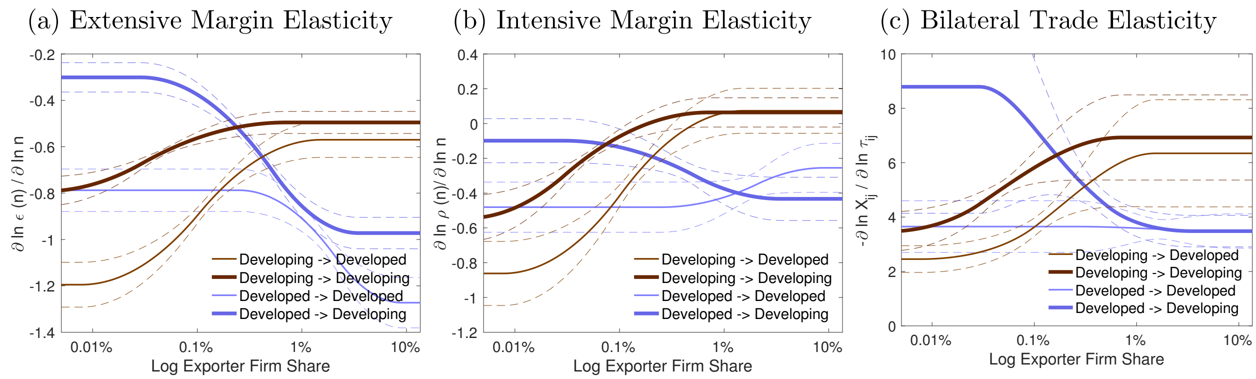
We note that other commonly-used parametric assumptions also yield heterogeneous trade elasticities from firm decisions. However, in contrast with our estimates, Figure 1 shows that all elasticity functions are decreasing on the exporter firm share under the assumption that the distribution of firm fundamentals is either log-normal, truncated Pareto, or a combination of both. Appendix Figures OA.2 indicates that their implied elasticity functions lie outside the confidence interval of our semiparametric estimates for at least part of the support.

5.2.2 Elasticity Heterogeneity with Respect to Country Development

We investigate next how the elasticity functions of firm exports vary across markets. Formally, we relax the assumption of a single group behind the estimates above, and instead assume that the distribution of revenue and entry potentials has the same shape within groups of markets with similar observable characteristics.

We first estimate elasticity functions that are specific to the level of development of the origin country. This allows developing and developed countries to differ in terms of the dispersion of firm-level productivity, in line with the evidence in Hsieh and Olken (2014). Figure 4 reports estimates separately for markets whose origin country is developing (dark brown) and developed (light purple), as defined by the World Bank (see Appendix Table OA.2). Estimated elasticity functions for developing origins are increasing on the exporter

Figure 5: Semiparametric Gravity of Firm Exports – Developed and Developing Countries



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination. We assume that there are four groups of markets ($G = 4$) defined by whether either the origin or destination is developed or developing, as defined in Table OA.2. Panels (a) and (b) report respectively the elasticity of $\epsilon_g(n)$ and $\rho_g(n)$ with respect to the exporter firm share n, $\varepsilon_g(n)$ and $\varrho_g(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

firm share and qualitatively similar to the pooled estimates reported in Figure 3 (due to developing origins being over-represented in our sample). In contrast, estimated elasticity functions for developed origins are decreasing on the exporter firm share. We note that the estimates for the two groups are statistically different not only from each other, but also across levels of the exporter firm share.

Our estimates suggest that developing countries have a higher density of firms with low levels of entry and revenue potentials. As a result, Appendix Figure OA.4 shows that developing origins exhibit stronger impacts of trade shocks on both the extensive and intensive margins in markers with high n_{ij} (that have relatively more low entry potential firms). In comparison, developed origins have a higher relative mass of firms with high entry and revenue potentials, which leads to stronger responses to trade shocks in markets with a low n_{ij} (that have relatively more high entry potential firms). Our estimates indicate that the evidence in Bas et al. (2017) for French exports across destinations also holds for a wider set of developed countries; that is, their trade elasticity decreases with the exporter firm share. We also note that the qualitatively distinct shape of the trade elasticity function for the two country groups is consistent with Hsieh and Klenow (2009) and Hsieh and Olken (2014), who show that developing countries have a fatter tail of low-productivity firms (which have low entry and revenue potentials).[27]

We then investigate whether the elasticity functions vary systematically with the development level of the destination country. This could be the case if the destination's income

[27]Appendix Figure OA.5 reports estimates for four country groups based on income level, as defined by the World Bank. Despite wider confidence intervals due to fewer markets in each group, we estimate elasticity functions that are steeper for less developed countries, in line with Figure 4.

affects the distribution of trade costs or variety quality (Waugh, 2010; Khandelwal, 2010). In Figure 5, we consider four groups based on whether either the origin or destination are developed. For developing origins, our estimates suggest that the elasticity functions have a similar shape in developing or developed destinations. However, the adjustment margins of firm exports in developed origins depend on the level of development of the destination. As panel (a) shows, the extensive margin elasticity is closer to zero in developing than in developed destinations, and more so in markets with low n_{ij}. Panel (b) indicates that the intensive margin elasticity is decreasing on the exporter firm share for developing destinations, but slightly increasing for developed destinations. Interestingly, our estimated trade elasticity between developed countries is roughly constant at four, which is remarkably close to existing estimates such as those in Simonovska and Waugh (2014) –see also the summary of structural gravity estimates discussed in Head and Mayer (2014)–. For other markets instead, our estimates indicate that the trade elasticity varies with exporter firm shares.

5.3 Robustness

Within-Sector Variation. Our estimates so far have pooled together firms in all sectors. It is possible however that the elasticity heterogeneity documented above is driven by cross-market variation in sectoral firm composition. To address this concern, we now use within-sector variation to estimate the elasticity functions by defining markets as origin-destination-sector triplets and fixed effects as origin-sector and destination-sector.[28] Appendix Figure OA.6 shows that the shape of the within-sector estimates are similar to the baseline estimates above, with wider confidence intervals due to the smaller number of countries in our sample. In addition, Appendix Figures OA.7-OA.9 report sector-specific estimates of the elasticity functions that are consistent with the multi-sector model discussed in Section 2.4. We do so under the assumption that the sector-specific elasticity functions are the same in all origin-destination pairs due to the lower number of countries in our sector-level sample. While most sectors have similar shapes for the extensive elasticity function, they differ in their intensive margin elasticities. Combining the two margins, the bilateral trade elasticity has a similar shape in all sectors, despite its level varying across sectors.

Other Dimensions of Elasticity Heterogeneity. Appendix Figure OA.10 investigates whether the elasticity functions vary with determinants of market integration. This could be the case

[28]We build the sample of origin-destination-sector triplets using the same data sources described in Section 5.1, which provide sector-level data for 46 origins (see Table OA.1). Our sector definition follows Boehm et al. (2023), as reported in Appendix Figure OA.9. The vector z_{ij} remains the same, but we use instead the simple average of import tariffs across 6-digit HS in each sector.

for example if deeper levels of integration have a disproportional impact on the trade costs of smaller firms. Panel (a) shows that deeply integrated markets, defined as those with a trade agreement and a common currency, also have a trade elasticity that is increasing on the exporter firm share, with a threshold shifted to the right due to the higher levels of exporter firm share in this subsample of markets. Finally, panel (b) reports that estimates are also similar when we consider two market groups defined by whether they either have a common language or colonial ties. Overall, our estimates indicate that these determinants of market integration do not affect the strength of the adjustment margins of firm exports conditional on the exporter firm share.

Alternative Cost Shifter. Our baseline estimates impose that import tariffs do not affect fixed costs. In panels (a) and (b) of Appendix Figure OA.11, we show that estimates are similar if we assume instead that import tariffs affect both variable and fixed trade costs. In addition, panels (c) and (d) of Appendix Figure OA.11 report estimates with σ given by the 25th and 75th percentiles of the estimates in Redding and Weinstein (2024).

Our estimation strategy also imposes that bilateral average tariffs are orthogonal to unobserved shifters of bilateral trade conditional on bilateral gravity variables in z_{ij} and origin and destination fixed effects. We evaluate the robustness of our estimates with respect to this assumption by considering an instrumental variable for import tariffs inspired by Boehm et al. (2023). In particular, we define $z_{ij}^{\text{tariffIV}} = \Delta_{2002-2012} \ln(1 + \text{MFNtariff}_j) \times 1[i, j \in \text{WTO}_{2002}] \times 1[i, j \notin \text{FTA}_{2002}]$, so that we only leverage bilateral variation in tariffs stemming MFN tariff reductions in the decade preceding our sample year that affected WTO members without a free trade agreement. As such, we do not use in estimation time-invariant bilateral determinants of tariffs nor variation induced by tariff reductions on specific partners. Appendix Figure OA.12 reports that estimates obtained with this alternative set of moments are similar to our baseline estimates. However, confidence intervals are wider because z_{ij}^{tariffIV} explains only 18% of the variation in bilateral tariffs.

Alternative Specifications. Our baseline confidence intervals are valid under Assumption 4. We now instead follow Chen et al. (2024) to provide confidence intervals under the assumption that (36) is a basis for the nonparametric estimation of $\epsilon(n)$ and $\rho(n)$. Appendix Figure OA.13 shows that this weaker assumption only slightly widens confidence intervals.

Our baseline estimates allow the elasticity functions to differ across three intervals of the support, as specified in Assumption 4 with $M = 3$. In Appendix Figure OA.14, we investigate the robustness with respect to this specification choice by allowing elasticity functions to vary across five intervals of the support; that is, we specify $M = 5$ in Assumption 4. This alternative specification yields similar estimates, albeit less precise.

Alternative Sample. We show that estimates are similar when (i) we use data for the years of 2010 or 2014 with similar sample coverage (Appendix Figures OA.15–OA.16), (ii) we measure n_{ii} using three-year survival rates (Appendix Figure OA.17), and (iii) we exclude from the sample origin countries with imputed survival rates (Appendix Figure OA.18).[29]

6 Quantifying The Aggregate Implications of Firm Export Decisions

We conclude by quantifying the contribution of firm export decisions for the aggregate impact of changes in trade costs. In doing so, we combine the elasticity estimates of Section 5 with the theoretical results of Section 3.

6.1 Uniform Changes in Bilateral Trade Costs

We first consider a uniform reduction of 1% in the bilateral trade cost between all origins and destinations starting from the observed equilibrium for $\{X^0_{ij}, n^0_{ij}\}$ in 2012. We focus on a uniform shock because its heterogeneous impact across countries comes solely from heterogeneity in trade elasticities and initial conditions. We then use Proposition 3.b to compute the model's counterfactual predictions for changes in all outcomes, which we feed into expression (26) to obtain the associated welfare responses and its components.

Panel A of Table 1 reports counterfactual predictions by country group. We use the semiparametric estimates in Figure 5, which allow for elasticity heterogeneity with respect to both exporter firm shares and country development level. The second column reports the average welfare response across all countries (first row), the subset of developed countries (second row) and the subset of developing countries (third row), weighted by each country's aggregate expenditure in 2012 and normalized by the shock size of 0.01. The other columns display the average of each component of welfare responses in (26) divided by the overall change reported in the second column.

The first row shows that, if trade costs were to decline by 1% for all countries, then average global welfare would increase by 0.032%. In line with the discussion in Section 3.2, the average global welfare response is entirely given by the technology term, since the efficiency of the equilibrium implies that all other terms represent redistribution across countries and,

[29]We note that we drop 75% of the origin-destination pairs in this case, which widens confidence internals.

Table 1: Impact of Reductions in Bilateral Trade Costs on Welfare and its Components

Group of Countries	Welfare Elasticity (×100)	Contribution to Welfare Elasticity				
		Neoclassical Components			Firm Components	
		Technology	Terms of trade	Substitution	Entry	Selection
Panel A: For all origins and destinations						
All	3.17	101.6 %	-2.0 %	0.6 %	3.8 %	-4.1 %
Developed	4.05	93.3 %	-3.2 %	0.5 %	-0.7 %	10.1 %
Developing	1.68	136.0 %	3.2 %	1.3 %	22.0 %	-62.4 %
Panel B: For developing origins (beneficiaries) in developed destinations (donors)						
All	0.32	96.7 %	3.0 %	0.7 %	13.6 %	-14.0 %
Donors	0.31	146.4 %	-60.5 %	0.9 %	-0.8 %	14.1 %
Beneficiaries	0.25	0.0 %	152.0 %	0.5 %	46.1 %	-98.6 %

Note. Starting from the observed equilibrium in 2012, we compute the counterfactual equilibrium implied by a reduction of 1% in bilateral trade costs, i.e. $\tau_{ij} = 0.99$ for all $i \neq j$, between all countries (Panel A), and from developing origins in the GSP list of the developed destinations in our sample conceding preferential treatment under GSP rules (Panel B). For each group of countries, the second column of each panel reports 100 times the average log-change in real wage, weighted by each country's aggregate expenditure in 2012 and normalized by the shock size of 0.01. The remaining columns report the average of each component in (26) divided by the value reported in the second column. Counterfactual predictions computed with Proposition 3.b, given the elasticity estimates reported in Figure 5. We display welfare changes and its components for each country in Appendix Figure OA.19.

thus, tend to cancel each other at the world-level –bearing small differences because trade is not balanced in our simulations–.

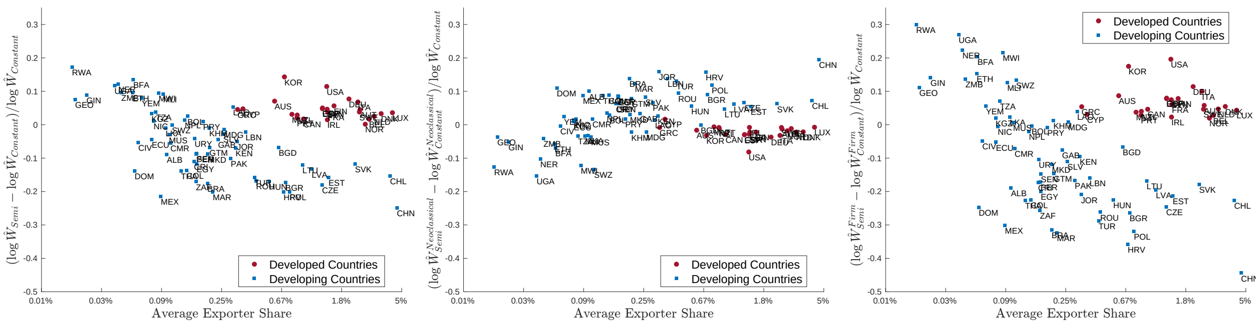
The remaining rows of Panel A indicate that the response of welfare to a uniform reduction in trade costs is larger for developed than developing countries. The primary reason for this difference is the larger technology term for developed countries, which follows from their higher trade openness in 2012. The other two neoclassical terms are small for both groups of countries; terms of trade because relative wages change little for a uniform shock across all countries, and demand substitution because it is second-order for small shocks.

The difference in welfare responses for developed and developing countries is further amplified by the firm components, which increase welfare for developed countries but reduce welfare for developing countries. This is a direct consequence of the elasticity estimates in Figure 5, as we now explain separately for the firm entry and selection channels.

Consider first the firm entry component, whose response follows the intuition in Section 3.1. For developed countries, we estimate a trade elasticity with other developed countries that does not vary much with the exporter firm share. As a result, since developed countries mainly trade with other developed countries, they experience small changes in firm entry; that is, $\hat{N}_i$ is close to zero for developed countries (see Appendix Figure OA.20.a). In contrast, among developing countries, the increasing trade elasticity in Figure 5 leads to domestic firm entry following the increase in the number of exporters caused by the reduction in trade costs. Thus, firm entry has a positive contribution for welfare in developing countries, but

Figure 6: Impact of a Uniform Reduction in Bilateral Trade Costs on Welfare and its Components: The Role of Parametric Assumptions

(a) Welfare Change (b) Neoclassical Components (c) Firm Components



Note. We consider the impact of a reduction of 1% in bilateral trade costs between all countries starting from the observed equilibrium in 2012, i.e. $\tau_{ij} = 0.99$ for all $i \neq j$,. Panel (a) reports in the vertical axis is the difference in welfare responses predicted by the semiparametric and constant-elasticity specifications for each country, divided by the welfare response implied by the constant-elasticity benchmark, and the horizontal axis is the log of the average exporter share of that country in 2012. The other two panels report analogous scatter plots, but the vertical axis is instead the difference in components of predicted welfare responses, divided by the overall welfare response implied by constant-elasticity benchmark. Panel (b) does this for the sum of the neoclassical components associated with technology, terms of trade, and demand substitution in (26), and panel (c) for the sum of the firm components associated with entry and selection in (26). Semiparametric and constant-elasticity predictions use the elasticity estimates in Figures 5 and OA.2, respectively.

this contribution is only equivalent to 22% of the average welfare gains.

The last column of the table reports the contribution of firm selection for welfare. It is positive and equivalent to 10% of gains for developed countries, but it is negative and equivalent to -62% of gains for developing countries. The trade cost reduction causes an increase in the average number of foreign varieties in all countries, but the increasing extensive margin elasticity in developing countries induces a larger drop in domestic firm selection as implied by equation (23). This leads to a decline in the expenditure-weighted average number of firms operating in developing countries.

We finally compare the predictions obtained using elasticity estimates from semiparametric and constant-elasticity specifications, as reported in Figures 5 and OA.2, respectively. In Figure 6 panel (a), the vertical axis is the difference in welfare responses predicted by the semiparametric and constant-elasticity specifications for each country, divided by the welfare response implied by the constant-elasticity benchmark. The horizontal axis is the log of the average exporter share of that country in 2012, which is what our estimated elasticities are a function of. In the other panels, the vertical axis is instead the difference in a component of welfare, divided by the overall welfare response implied by the constant-elasticity benchmark.

Panel (a) shows that parametric assumptions may have substantial effects on welfare predictions for different groups of countries. Relative to the constant-elasticity benchmark, our semiparametric estimates yield predicted welfare responses that are typically lower for developing countries. The average difference is 10% across developing countries, but it is as high as 21% and 25% for Mexico and China, respectively. In contrast, our semiparametric estimates yield larger welfare gains for developed countries, with an average of 5% and largest differences of 12% and 11% for South Korea and the United States, respectively. The deviations are systematically related with the country's average exporter firm share, with correlations of -0.71 and -0.24 for developing and developed countries, respectively.

Panel (b) shows that only a small fraction of the deviation comes from different predictions for the sum of neoclassical components, which is mainly driven by distinct terms of trade predictions, as the two specifications have identical technology terms and small substitution terms. Panel (c) indicates that the firm component is the main force behind the deviation between the two specifications and its correlation with the average exporter firm share. This follows from our estimates of the trade elasticity, which are systematically related with the firm exporter shares through the export decisions of heterogeneous firms. In fact, Appendix Figure OA.21 shows that heterogeneity in initial exporter firm share induces heterogeneity in trade elasticities and, as a result, large differences between the two specifications in predicted responses of trade flows, along both the extensive and intensive margins.

6.2 Heterogeneous Changes in Bilateral Trade Costs

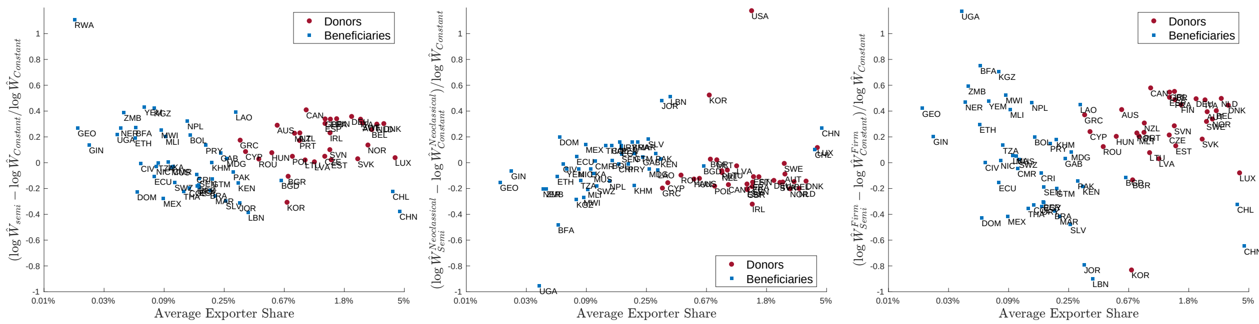
Our next exercise simulates an asymmetric shock to bilateral trade costs across countries that is motivated by the rules of the Generalized System of Preference (GSP). Under those rules, developed countries concede preferentially lower import barriers to a subset of developing countries (i.e., those in the country's GSP list). We consider a counterfactual in which developed countries reduce further barriers on imports from countries in their GSP lists. In particular, we again reduce bilateral trade costs by 1%, but now only for developing origins in the GSP list of each developed destination in our sample that concedes preferential treatment under GSP rules.[30] This exercise leverages a realistic policy choice to illustrate how our estimates of heterogeneous trade elasticities interact with heterogeneous changes in bilateral trade costs.

Panel (b) of Table 1 reports average welfare gains for all countries, the developed countries reducing import costs (donors), and the developing countries benefiting from the reduction (beneficiaries). The second column indicates that this shock has a smaller impact on global welfare relative to the uniform reduction in Panel (a), given that it affects only a subset of

[30] We obtain the list of beneficiaries of GSP in 2012 for the following developed countries in our sample: the European Union, the United States, Japan, Australia, Canada, New Zealand, and South Korea. In reality, these countries reduce tariffs imposed on imports from the developing countries in their GSP lists. Our counterfactual exercise instead considers a hypothetical reduction in import barriers that does not affect tax revenue. As such, it should be seen as a reduction in non-tariff barriers, like sanitary and inspection requirements or technical barriers.

Figure 7: Impact of Reducing the Cost of Exporting from Developing to Developed Countries on Welfare and its Components: The Role of Parametric Assumptions

(a) Welfare Change (b) Neoclassical Components (c) Firm Components

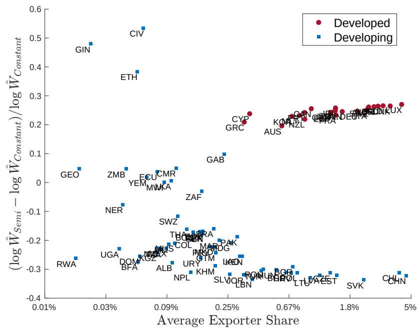


Note. We consider the impact of a reduction of 1% in bilateral trade costs from developing countries in the GSP list to developed countries that concede preferential treatment to countries in the GSP list. Panel (a) reports in the vertical axis is the difference in welfare responses predicted by the semiparametric and constant-elasticity specifications for each country, divided by the welfare response implied by the constant-elasticity benchmark, and the horizontal axis is the log of the average exporter share of that country in 2012. The other two panels report analogous scatter plots, but the vertical axis is instead the difference in components of predicted welfare responses, divided by the overall welfare response implied by the constant-elasticity benchmark. Panel (b) does this for the sum of the neoclassical components associated with technology, terms of trade, and demand substitution in (26), and panel (c) for the sum of the firm components associated with entry and selection in (26). Semiparametric and constant-elasticity predictions use the elasticity estimates in Figures 5 and OA.2, respectively. We report the same ranges for all figures and omit the United States in Panel (a), as it extends below the displayed range as the denominator is very close to zero.

the trading partners in the world. We find that again welfare gains are larger for developed countries, but now due to different mechanisms. The shock only reduces import costs for donor countries, so only these countries have a positive technology term. Because now the shock is heterogeneous, there is a substantial contribution of terms of trade. Donor countries experience a deterioration in their terms of trade, which reduces welfare by an equivalent of 60% of their overall gain. The opposite is true for developing countries. The contributions of the firm components in this case are qualitatively similar to those obtained with the uniform reduction in trade costs, but larger in magnitude.

In Figure 7, we present an analog of Figure 6 to investigate the role of parametric assumptions for this alternative counterfactual. Panel (a) shows that our semiparametric estimates yield welfare responses that can be substantially different from those implied by the constant-elasticity benchmark, with differences of more than 30% in absolute value for several developing and developed countries. Interestingly, panel (b) shows that neoclassical terms are more important for heterogeneous changes in trade costs, since they generate movements in terms of trade. However, such differences are only weakly correlated with the country's average exporter share. Panel (c) indicates that firm components also lead to substantial deviations in welfare predictions, again systematically linked to the country's average exporter firm share. This pattern reflects substantial differences in predicted responses in firm export margins, as detailed in Appendix Figure OA.22.

Figure 8: The Gains From Trade



Note. Gains from trade is minus the real wage change implied by moving from the observed equilibrium in 2012 to autarky, computed with the formula in Corollary 1. The vertical axis is the difference in welfare responses predicted by the semiparametric and constant-elasticity specifications for each country, divided by the welfare response implied by the constant-elasticity benchmark, and the horizontal axis is the log of the average exporter share of that country in 2012. Semiparametric and constant-elasticity predictions use the elasticity estimates in Figures 5 and OA.2, respectively.

6.3 Moving From Autarky to Free Trade

Our final counterfactual exercise relies on Corollary 1 to compute the welfare gains from trade; i.e., a move from the equilibrium in 2012 to autarky. Figure 8 shows that our semiparametric specification yields gains from trade that are higher for developed countries (by an average of 24%), but lower for developing countries (by an average of 17%) relative to the constant-elasticity benchmark. The difference can be substantial: it is 53% higher for Guinea on one extreme, but 34% lower for Slovakia on the other extreme.

These results follow again from the intuition in Section 3.1. The differences in predicted gains from trade come entirely from entry and selection of domestic firms. When the average exporter firm share is higher, more resources are used for exporting in the trade equilibrium, which creates competitive pressure in domestic firm entry and selection through the slope of the trade elasticity function. Indeed, Appendix Figure OA.23 shows that both margins are systematically related to the country's average exporter share in 2012, with larger changes for developing countries due to their more pronounced heterogeneity in trade elasticities. Heterogeneous firms and resulting elasticities imply heterogeneous effects on the gains from trade across developed and developing countries.

Other Parametric Assumptions. Appendix Figure OA.24 compares the gains from trade implied by our semiparametric estimates to those obtained under the assumption that the firm-productivity distribution is either Truncated Pareto in panel (a) (as in Melitz and Redding (2015)) or Log-normal in panel (b) (as in Head et al. (2014)). Because of the different shape of the elasticity functions, we find substantial deviations (averaging 69% with truncated Pareto and 45% with log-normal in absolute terms). As before, using the decomposition in equation (28), all differences in the predicted gains from trade are due to entry and selection of domestic firms. These discrepancies in predictions underscore the critical need to ensure that the elasticity functions derived from these alternative parameterizations are aligned with their empirical counterparts reported in Section 5.

7 Conclusion

We propose a new way to measure the aggregate implications of firm heterogeneity in monopolistic competition model with CES demand. We show that firm heterogeneity affects the extensive and intensive margins of firm exports through two nonparametric elasticity functions, which summarize all the key partial and general equilibrium predictions of the model. We estimate our model's semiparametric gravity equations for firm export margins, which indicate that trade elasticities vary with the number of exporters and the country's development level. Compared to constant-elasticity gravity models, our estimates yield gains from trade that are larger in developed countries but smaller in developing countries.

We view our work as a step toward moving beyond the constant-elasticity-of-trade paradigm in international trade. While this framework has provided a remarkable service to the field, it has also imposed constrains on the positive and normative implications of models. An important next step would be to apply our insights to a more flexible model for international trade: one that features multiple sectors and factors but does not rely on constant elasticities, either in demand (e.g., CES) or in supply (e.g., due to a Pareto distribution).

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.

Allen, Treb, Costas Arkolakis, and Yuta Takahashi, ‘‘Universal gravity,’’ *NBER Working Paper*, 2014, (w20787).

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

Antras, Pol, Teresa C Fort, and Felix Tintelnot, ‘‘The margins of global sourcing: Theory and evidence from us firms,’’ *American Economic Review*, 2017, *107* (9), 2514--2564.

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 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.

Atkeson, Andrew and Ariel Burstein, "Innovation, Firm Dynamics, and International Trade," *Journal of Political Economy*, 2010, *118* (3), 433--489.

Bai, Yan, Keyu Jin, and Dan Lu, "Misallocation under trade liberalization," *American Economic Review*, 2024, *114* (7), 1949--1985.

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

Baqaee, David Rezza, Emmanuel Farhi, and Kunal Sangani, "The darwinian returns to scale," *Review of Economic Studies*, 2024, *91* (3), 1373--1405.

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, "Exporting and Productivity in the USA," *Oxford Review of Economic Policy*, 2004, *20* (3), 343--357.

—, **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, "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 1994, *25* (2), 242--262.

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

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Gains From Input Trade in Firm-Based Models of Importing," 2015.

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

Boehm, Christoph E, Andrei A Levchenko, and Nitya Pandalai-Nayar, ''The long and short (run) of trade elasticities,'' *American Economic Review*, 2023, *113* (4), 861--905.

Broda, Cristian and David Weinstein, ''Globalization and the Gains from Variety,'' *Quarterly Journal of Economics*, 2006, *121* (2), 541--585.

Burstein, Ariel and Jonathan Vogel, ''International trade, technology, and the skill premium,'' *Journal of Political Economy*, 2017, *125* (5), 1356--1412.

Bustos, Paula, ''Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinean Firms,'' *American Economic Review*, 2011, *101* (1), 304--340.

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

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.

_, Timothy Christensen, and Sid Kankanala, ''Adaptive estimation and uniform confidence bands for nonparametric structural functions and elasticities,'' *Review of Economic Studies*, March 2024.

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.

_ and Andrés Rodríguez-Clare, ''The US gains from trade: Valuation using the demand for foreign factor services,'' *Journal of Economic Perspectives*, 2018, *32* (2), 3--24.

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

Dhingra, Swati and John Morrow, ''The Impact of Integration on Productivity and Welfare Distortions Under Monopolistic Competition,'' *mimeo, LSE*, 2012.

Dickstein, Michael J and Eduardo Morales, ''What do exporters know?,'' *The Quarterly Journal of Economics*, 2018, *133* (4), 1753--1801.

Dixit, Avinash and Victor Norman, *Theory of international trade: A dual, general equilibrium approach*, Cambridge University Press, 1980.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz, "Dissecting Trade: Firms, Industries, and Export Destinations," *American Economic Review, Papers and Proceedings*, 2004, *94* (2), 150--154.

— , — , and — , "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 2011, *79* (5), 1453--1498.

Egger, Peter and Ruobing Huang, *Market Versus Optimum Diversity in Open Economies: Theory and Quantitative Evidence*, Centre for Economic Policy Research, 2023.

Egger, Peter H, Katharina Erhardt, and Sergey Nigai, "Empirical productivity distributions and international trade," 2023.

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. and David Weinstein, "Globalization, Markups, and the U.S. Price Level," *NBER working paper*, 2010, *15749.*

Fernandes, Ana M, Peter J Klenow, Sergii Meleshchuk, Martha Denisse Pierola, and Andrés Rodríguez-Clare, "The intensive margin in trade: How big and how important?," *American Economic Journal: Macroeconomics*, 2023, *15* (3), 320--354.

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 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.

Heckman, James J and Bo E Honore, "The empirical content of the Roy model," *Econometrica: Journal of the Econometric Society*, 1990, pp. 1121--1149.

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

Hsieh, Chang-Tai and Benjamin A Olken, "The missing "missing middle"," *Journal of Economic Perspectives*, 2014, *28* (3), 89--108.

— and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, *124* (4), 1403--1448.

— , Nicholas Li, Ralph Ossa, and Mu-Jeung Yang, "Accounting for the new gains from trade liberalization," *Journal of International Economics*, 2020, *127*, 103370.

Khandelwal, Amit, "The long and short (of) quality ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450--1476.

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

Kugler, Maurice and Eric Verhoogen, "The Quality-Complementarity Hypothesis: Theory and Evidence from Colombia," 2008. Manuscript, Columbia University.

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, "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.

Morales, Eduardo, Gloria Sheu, and Andrés Zahler, "Extended gravity," *The Review of economic studies*, 2019, *86* (6), 2668--2712.

Negishi, Takashi, "Welfare economics and existence of an equilibrium for a competitive economy," *Metroeconomica*, 1960, *12* (2-3), 92--97.

Newey, Whitney K and Daniel McFadden, "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 1994, *4*, 2111--2245.

_ 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.

Pavcnik, Nina, "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, 2002, *69* (1), 245--276.

Perla, Jesse, Christopher Tonetti, and Michael E Waugh, "Equilibrium technology diffusion, trade, and growth," *American Economic Review*, 2021, *111* (1), 73--128.

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

Redding, Stephen J and David E Weinstein, "Measuring aggregate price indices with taste shocks: Theory and evidence for CES preferences," *The Quarterly Journal of Economics*, 2020, *135* (1), 503--560.

__ **and** __ , "Accounting for trade patterns," *Journal of International Economics*, 2024, *150*, 103910.

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

Simonovska, Ina and Michael E Waugh, "The elasticity of trade: Estimates and evidence," *Journal of International Economics*, 2014, *92* (1), 34--50.

Tintelnot, Felix, "Global production with export platforms," *The Quarterly Journal of Economics*, 2017, *132* (1), 157--209.

Trefler, Daniel, "The Long and Short of the Canada-US Free Trade Agreement," *American Economic Review*, 2004, *94* (4), 870--895.

Waugh, Michael E, "International trade and income differences," *American Economic Review*, 2010, *100* (5), 2093--2124.

Zhelobodko, Evgeny, Sergey Kokovin, Mathieu Parenti, and Jacques-François Thisse, "Monopolistic Competition in General Equilibrium: Beyond the CES," *mimeo*, 2011.

Online Appendix

A Theory Appendix: Proofs and Additional Results

A.1 Proofs

A.1.1 Proof of Part b of Proposition 2

Equilibrium Efficiency. To prove the efficiency of the equilibrium, we show that we can find positive weights for the social planner problem so that its outcomes are the same as the competitive equilibrium.

Denote $v_i(\omega) = \{a_i(\omega), b_{ij}(\omega), \tau_{ij}(\omega), f_{ij}(\omega)\}_j$ with distribution $v_i(\omega) \sim \bar{H}_i(v)$. The Planner's problem is

$$\max_{\{q_{ij}(v), D_{ij}(v), N_i^p\}} \sum_j \chi_j \left[\sum_i N_i^p \int (\bar{b}_{ij} b_{ij})^{\frac{1}{\sigma}} (q_{ij}(v))^{\frac{\sigma-1}{\sigma}} D_{ij}(v) d\bar{H}_i(v) \right]$$

subject to

$$N_i^p \bar{F}_i + \sum_j N_i^p \int \left(\frac{\bar{\tau}_{ij}}{\bar{a}_i} \frac{\tau_{ij}}{a_i} q_{ij}(v) + \bar{f}_{ij} f_{ij} \right) D_{ij}(v) d\bar{H}_i(v) = \overline{L}_i$$

$$D_{ij}(v) \in \{0, 1\}$$

We use the definitions in (5) and (7) to re-write the problem in terms of revenue and entry potentials: by defining $\tilde{q}_{ij} = (\bar{b}_{ij} b_{ij})^{\frac{1}{\sigma-1}} q_{ij}$,

$$\max_{\{\tilde{q}_{ij}(r,e), D_{ij}(r,e), N_i^p\}} \sum_j \chi_j \left[\sum_i N_i^p \int (\tilde{q}_{ij}(r,e))^{\frac{\sigma-1}{\sigma}} D_{ij}(r,e) dH_{ij}(r,e) \right]$$

subject to

$$N_i^p \bar{F}_i + \sum_j N_i^p \int \left(\frac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} \tilde{q}_{ij}(r,e) + \bar{f}_{ij} \frac{r}{e} \right) D_{ij}(r,e) dH_{ij}(r,e) = \overline{L}_i$$

$$D_{ij}(v) \in \{0, 1\}$$

Thus, the Lagrangean is

$$\begin{aligned} \mathcal{L} \;=\; & \sum_j \chi_j \left[\sum_i N_i^p \int (\tilde{q}_{ij}(r,e))^{\frac{\sigma-1}{\sigma}} D_{ij}(r,e) dH_{ij}(r,e) \right] \\ + \; & \sum_i \lambda_i^p \left[\overline{L}_i - N_i^p \bar{F}_i - \sum_j N_i^p \int \left(\tfrac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} \tilde{q}_{ij}(r,e) + \bar{f}_{ij} \tfrac{r}{e} \right) D_{ij}(r,e) dH_{ij}(r,e) \right] \end{aligned}$$

The first-order conditions of the problem imply that any solution must satisfy

$$\tilde{q}_{ij}(r,e) = (\bar{r}_{ij} r)^{\frac{\sigma}{\sigma-1}} \left(\frac{\lambda_i^p}{\chi_j} \right)^{-\sigma} \tag{OA.1}$$

$$D_{ij}(r,e) = 1 \Leftrightarrow \chi_j (\tilde{q}_{ij}(r,e))^{\frac{\sigma-1}{\sigma}} \geq \lambda_i^p \left(\frac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} \tilde{q}_{ij}(r,e) + \bar{f}_{ij} \frac{r}{e} \right) \tag{OA.2}$$

$$\sum_j \chi_j \int (\tilde{q}_{ij}(r,e))^{\frac{\sigma-1}{\sigma}} D_{ij}(r,e) dH_{ij}(r,e) = \lambda_i^p \left[\bar{F}_i + \sum_j \int \left(\frac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} \tilde{q}_{ij}(r,e) + \bar{f}_{ij} \frac{r}{e} \right) D_{ij}(r,e) dH_{ij}(r,e) \right] \tag{OA.3}$$

$$N_i^p \bar{F}_i + \sum_j N_i^p \int \left(\frac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} \tilde{q}_{ij}(r,e) + \bar{f}_{ij} \frac{r}{e} \right) D_{ij}(r,e) dH_{ij}(r,e) = \overline{L}_i \tag{OA.4}$$

Substituting (OA.1) into (OA.2),

$$D_{ij}(r,e) = 1 \Leftrightarrow e \geq e_{ij}^p \equiv \frac{\sigma \bar{f}_{ij}}{\bar{r}_{ij}} \left(\frac{\lambda_i^p}{\chi_j} \right)^{\sigma}. \tag{OA.5}$$

Substituting (OA.1) and (OA.5) into (OA.3),

$$\sum_j \frac{1}{\sigma} \bar{r}_{ij} \left((\lambda_i^p)^{1-\sigma} (\chi_j)^{\sigma} \right) \int_{e_{ij}^p}^{\infty} r dH_{ij}(r,e) = \lambda_i^p \left[\bar{F}_i + \sum_j \bar{f}_{ij} \int_{e_{ij}^p}^{\infty} \frac{r}{e} dH_{ij}(r,e) \right] \tag{OA.6}$$

$$\Rightarrow \overline{L}_i - N_i^p \bar{F}_i = \sum_j N_i^p \int \left(\frac{\sigma-1}{\sigma} (\bar{r}_{ij} r)^{\frac{1}{1-\sigma}} (\bar{r}_{ij} r)^{\frac{\sigma}{\sigma-1}} \left(\frac{\lambda_i^p}{\chi_j} \right)^{-\sigma} + \bar{f}_{ij} \frac{r}{e} \right) D_{ij}(r,e) dH_{ij}(r,e)$$

Substituting (OA.1), (OA.5) and (OA.3) into (OA.4)

$$\sum_j \bar{r}_{ij} \left((\lambda_i^p)^{1-\sigma} (\chi_j)^{\sigma} \right) \int_{e_{ij}^p}^{\infty} r dH_{ij}(r,e) = \frac{\lambda_i^p \overline{L_i}}{N_i^p}. \tag{OA.7}$$

Using the change of variable $n = 1 - H_{ij}^e(e)$ in (OA.6)–(OA.7),

$$\epsilon_{ij}(n_{ij}^p) = \frac{\sigma \bar{f}_{ij}}{\bar{r}_{ij}} \frac{(\lambda_i^p)^{\sigma}}{(\chi_j)^{\sigma}} \tag{OA.8}$$

$$\sum_j \frac{1}{\sigma} n_{ij}^p \bar{x}_{ij}^p = \lambda_i^p \left[\bar{F}_i + \sum_j \bar{f}_{ij} \int_0^{n_{ij}^p} \frac{\rho_{ij}(n)}{\epsilon_{ij}(n)} (1 + \varrho_{ij}(n)) dn \right] \tag{OA.9}$$

$$\sum_j N_i^p n_{ij}^p \bar{x}_{ij}^p = \lambda_i^p \overline{L}_i \tag{OA.10}$$

with

$$\bar{x}_{ij}^p \equiv \bar{r}_{ij} \left((\lambda_i^p)^{1-\sigma} (\chi_j)^{\sigma} \right) \rho_{ij}(n_{ij}^p). \tag{OA.11}$$

Thus, given a set of positive weights $\{\chi_j\}$, the system (OA.8)–(OA.11) must be solved by any efficient allocation with firm export share n_{ij}^p, average firm exports $\bar{x}_{ij}^p$, mass of firms N_i^p, and multipliers λ_i^p. We note that, if we set the weight to be equal to the destination shifter of trade flows, $(\chi_j)^{\sigma} = P_j^{\sigma-1} E_j$, the system above becomes

$$\epsilon_{ij}(n_{ij}^p) = \frac{\sigma \bar{f}_{ij}}{\bar{r}_{ij}} \frac{(\lambda_i^p)^{\sigma}}{P_j^{\sigma-1} E_j} \tag{OA.12}$$

$$\sum_j \frac{1}{\sigma} n_{ij}^p \bar{x}_{ij}^p = \lambda_i^p \left[\bar{F}_i + \sum_j \bar{f}_{ij} \int_0^{n_{ij}^p} \frac{\rho_{ij}(n)}{\epsilon_{ij}(n)} (1 + \varrho_{ij}(n)) dn \right] \tag{OA.13}$$

$$\sum_j N_i^p n_{ij}^p \bar{x}_{ij}^p = \lambda_i^p \overline{L_i} \tag{OA.14}$$

$$\bar{x}_{ij}^p \equiv \bar{r}_{ij}\left((\lambda_i^p)^{1-\sigma} P_j^{\sigma-1} E_j\right) \rho_{ij}(n_{ij}^p). \tag{OA.15}$$

Given the equilibrium conditions in Proposition 2, one solution of the system in (OA.12)–(OA.15) is $\lambda_i^p = w_i$, $N_i^p = N_i$, $n_{ij}^p = n_{ij}$ and $\bar{x}_{ij}^p = x_{ij}$, where the efficient set of varieties from i available in j determined by (OA.5) is identical to the equilibrium set given by (6). This implies that the equilibrium is a solution of the planner's problem for $\chi_j = (P_j^{\sigma-1} E_j)^{1/\sigma}$ and thus it is efficient.

A.1.2 Proof of Section 3.1

Proof of Proposition 3

Part a: Small shocks. We start by totally differentiating the equilibrium equations for the extensive and intensive margins of firm-level exports. We simplify the notation by defining $\varepsilon_{ij}^0 \equiv \varepsilon_{ij}(n_{ij}^0)$, $\varrho_{ij}^0 \equiv \varrho_{ij}(n_{ij}^0)$, and $\theta_{ij}^0 \equiv \theta_{ij}(n_{ij}^0)$. Equations (11) and (13) respectively imply that

$$\varepsilon_{ij}^0 d\ln n_{ij} = d\ln \bar{f}_{ij} - d\ln \bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1) d\ln P_j - d\ln E_j \tag{OA.16}$$

$$d\ln \bar{x}_{ij} = \varrho_{ij}^0 d\ln n_{ij} + d\ln \bar{r}_{ij} + (1-\sigma) d\ln w_i + (\sigma-1) d\ln P_j + d\ln E_j. \tag{OA.17}$$

We can then use these equations to obtain an expression for the change in bilateral trade flows, $d\ln X_{ij} = d\ln n_{ij} + d\ln \bar{x}_{ij} + d\ln N_i$:

$$d\ln X_{ij} = \theta_{ij}^0 d\ln \bar{r}_{ij} + (1-\theta_{ij}^0) d\ln \bar{f}_{ij} + (1-\theta_{ij}^0\sigma) d\ln w_i + \theta_{ij}^0(\sigma-1) d\ln P_j + \theta_{ij}^0 d\ln E_j + d\ln N_i, \tag{OA.18}$$

where we use the definition of θ_{ij}^0 in (15).

We now turn to the free entry condition in (16). When combined with the labor market clearing condition in (20), (16) is equivalent to

$$\frac{1}{N_i} = \frac{\sigma \bar{F}_i}{\bar{L}_i} + \sum_j \frac{\sigma \bar{f}_{ij}}{\bar{L}_i} \int_0^{n_{ij}} \frac{\rho_{ij}(n')}{\epsilon_{ij}(n')} (1 + \varrho_{ij}(n')) dn'. \tag{OA.19}$$

By totally differentiating this expression, we get that

$$\begin{aligned} -\tfrac{1}{N_i^0} d\ln N_i \quad &= \quad \tfrac{\sigma \bar{F}_i^0}{\bar{L}_i^0} d\ln(\bar{F}_i/\bar{L}_i) + \textstyle\sum_j \tfrac{\sigma \bar{f}_{ij}^0}{\bar{L}_i^0} \left[\int_0^{n_{ij}^0} \tfrac{\rho_{ij}(n')}{\epsilon_{ij}(n')}(1+\varrho_{ij}(n'))dn'\right] d\ln(\bar{f}_{ij}/\bar{L}_i) \\ &+ \quad \textstyle\sum_j \tfrac{\sigma \bar{f}_{ij}^0}{\bar{L}_i^0} \tfrac{n_{ij}^0 \rho_{ij}(n_{ij}^0)}{\epsilon_{ij}(n_{ij}^0)} (1+\varrho_{ij}^0) d\ln n_{ij}. \end{aligned}$$

Note that, by adding (11) and (13), $\sigma \bar{f}_{ij} w_i = \bar{x}_{ij}\epsilon_{ij}(n_{ij})/\rho_{ij}(n_{ij}) = X_{ij}\epsilon_{ij}(n_{ij})/N_i n_{ij}\rho_{ij}(n_{ij})$. Thus, the expression above can be written as

$$-d\ln N_i = (1 - \sum_j y_{ij}^0 \gamma_{ij}(n_{ij}^0)) d\ln(\bar{F}_i/\bar{L}_i) + \sum_j y_{ij}^0 \gamma_{ij}(n_{ij}^0) d\ln(\bar{f}_{ij}/\bar{L}_i) + \sum_j y_{ij}^0 (1+\varrho_{ij}^0) d\ln n_{ij} \tag{OA.20}$$

with $y_{ij} \equiv X_{ij}/\sum_{j'} X_{ij'}$ the share of sales to destination j in the output of origin i, and

$$\gamma_{ij}(n) \equiv \frac{\epsilon_{ij}(n)}{n\rho_{ij}(n)} \int_0^n \frac{\rho_{ij}(n')}{\epsilon_{ij}(n')}(1+\varrho_{ij}(n'))dn. \tag{OA.21}$$

Finally, using the definition of θ^0_{ij} in (15), we get that

$$-d\ln N_i = (1-\sum_j y^0_{ij}\gamma_{ij}(n^0_{ij}))d\ln(\bar{F}_i/\bar{L}_i) + \sum_j y^0_{ij}\gamma_{ij}(n^0_{ij})d\ln(\bar{f}_{ij}/\bar{L}_i) + \sum_j y^0_{ij}(1-\theta^0_{ij})\varepsilon_{ij} d\ln n_{ij} \tag{OA.22}$$

which in combination with OA.16 implies that

$$\begin{aligned} d\ln N_i \quad = & \quad d\ln(\bar{L}_i) - (1-\textstyle\sum_j y^0_{ij}\gamma_{ij}(n^0_{ij}))d\ln\bar{F}_i - \textstyle\sum_j y^0_{ij}\gamma_{ij}(n^0_{ij})d\ln\bar{f}_{ij} \\ - & \quad \textstyle\sum_j y^0_{ij}(1-\theta^0_{ij})\left(d\ln\bar{f}_{ij} - d\ln\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - d\ln E_j\right). \end{aligned} \tag{OA.23}$$

The budget balance condition in (19) implies that

$$d\ln E_j = \iota^0_j(d\ln w_j + d\ln\bar{L}_j) + (1-\iota^0_j)d\ln\bar{T}_j = \sum_i x^0_{ij} d\ln X_{ij} \tag{OA.24}$$

with $\iota_j = \sum_{j'} X_{jj'}/\sum_{j'} X_{j'j}$ the income-to-spending ratio and x_{ij} the share of origin i in destination j's spending.

The labor market clearing condition in (20) implies that

$$\sum_j y^0_{ij} d\ln X_{ij} = d\ln w_i + d\ln\bar{L}_i. \tag{OA.25}$$

The system of equations (OA.18), (OA.23), (OA.24) and (OA.25) determines $\{d\ln X_{ij}, d\ln P_i, d\ln N_i, d\ln w_i\}$ as a function of shocks in exogenous fundamentals, $\{d\ln\bar{r}_{ij}, d\ln\bar{f}_{ij}, d\ln\bar{L}_i, d\ln\bar{T}_i, d\ln\bar{F}_i\}$. To establish the proposition, consider the special case of this system for shocks in bilateral revenue shifters:

$$d\ln X_{ij} = \theta^0_{ij} d\ln\bar{r}_{ij} + (1-\theta^0_{ij}\sigma)d\ln w_i + \theta^0_{ij}(\sigma-1)d\ln P_j + \theta^0_{ij}\iota^0_j d\ln w_j + d\ln N_i, \tag{OA.26}$$

$$d\ln N_i \quad = \quad \textstyle\sum_j y^0_{ij}(\theta^0_{ij}-1)\left(-d\ln\bar{r}_{ij} + \sigma d\ln w_i - (\sigma-1)d\ln P_j - \iota^0_j d\ln w_j\right)\ , \tag{OA.27}$$

$$\iota^0_j d\ln w_j = \sum_i x^0_{ij} d\ln X_{ij}, \tag{OA.28}$$

$$\sum_j y^0_{ij} d\ln X_{ij} = d\ln w_i. \tag{OA.29}$$

The proposition follows from the observation that, given any shock $\{d\ln\bar{r}_{ij}\}$, the system (OA.26)–(OA.29) can be solved only with knowledge of the (i) the demand elasticity of substitution σ, (ii) the bilateral trade matrix at the initial equilibrium $\{X^0_{ij}\}$ (since it implies $\{y^0_{ij}, x^0_{ij}, \iota^0_j\}$ by definition), and (iii) the bilateral trade elasticity matrix at the initial equilibrium $\{\theta^0_{ij}\}$.

To further establish the expression in (22), we note that

$$
\begin{array}{rcl}
d\ln w_i & = & \sum_j y_{ij}^0 \left(d\ln N_i + d\ln n_{ij} + d\ln \bar{x}_{ij}\right) \\
d\ln w_i & = & \sum_j y_{ij}^0 \left(d\ln N_i + d\ln n_{ij} + d\ln w_i + (\varrho_{ij}^0 - \varepsilon_{ij}^0) d\ln n_{ij}\right) \\
d\ln N_i & = & -\sum_j y_{ij}^0 \left((1+\varrho_{ij}^0)/\varepsilon_{ij}^0 - 1\right) \varepsilon_{ij}^0 d\ln n_{ij} \\
d\ln N_i & = & \sum_j y_{ij}^0 \theta_{ij}^0 \varepsilon_{ij}^0 d\ln n_{ij}
\end{array}
$$

where the first equality follows from (OA.29), the second equality follows from $\sigma \bar{f}_{ij} w_i = \bar{x}_{ij}\epsilon_{ij}(n_{ij})/\rho_{ij}(n_{ij})$ (as implied by the sum of (11) and (13)), the third equality from $\sum_j y_{ij}^0 = 1$, and the last equality from the definition of θ_{ij}^0 in (15). This expression implies that

$$\sum_j y_{ij}^0 \varepsilon_{ij}^0 d\ln n_{ij} = 0, \tag{OA.30}$$

since equation (OA.27) is equivalent to

$$d\ln N_i = \sum_j y_{ij}^0 (\theta_{ij}^0 - 1)\varepsilon_{ij}^0 d\ln n_{ij}. \tag{OA.31}$$

Thus, (OA.30) implies that

$$d\log n_{ii} = -\sum_{j\neq i} \frac{y_{ij}^0 \varepsilon_{ij}^0}{y_{ii}^0 \varepsilon_{ii}^0} d\ln n_{ij}$$

By substituting this expression into (22),

$$d\ln N_i = \sum_{j\neq i} (\theta_{ij}^0 - \theta_{ii}^0) y_{ij}^0 \varepsilon_{ij}^0 d\log n_{ij}. \tag{OA.32}$$

Part b: Large shocks. Let a variable with a "hat" ($\hat{y}_i \equiv y_i'/y_i^0$) denote the ratio between that variable at the initial equilibrium, y_i^0, and the counterfactual equilibrium, y_i'. We now characterize the system that determines changes in equilibrium outcomes for any arbitrary change in fundamentals, $\{\hat{\bar{r}}_{ij}, \hat{\bar{f}}_{ij}, \hat{\bar{L}}_i, \hat{\bar{T}}_i, \hat{\bar{F}}_i\}$. Equations (11) and (13) respectively imply that

$$\frac{\epsilon_{ij}(n_{ij}^0 \hat{n}_{ij})}{\epsilon_{ij}(n_{ij}^0)} = \frac{\hat{\bar{f}}_{ij}}{\hat{\bar{r}}_{ij}} \frac{\hat{w}_i^\sigma}{\hat{P}_j^{\sigma-1}\hat{E}_j}, \tag{OA.33}$$

$$\hat{\bar{x}}_{ij} = \frac{\rho_{ij}(n_{ij}^0 \hat{n}_{ij})}{\rho_{ij}(n_{ij}^0)} \hat{\bar{r}}_{ij} \hat{w}_i^{1-\sigma} \hat{P}_j^{\sigma-1} \hat{E}_j. \tag{OA.34}$$

By definition, changes in bilateral trade flows are given by

$$\hat{X}_{ij} = \hat{N}_i \hat{n}_{ij} \hat{\bar{x}}_{ij}. \tag{OA.35}$$

Using the fact that $\sigma \bar{f}_{ij} w_i = \bar{x}_{ij}\epsilon_{ij}(n_{ij})/\rho_{ij}(n_{ij})$ and the definition of $\gamma_{ij}(n)$ in (OA.21), the version of the free entry condition in (OA.19) is equivalent to

$$\frac{1}{N_i} = \frac{\sigma \bar{F}_i}{\bar{L}_i} + \sum_j \frac{n_{ij}\bar{x}_{ij}}{w_i \bar{L}_i} \gamma_{ij}(n_{ij}),$$

which, in combination with the fact that $\sigma N_i^0 \bar{F}_i^0 / \bar{L}_i^0 = 1 - \sum_j y_{ij}^0 \gamma_{ij}(n_{ij}^0)$, implies that

$$\frac{1}{\hat{N}_i} = (1 - \sum_j y_{ij}^0 \gamma_{ij}(n_{ij}^0)) \frac{\hat{\bar{F}}_i}{\hat{\bar{L}}_i} + \sum_j y_{ij}^0 \frac{\hat{n}_{ij} \hat{\bar{x}}_{ij}}{\hat{w}_i \hat{\bar{L}}_i} \gamma_{ij}(n_{ij} \hat{n}_{ij}). \tag{OA.36}$$

Finally, the equations for budget balance in (19) and market clearing in (20) immediately imply that

$$\hat{E}_j = \sum_i x_{ij}^0 \hat{X}_{ij} = \iota_j^0 \hat{w}_j \hat{\bar{L}}_j + (1 - \iota_j^0) \hat{\bar{T}}_j, \tag{OA.37}$$

$$\sum_j y_{ij}^0 \hat{X}_{ij} = \hat{w}_i \hat{\bar{L}}_i. \tag{OA.38}$$

Part b of the proposition follows from the fact that, for any shock in fundamentals $\{\hat{\bar{r}}_{ij}, \hat{\bar{f}}_{ij}, \hat{\bar{L}}_i, \hat{\bar{T}}_i, \hat{\bar{F}}_i\}$, counterfactual changes $\{\hat{n}_{ij}, \hat{\bar{x}}_{ij}, \hat{X}_{ij}, \hat{P}_i, \hat{N}_i, \hat{w}_i\}$ are given by the solution of the system (OA.33)–(OA.38), which depends on the (i) the demand elasticity of substitution σ, (ii) the elasticity functions $\rho_{ij}(n)$ and $\epsilon_{ij}(n)$ (since they imply $\gamma_{ij}(n)$ by definition), and (iii) the exporter firm share and bilateral trade matrices at the initial equilibrium $\{n_{ij}^0, X_{ij}^0\}$ (since $\{X_{ij}^0\}$ implies $\{y_{ij}^0, x_{ij}^0, \iota_j^0\}$ by definition).

A.1.3 Proofs for Section 3.2

Proof of Equation (27). Consider small changes in revenue shifters under trade balance ($\sum_i X_{ij}^0 = \sum_i X_{ji}^0$ for all j). We now derive an expression for the change in the average real wage across countries, weighted by their initial spending:

$$d \ln W \equiv \sum_j \frac{E_j^0}{E^0} d \ln \frac{w_j}{P_j}$$

with $E^0 \equiv \sum_j E_j^0$.

Note that, up to a first-order approximation, $\sum_i x_{ij}^0 \ln \hat{x}_{ij} = \sum_i x_{ij}^0 d \ln x_{ij} \approx \sum_i dx_{ij} = 0$. Thus, up to a first-order approximation, equation (26) becomes

$$d \ln \frac{w_j}{P_j} \approx \sum_i \frac{x_{ij}^0}{\sigma - 1} d \ln \bar{r}_{ij} + \sum_i \frac{x_{ij}^0}{\sigma - 1} d \ln \frac{w_j}{w_i} + \sum_i \frac{x_{ij}^0}{\sigma - 1} d \ln N_i + \sum_i \frac{x_{ij}^0}{\sigma - 1} (1 + \varrho_{ij}^0) d \ln n_{ij}. \tag{OA.39}$$

Thus,

$$d \ln W \approx \sum_{i,j} \frac{X_{ij}^0}{E^0} \frac{d \ln \bar{r}_{ij}}{\sigma - 1} + \sum_{i,j} \frac{X_{ij}^0}{E^0} d \ln(w_j / w_i) + \frac{1}{\sigma - 1} \sum_{i,j} \frac{X_{ij}^0}{E^0} \left(d \ln N_i + (1 + \varrho_{ij}^0) d \ln n_{ij} \right). \tag{OA.40}$$

We now establish that the second and third terms in this expression are equal to zero. Consider the second term:

$$\begin{aligned} \textstyle\sum_{i,j} X_{ij}^0 d \ln(w_j / w_i) &= \textstyle\sum_j \left(\sum_i X_{ij}^0 \right) d \ln w_j - \sum_i \left(\sum_j X_{ij}^0 \right) d \ln w_i \\ &= \textstyle\sum_j \left(\sum_i X_{ij}^0 \right) d \ln w_j - \sum_j \left(\sum_i X_{ji}^0 \right) d \ln w_j \\ &= \textstyle\sum_j \left(\sum_i X_{ij}^0 - \sum_i X_{ji}^0 \right) d \ln w_j \\ &= 0 \end{aligned}$$

where the last equality follows from trade balance.

Turning to the third term, note that

$$\begin{aligned}\sum_{i,j} X_{ij}^0 \left(d\ln N_i + (1+\varrho_{ij}^0) d\ln n_{ij}\right) &= \sum_i (\sum_j X_{ij}^0) d\ln N_i - \sum_i \left(\sum_j X_{ij}^0 (\theta_{ij}^0 - 1)\varepsilon_{ij}^0 d\ln n_{ij}\right) \\ &= \sum_i (\sum_j X_{ij}^0) d\ln N_i - \sum_i (\sum_j X_{ij}^0) d\ln N_i \\ &= 0\end{aligned}$$

where the first equality uses the definition of θ_{ij}^0 in (15), and the second equality uses (OA.31).

Thus,

$$d\ln W \approx \sum_{i,j} \frac{X_{ij}^0}{E^0} \frac{d\ln \bar{r}_{ij}}{\sigma - 1}.$$

Special Case of Symmetric Countries. Consider small changes in trade costs in a world economy with symmetric countries such that

$$E_i^0 = E^0, \quad X_{ij}^0 = X_{ji}^0, \quad \theta_{ij}^0 = \theta_{ji}^0, \quad d\ln N_i = d\ln N, \quad d\ln \epsilon_{ij}(n_{ij}) = d\ln \epsilon_{ji}(n_{ji}). \tag{OA.41}$$

From equation (OA.27),

$$d\ln N = d\ln N_i = \frac{1}{E^0} \sum_j X_{ij}^0 (\theta_{ij}^0 - 1) d\ln \epsilon_{ij}(n_{ij}). \tag{OA.42}$$

This implies that the firm entry and firm selection terms cancel each other for every country:

$$\begin{aligned}\tfrac{1}{\sigma-1}\sum_i x_{ij}^0 \left[d\ln N_i + (1+\varrho_{ij}^0) d\ln n_{ij}\right] &= \tfrac{1}{\sigma-1}\sum_i \tfrac{X_{ij}^0}{E_j^0}\left[d\ln N_i - (\theta_{ij}^0 - 1) d\ln \epsilon_{ij}(n_{ij})\right] \\ &= \tfrac{1}{\sigma-1}\left[d\ln N - \tfrac{1}{E^0}\sum_i X_{ji}^0 (\theta_{ji}^0 - 1) d\ln \epsilon_{ji}(n_{ji})\right] \\ &= \tfrac{1}{\sigma-1}\left[d\ln N - d\ln N\right] \\ &= 0\end{aligned}$$

where the first equality uses the definition of θ_{ij}^0 in (15), the second equality uses the symmetry assumption in (OA.41), and the third equality uses (OA.42).

Note that, in this case, the terms of trade term is also equal to zero, since $d\ln w_j = d\ln w$ for all j. Thus, the first-order approximation for welfare in (OA.39) only contains the technology term.

Constant-Elasticity Benchmark. Consider small changes in trade costs under trade balance. We assume that the economy is given by the constant-elasticity benchmark in (24) with identical elasticities in all countries:

$$\bar{\varrho}_{ij} = \bar{\varepsilon}_{ij} = -1/\theta,$$

and thus $\bar{\theta}_{ij} = \theta$.

The resource constraint in (OA.30) implies that $\sum_j y_{ij}^0 d\ln n_{ij} = 0$. Thus, the free entry condition in (OA.31) implies that

$$d\ln N_i = \sum_j y_{ij}^0 \bar{\theta}_{ij} \bar{\varepsilon}_{ij} d\ln n_{ij} = -\sum_j y_{ij}^0 d\ln n_{ij} = 0.$$

From (OA.18) and (OA.28), we obtain the following expression for the price index:

$$\theta(\sigma-1)d\ln P_j = (1-\theta)d\ln w_j - \sum_i x_{ij}^0\left(\theta d\ln\bar{r}_{ij} + (1-\theta\sigma)d\ln w_i\right) \tag{OA.43}$$

Using this expression, we can then use (OA.18) to re-write the market clearing condition in (OA.29) to obtain the following system of equations determining wages:

$$\theta\sigma d\ln w_i - \sum_j\left[y_{ij}^0 + (\theta\sigma-1)\sum_d y_{id}^0 x_{jd}^0\right]d\ln w_j = \sum_j y_{ij}^0\left(\theta d\ln\bar{r}_{ij} - \sum_o x_{oj}^0\theta d\ln\bar{r}_{oj}\right). \tag{OA.44}$$

Constant-Elasticity Benchmark with Two Countries. We now focus on the special case with two countries, Home ($i=H$) and Foreign ($i=F$), where $d\ln\bar{r}_{HH} = d\ln\bar{r}_{FF} = 0$. We define Foreign's wage as the numeraire, $d\ln w_F = 0$, and denote Home's wage change as $d\ln w_H = d\ln w$. Home's labor market clearing condition determines the equilibrium change in relative wages:

$$d\ln w = \frac{-y_{HH}^0 x_{FH}^0\theta d\ln\bar{r}_{FH} + y_{HF}^0\left(1-x_{HF}^0\right)\theta d\ln\bar{r}_{HF}}{\left(\theta\sigma - y_{HH}^0 - (\theta\sigma-1)\sum_{d=H,F} y_{Hd}^0 x_{Hd}^0\right)}. \tag{OA.45}$$

Using (OA.43), we solve for Foreign's price index change,

$$\theta(\sigma-1)d\ln P_F = -x_{HF}^0\left(\theta d\ln\bar{r}_{HF} + (1-\theta\sigma)d\ln w\right),$$

which we plug into the extensive margin expression in (OA.16) to characterize firm selection:

$$d\ln n_{FF} = -x_{HF}^0\left(\theta d\ln\bar{r}_{HF} - \theta\sigma d\ln w\right) - x_{HF}^0 d\ln w$$

$$d\ln n_{HF} = x_{FF}^0(\theta d\ln\bar{r}_{HF} - \theta\sigma d\ln w) - x_{HF}^0 d\ln w$$

Thus, up to a first-order approximation, the decomposition in (26) becomes

$$d\ln\frac{w_F}{P_F} = \underbrace{\left(\frac{x_{HF}^0}{\sigma-1}d\ln\bar{r}_{HF}\right)}_{\text{Technology}} + \underbrace{\left(-x_{HF}^0 d\ln w\right)}_{\text{Terms of trade}} + \underbrace{0}_{\text{Firm entry}} + \underbrace{\left(\frac{1-\theta}{\theta(\sigma-1)}x_{HF}^0 d\ln w\right)}_{\text{Firm selection}}. \tag{OA.46}$$

In combination with (OA.45), (OA.46) implies that, when countries are asymmetric, responses in terms of trade and firm selection have first-order impacts on welfare. Note however that, when countries are symmetric as defined in (OA.41), we have that $d\ln w = 0$ and thus both terms are second-order.

A.1.4 Proofs of Section 3.3

Proof of Corollary 1. We consider a counterfactual exercise in which an economy without international transfers moves to autarky. Specifically, we assume that $\hat{\bar{r}}_{ij} \to 0$ for all $i \neq j$, that $\iota_i^0 = 1$ for all i, and that $\hat{\bar{F}}_i = \hat{\bar{f}}_{ij} = \hat{\bar{L}}_i = \hat{\bar{r}}_{ii} = 1$ for all i and j. We set the wage of country j to be the numerarie, $w_j \equiv 1$, so that $\hat{w}_j = 1$ and $\hat{E}_j = 1$.

By noticing that $\hat{x}_{jj} = 1/x^0_{jj}$, equation (25) implies that

$$\hat{P}_j^{1-\sigma} = x^0_{jj}\hat{N}_j\hat{n}_{jj}\frac{\rho_{jj}(n^0_{jj}\hat{n}_{jj})}{\rho_{jj}(n^0_{jj})}. \tag{OA.47}$$

We then use (OA.33) to substitute for $\hat{P}_j^{1-\sigma}$:

$$\frac{\epsilon_{jj}(n^0_{jj}\hat{n}_{jj})}{\epsilon_{jj}(n^0_{jj})} = x^0_{jj}\hat{n}_{jj}\hat{N}_j\frac{\rho_{jj}(n^0_{jj}\hat{n}_{jj})}{\rho_{jj}(n^0_{jj})}. \tag{OA.48}$$

Finally, since $\hat{N}_i\hat{n}_{ij}\hat{\bar{x}}_{ij} = 0$ for all $i \neq j$ and $y^0_{jj}\hat{y}_{jj} = y^0_{jj}\hat{N}_j\hat{n}_{jj}\hat{\bar{x}}_{jj} = 1$, the free entry condition in (OA.36) becomes

$$\hat{N}_j = \frac{1-\gamma_{jj}(n^0_{jj}\hat{n}_{jj})}{1-\sum_d y^0_{jd}\gamma_{jd}(n^0_{jd})}. \tag{OA.49}$$

The system (OA.47)–(OA.49) determines $\{\hat{n}_{jj}, \hat{N}_j, \hat{P}_j\}$ with $\hat{w}_j = 1$.

Proof of Equation (31). Equation (25) implies that

$$\begin{aligned}(\sigma-1)d\ln\tfrac{w_j}{P_j} &= -d\ln x_{jj}/N_j + (1+\varrho^0_{jj})d\ln n_{jj} \\ &= -d\ln x_{jj}/N_j + (1-\theta^0_{jj})\varepsilon^0_{jj}d\ln n_{jj}\end{aligned}$$

where the second equality uses the definition of θ^0_{jj} in (15).

Now note that, under trade balance ($d\ln E_j = d\ln w_j$) and no domestic shocks ($\ln\bar{f}_{jj} = d\ln\bar{r}_{jj} = 0$), the extensive margin equation in (OA.16) implies that

$$\varepsilon^0_{jj}d\ln n_{jj} = (\sigma-1)d\ln\frac{w_j}{P_j}.$$

Equation (31) immediately follows from the two expressions above.

A.2 Extensions

This appendix presents extensions of our baseline framework. Section A.2.1 relaxes the assumption of CES demand in our baseline framework by allowing for a general class of demand functions with a single aggregator. In Section A.2.2, we extend our model to include import tariffs that generate government revenue, as well as heterogeneous firms in multiple sectors whose production function uses multiple factors and sector-specific inputs. In Section A.2.3, we relax the assumption of full support in the distribution of entry potentials to allow for zero trade flows between countries. Section A.2.4 extends our baseline framework to allow firms to produce multiple products.

A.2.1 Non-CES Demand and Variable Markups

Our baseline model considers a nonparametric distribution of firm fundamentals, while maintaining the typical parametric assumption of CES demand. We now show how our insights generalize for a class of single aggregator demand functions that allow for variable markups.

Environment

We maintain the same environment of Section 2.1, except that preferences are now given by (21) in Section 2.4. To simplify notation, we drop the components of bilateral shifters that are common to all firms, and introduce them below when deriving the expressions for the margins of firm-level exports.

Entry and Revenue Potentials. We consider a monopolistic competitive environment in which firms take w_i and D_j as given. The firm's profit maximization problem conditional on entering market j is:

$$\max_p \left(p - w_i \frac{\tau_{ij}(\omega)}{a_i(\omega)}\right) \frac{1}{b_{ij}(\omega)} q_j \left(\frac{1}{b_{ij}(\omega)} \frac{p}{D_j}\right) - w_i f_{ij}(\omega),$$

with an associated FOC of

$$q_j \left(\frac{1}{b_{ij}(\omega)} \frac{p}{D_j}\right) + \left(\frac{1}{b_{ij}(\omega)} \frac{p}{D_j} - \frac{w_i/D_j}{r_{ij}(\omega)}\right) q_j' \left(\frac{1}{b_{ij}(\omega)} \frac{p}{D_j}\right) = 0, \tag{OA.50}$$

where we define the revenue potential in j of firm ω from i as

$$r_{ij}(\omega) \equiv \frac{b_{ij}(\omega) a_i(\omega)}{\tau_{ij}(\omega)}. \tag{OA.51}$$

The equilibrium condition in (OA.50) implicitly defines the optimal price of firm ω:

$$\frac{1}{b_{ij}(\omega)} \frac{p_{ij}(\omega)}{D_j} = \mathcal{P}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right).$$

This implies that, conditional on selling in j, firm ω from i has revenue, variable cost and variable profits given by

$$\frac{R_{ij}(\omega)}{D_j} = \mathcal{R}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) \equiv \mathcal{P}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) q_j \left(\mathcal{P}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right)\right) \tag{OA.52}$$

$$\frac{C_{ij}(\omega)}{D_j} = \mathcal{C}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) \equiv \frac{w_i/D_j}{r_{ij}(\omega)} q_j \left(\mathcal{P}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right)\right) \tag{OA.53}$$

$$\frac{\Pi_{ij}(\omega)}{D_j} = \Pi_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) \equiv \mathcal{R}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) - \mathcal{C}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) \tag{OA.54}$$

We assume that the demand function in (21) implies that firms with a higher marginal cost have lower revenue and variable profit,

$$\mathcal{R}_j' < 0 \quad \text{and} \quad \Pi_j' < 0, \tag{OA.55}$$

with $\lim_{x\to 0} \mathcal{R}_j(x) = \infty$ and $\lim_{x\to\infty} \mathcal{R}_j(x) = 0$.[31]

Firm ω from i decides to sell in j if, and only if, $\Pi_{ij}(\omega) \geq w_i f_{ij}(\omega)$ which is equivalent to

$$\bar{\Pi}_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right) \equiv \frac{\Pi_j \left(\frac{w_i/D_j}{r_{ij}(\omega)}\right)}{\frac{w_i/D_j}{r_{ij}(\omega)}} > r_{ij}(\omega) f_{ij}(\omega)$$

[31]This is a mild restriction that arises from assumptions about the second derivative of the demand function.

Note that $\bar{\Pi}_j' < 0$ since $\Pi_j' < 0$. Thus,

$$\Omega_{ij} \equiv \{\omega : e_{ij}(\omega) > w_i/D_j\} \quad \text{such that} \quad e_{ij}(\omega) \equiv r_{ij}(\omega)\bar{\Pi}_j^{-1}\left(r_{ij}(\omega)f_{ij}(\omega)\right). \tag{OA.56}$$

Extensive and Intensive Margins of Firm Exports

As in Section 2.2, we consider the distribution of $(r_{ij}(\omega), e_{ij}(\omega))$ for each origin i and destination j. We now explicitly introduce shifters of entry and revenue potentials:

$$r_{ij}(\omega) \sim H_{ij}^r\left(r/\bar{r}_{ij}|e\right) \qquad \text{and} \qquad e_{ij}(\omega) \sim H_{ij}^e(e/\bar{r}_{ij}), \tag{OA.57}$$

where H_{ij}^e satisfies Assumption 1.

Extensive margin of firm-level exports. The entry condition in (OA.56) implies that $1 - n_{ij} = Pr(e_{ij}(\omega) < w_i/D_j) = H_{ij}^e(d_{ij})$ with $d_{ij} \equiv w_i/D_j\bar{r}_{ij}$. Let us define again the extensive margin elasticity function as $\epsilon_{ij}(n) \equiv (H_{ij}^e)^{-1}(1-n)$ such that $\epsilon_{ij}(n)$ is strictly decreasing, $\epsilon_{ij}(1) = 0$, and $\lim_{n\to 0}\epsilon_{ij}(n) = \infty$. Thus,

$$\ln \epsilon_{ij}(n_{ij}) = -\ln \bar{r}_{ij} + \ln w_i - \ln D_j. \tag{OA.58}$$

Intensive margin of firm-level exports. Given the profit maximization problem above, average firm exports are given by $\bar{x}_{ij} = D_j\mathbb{E}[\mathcal{R}_j(w_i/D_{ij}r_{ij}(\omega))|\omega \in \Omega_{ij}]$. The entry decision in (OA.56) implies that

$$\bar{x}_{ij} = \frac{D_j}{n_{ij}} \int_{d_{ij}}^{\infty} \mathbb{E}\left[\mathcal{R}_j(d_{ij}/r)|e\right] dH_{ij}^e(e)$$

with $\mathbb{E}\left[\mathcal{R}_j(d_{ij}/r)|e\right] \equiv \int \mathcal{R}_j(d_{ij}/r)dH_{ij}^r(r|e)$. Let us define $\tilde{\rho}_{ij}(d) \equiv \int_d^{\infty} \mathbb{E}\left[\mathcal{R}_j(d/r)|e\right] dH_{ij}^e(e)$. Since $\tilde{\rho}_{ij}'(d) < 0$, $\lim_{d\to 0}\tilde{\rho}_{ij}(d) = \infty$ and $\lim_{d\to\infty}\tilde{\rho}_{ij}(d) = 0$, $\tilde{\rho}_{ij}(d)$ is invertible and we can define $\rho_{ij}(x) \equiv \tilde{\rho}_{ij}^{-1}(x)$ such that

$$\ln \rho_{ij}(\bar{x}_{ij}n_{ij}/D_j) = -\ln \bar{r}_{ij} + \ln w_i - \ln D_j. \tag{OA.59}$$

We can now extend Proposition 1 for our setting with non-CES demand of the form in equation (21).

Proposition 1 (non-CES demand). *Consider the monopolistic competition model with non-CES demand in the environment of Appendix Section A.2.1 under (OA.57). Then, for any origin i and destination j, the exporter firm share, n_{ij}, and the average firm exports, $\bar{x}_{ij}$, are given by equations (OA.58) and (OA.59), which depend on country-level endogenous variables, exogenous bilateral shifters, and two elasticity functions of the exporter firm share $n \in [0,1]$, $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$.*

Intensive margin of firm-level exports across percentiles of the distribution of firm exports. We now characterize an expression for percentile π of the distribution of firm-level exports from i to j. We start by deriving the distribution of firm-level exports from i to j:

$$\begin{aligned}
H_{ij}^R(\bar{R}) &\equiv Pr(R_{ij}(\omega) < \bar{R} | \omega \in \Omega_{ij}) \\
&= \frac{1}{n_{ij}} Pr(\mathcal{R}_j(w_i/D_j r_{ij}(\omega)) < \bar{R}/D_j,\ e(\omega) > d_{ij}) \\
&= \frac{1}{n_{ij}} Pr(w_i/D_j r_{ij}(\omega) > \mathcal{R}_j^{-1}\left(\bar{R}/D_j\right),\ e(\omega) > d_{ij}) \\
&= \frac{1}{n_{ij}} \int_{d_{ij}}^{\infty} H_{ij}^r\left(d_{ij}/\mathcal{R}_j^{-1}\left(\bar{R}/D_j\right)|e\right) dH_{ij}^e(e) \\
&= \frac{1}{n_{ij}} \int_0^{n_{ij}} H_{ij}^r\left(\epsilon_{ij}(n_{ij})/\mathcal{R}_j^{-1}\left(\bar{R}/D_j\right)|\epsilon_{ij}(n)\right) dn
\end{aligned}$$

where the second row uses (OA.52), the third row uses $\mathcal{R}_j(.)$ invertible with $\mathcal{R}_j' < 0$, the fourth row uses (OA.57), and the last row uses change of variables $n = 1 - H_{ij}^e(e)$.

We now define $F_{ij}(R|n) \equiv \frac{1}{n}\int_0^n H_{ij}^r\left(R|\epsilon_{ij}(n')\right) dn'$. Note that $F_{ij}(.|n)$ is invertible for any n since $\partial F_{ij}(R|n)/\partial R > 0$, $F_{ij}(0|n) = 0$ and $\lim_{R\to\infty} F_{ij}(R|n) = 1$. The distribution of firm-level exports from i to j can be written as

$$H_{ij}^R(\bar{R}) = F_{ij}\left(\epsilon_{ij}(n_{ij})/\mathcal{R}_j^{-1}\left(\bar{R}/D_j\right)|n_{ij}\right).$$

We denote the revenue of firms in percentile π of the distribution of firm-level exports from i to j as x_{ij}^π, which is implicitly given by $\pi = H_{ij}^R(x_{ij}^\pi)$. Since $F_{ij}(.|n)$ is invertible for any n, we have that

$$\epsilon_{ij}(n_{ij})/\mathcal{R}_j^{-1}\left(x_{ij}^\pi/D_j\right) = F_{ij}^{-1}(\pi|n_{ij})$$

which implies that

$$x_{ij}^\pi = D_j\rho_{ij}^\pi(n_{ij}) \quad \text{with} \quad \rho_{ij}^\pi(n) \equiv \mathcal{R}_j\left(\epsilon_{ij}(n)/F_{ij}^{-1}(\pi|n)\right). \tag{OA.60}$$

We note that, by definition, when we know the functions $\rho_{ij}^\pi(n)$, $\epsilon_{ij}(n)$ and $\mathcal{R}_j(.)$, we can define the following function of π: $G_{ij}(\pi|n) \equiv \epsilon_{ij}(n)/\mathcal{R}_j^{-1}\left(\rho_{ij}^\pi(n)\right)$. Note that $\rho_{ij}^\pi(n)$ is increasing in π given n, which allows us to write $F_{ij}(R|n) = G_{ij}^{-1}(R|n)$. Thus, since $nF_{ij}(R|n) = \int_0^n H_{ij}^r\left(R|\epsilon_{ij}(n')\right) dn'$ by definition, $nG_{ij}^{-1}(R|n) = \int_0^n H_{ij}^r\left(R|\epsilon_{ij}(n')\right) dn'$ and

$$\frac{\partial\left[nG_{ij}^{-1}(R|n)\right]}{\partial n} = H_{ij}^r\left(R|\epsilon_{ij}(n)\right). \tag{OA.61}$$

Sufficient Statistics in General Equilibrium

We now outline the conditions that determine $\{w_i, D_i, N_i\}$ in general equilibrium. As in the baseline model, budget balance and labor market clearing are given by

$$\sum_i N_i n_{ij}\bar{x}_{ij} = w_j\bar{L}_j + \bar{T}_j, \tag{OA.62}$$

$$\sum_j N_i n_{ij}\bar{x}_{ij} = w_i\bar{L}_i. \tag{OA.63}$$

Thus, these two conditions can be written in terms of $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$ in equations (OA.58)-(OA.59), respectively.

We now turn to the free entry condition, which is still given by

$$\sum_j n_{ij}(\bar{x}_{ij} - \bar{c}_{ij}) = w_i\bar{F}_i, \tag{OA.64}$$

where $\bar{c}_{ij} \equiv \mathbb{E}[C_{ij}(\omega)|\omega \in \Omega_{ij}]$ is the mean cost of firms from i selling in j.

As in our baseline model, expressions (OA.58) and (OA.59) characterize n_{ij} and $\bar{x}_{ij}$ using the elasticity functions for the extensive and intensive margins of firm-level exports. Thus, it suffices to characterize the mean cost $\bar{c}_{ij}$, which can be written in terms of variable and fixed costs:

$$\bar{c}_{ij} = \mathbb{E}\left[D_j \mathcal{C}_j\left(w_i/D_j r_{ij}(\omega)\right)|\omega \in \Omega_{ij}\right] + \mathbb{E}\left[w_i f_{ij}(\omega)|\omega \in \Omega_{ij}\right]$$

Consider first the expected variable cost of firms from i operating in j:

$$\begin{aligned}
\mathbb{E}\left[D_j \mathcal{C}_j\left(\tfrac{w_i/D_j}{r_{ij}(\omega)}\right)|\omega \in \Omega_{ij}\right] &= D_j \tfrac{1}{n_{ij}} \int_{d_{ij}}^{\infty} \int \mathcal{C}_j\left(d_{ij}/r\right) dH_{ij}^r\left(r|e\right) dH_{ij}^e(e) \\
&= D_j \tfrac{1}{n_{ij}} \int_0^{n_{ij}} \int \mathcal{C}_j\left(\epsilon_{ij}(n_{ij})/r\right) dH_{ij}^r\left(r|\epsilon_{ij}(n)\right) dn
\end{aligned}$$

where the first equality uses (OA.57) and the second equality the change of variables $n = 1 - H_{ij}^e(e)$.

Turning to the mean fixed cost, we have that

$$\begin{aligned}
\mathbb{E}\left[w_i f_{ij}(\omega)|\omega \in \Omega_{ij}\right] &= w_i \mathbb{E}\left[\bar{\Pi}_j\left(e_{ij}(\omega)/r_{ij}(\omega)\right)/r_{ij}(\omega)|\omega \in \Omega_{ij}\right] \\
&= D_j d_{ij} \tfrac{1}{n_{ij}} \int_{d_{ij}}^{\infty} \int \bar{\Pi}_j\left(e/r\right)/r dH_{ij}^r\left(r|e\right) dH_{ij}^e(e) \\
&= D_j \epsilon_{ij}(n_{ij}) \tfrac{1}{n_{ij}} \int_0^{n_{ij}} \int \bar{\Pi}_j\left(\epsilon_{ij}(n)/r\right)/r dH_{ij}^r\left(r|\epsilon_{ij}(n)\right) dn
\end{aligned}$$

where the first equality uses (OA.56), the second equality uses (OA.57) and the third equality uses the change of variables $n = 1 - H_{ij}^e(e)$.

Combining these expressions, we get that

$$\bar{c}_{ij} = D_j \kappa_{ij}(n_{ij}) \tag{OA.65}$$

with

$$\kappa_{ij}(n_{ij}) \equiv \frac{1}{n_{ij}} \int_0^{n_{ij}} \int \left[\mathcal{C}_j\left(\epsilon_{ij}(n_{ij})/r\right) + \frac{\epsilon_{ij}(n_{ij})}{r} \bar{\Pi}_j\left(\epsilon_{ij}(n)/r\right)\right] dH_{ij}^r\left(r|\epsilon_{ij}(n)\right) dn. \tag{OA.66}$$

In order to compute $\kappa_{ij}(n)$ using (OA.66), one needs to know $H_{ij}^r(r|\epsilon_{ij}(n))$, $\epsilon_{ij}(n)$, $\mathcal{C}_j(.)$ and $\bar{\Pi}_j(.)$. Note that knowledge of the demand function $q_j(.)$ in (21) implies that we can compute $\mathcal{C}_j(.)$, $\mathcal{R}_j(.)$ and $\bar{\Pi}_j(.)$ using (OA.52)–(OA.54). Thus, it only remains to show how we can recover $H_{ij}^r(r|\epsilon_{ij}(n))$. We consider two cases. First, without dispersion in fixed costs, there is a one-to-one mapping between $r_{ij}(\omega)$ and $e_{ij}(\omega)$, given the definition in (OA.56). This implies that $H_{ij}^r(r|\epsilon_{ij}(n))$ is degenerate at a known value determined by $\epsilon_{ij}(n)$. Second, when there is dispersion in fixed costs, expression (OA.61) yields $H_{ij}^r(r|\epsilon_{ij}(n))$ from $\rho_{ij}^{\pi}(n)$ and $\epsilon_{ij}(n)$. Thus, in this case, knowledge of $q_j(.)$, $\rho_{ij}^{\pi}(n)$ and $\epsilon_{ij}(n)$ implies that we can compute $\kappa_{ij}(n)$ using (OA.66).

We can now extent our proposition outlining the sufficient statistics for computing aggregate variables in general equilibrium.

Proposition 2 (non-CES demand)**.** *Consider the monopolistic competition model with the demand function in (21) described in the environment of Appendix Section A.2.1 under (OA.57). Assume knowledge of the exogenous fundamentals $\{\bar{r}_{ij}, \bar{L}_i, \bar{T}_i, \bar{F}_i\}$, the demand function in (21), and the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Then:*

a. For a given $\kappa_{ij}(n)$, the equilibrium vector $\{D_i, N_i, w_i\}$ solves the system of equations (OA.62)–(OA.64) with n_{ij}, $\bar{x}_{ij}$ and $\bar{c}_{ij}$ respectively given by (OA.58), (OA.59), and (OA.65).

b. The function $\kappa_{ij}(n)$ is identified (i) from $\epsilon_{ij}(n)$ without fixed cost dispersion and (ii) from $\epsilon_{ij}(n)$ and $\rho_{ij}^{\pi}(n)$ with fixed cost dispersion.

A.2.2 Multi-Sector, Multi-Factor Heterogeneous Firm Model with Input-Output Links and Import Tariffs

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

Environment

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

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

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

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

$$c_j^s = \gamma_j^s \left(\frac{P_j^s}{P_j} \right)^{1-\lambda_j} \tag{OA.67}$$

where the consumption price index is

$$P_j = \left[\sum_s \gamma_j^s \left(P_j^s \right)^{1-\lambda_j} \right]^{\frac{1}{1-\lambda_j}}. \tag{OA.68}$$

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 varieties $\omega \in \Omega^s$ 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 $\sigma^s > 1$ 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, its price is the CES price index of intermediate inputs:

$$\left(P_j^s \right)^{1-\sigma^s} = \sum_i \int_{\Omega_{ij}^s} \left(\bar{b}_{ij}^s b_{ij}^s(\omega) \right) \left(p_{ij}^s(\omega) \right)^{1-\sigma^s} d\omega. \tag{OA.69}$$

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

$$L_i^s = \left[\sum_v \beta_i^{s,v}\left(L_i^{s,v}\right)^{\frac{\eta_i^s-1}{\eta_i^s}}\right]^{\frac{\eta_i^s}{\eta_i^s-1}} \quad \text{and} \quad M_i^s = \left[\sum_k \theta_i^{ks}\left(Q_i^k\right)^{\frac{\kappa_i^s-1}{\kappa_i^s}}\right]^{\frac{\kappa_i^s-1}{\kappa_i^s}}.$$

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

$$p_i^s = \left[\alpha_i^s \left(W_i^s\right)^{1-\mu_i^s} + \left(1-\alpha_i^s\right)\left(J_i^s\right)^{1-\mu_i^s}\right]^{\frac{1}{1-\mu_i^s}}, \tag{OA.70}$$

where

$$W_i^s = \left[\sum_v \beta_i^{s,v}\left(w_i^v\right)^{1-\eta_i^s}\right]^{\frac{1}{1-\eta_i^s}} \quad \text{and} \quad J_i^s = \left[\sum_k \theta_i^{ks}\left(P_i^k\right)^{1-\kappa_i^s}\right]^{\frac{1}{1-\kappa_i^s}}. \tag{OA.71}$$

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

$$l_i^{s,v} = \beta_i^{s,v}\left(\frac{w_i^v}{W_i^s}\right)^{1-\eta_i^s}\alpha_i^s\left(\frac{W_i^s}{p_i^s}\right)^{1-\mu_i^s} \quad \text{and} \quad m_i^{ks} = \theta_i^{ks}\left(\frac{P_i^k}{J_i^s}\right)^{1-\kappa_i^s}\left(1-\alpha_i^s\right)\left(\frac{J_i^s}{p_i^s}\right)^{1-\mu_i^s}. \tag{OA.72}$$

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. We also assume that country j imposes an ad-valorem tariff of t_{ij}^s on goods of sector s from country i. In order to sell q in market j, variety ω of country i faces a cost function given by

$$C_{ij}(\omega, q) = p_i^s(1+t_{ij}^s)\frac{\bar{\tau}_{ij}^s}{\bar{a}_i^s}\frac{\tau_{ij}^s(\omega)}{a_i^s(\omega)}q + p_i^s \bar{f}_{ij}^s f_{ij}^s(\omega)$$

where p_i^s is the price of the non-tradable input q_i^s in country i.

Entry and Revenue Potentials. We now define the two variables that determine firm-level revenue and entry in each sector. Given this production technology, the optimal price is $p_{ij}^s(\omega) = \frac{\sigma^s}{\sigma^s-1}\frac{\tau_{ij}^s(\omega)}{a_i^s(\omega)}\frac{(1+t_{ij}^s)\bar{\tau}_{ij}^s}{\bar{a}_i^s}p_i^s$ and the associated revenue is

$$R_{ij}^s(\omega) = \left((p_j^s)^{1-\sigma^s}(P_j^s)^{\sigma^s-1}E_j^s\right)\bar{r}_{ij}^s r_{ij}^s(\omega) \tag{OA.73}$$

where

$$r_{ij}^s(\omega) \equiv b_{ij}^s(\omega)\left(\frac{\tau_{ij}^s(\omega)}{a_i^s(\omega)}\right)^{1-\sigma^s} \quad \text{and} \quad \bar{r}_{ij}^s \equiv \bar{b}_{ij}^s\left(\frac{\sigma^s}{\sigma^s-1}\frac{(1+t_{ij}^s)\bar{\tau}_{ij}^s}{\bar{a}_i^s}\right)^{1-\sigma^s}. \tag{OA.74}$$

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

$$\Omega_{ij}^s = \{\omega : e_{ij}^s(\omega) \geq e_{ij}^{s,*}\} \tag{OA.75}$$

where

$$e^s_{ij}(\omega) \equiv \frac{r^s_{ij}(\omega)}{f^s_{ij}(\omega)}, \quad \text{and} \quad e^{s,*}_{ij} \equiv \frac{\bar{r}^s_{ij}}{\sigma^s \bar{f}^s_{ij}} \left[\left(\frac{p^s_i}{P^s_j} \right)^{\sigma^s} \frac{P^s_j}{E^s_j} \right]. \tag{OA.76}$$

Extensive and Intensive Margins of Firm Exports

We now turn to the characterization of entry and sales in each sector. We consider the distribution of $(r^s_{ij}(\omega), e^s_{ij}(\omega))$:

$$r^s_{ij}(\omega) \sim H^{r,s}_{ij}(r|e) \qquad \text{and} \qquad e^s_{ij}(\omega) \sim H^{e,s}_{ij}(e), \tag{OA.77}$$

where $H^{e,s}_{ij}$ 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 $n^s_{ij} = Pr\left[\omega \in \Omega^s_{ij}\right]$. We define $\epsilon^s_{ij}(n) \equiv \left(H^{e,s}_{ij}\right)^{-1}(1-n)$ such that

$$\ln \epsilon^s_{ij}(n^s_{ij}) = \ln \sigma^s \bar{f}^s_{ij}/\bar{r}^s_{ij} + \ln (p^s_i)^{\sigma^s} - \ln E^s_j \left(P^s_j\right)^{\sigma^s - 1}. \tag{OA.78}$$

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 $\bar{x}^s_{ij} \equiv \mathbb{E}\left[R^s_{ij}(\omega)|\omega \in \Omega^s_{ij}\right]$. Define the mean revenue potential of exporters when $n\%$ of $i's$ firms in sector s export to j as $\rho^s_{ij}(n) \equiv \frac{1}{n}\int_0^n \mathbb{E}[r|e = \epsilon^s_{ij}(n)]\, dn$. The change of variable $n = 1 - H^{e,s}_{ij}(e)$ implies that

$$\ln \bar{x}^s_{ij} - \ln \rho^s_{ij}(n^s_{ij}) = \ln \left(\bar{r}^s_{ij}\right) + \ln (p^s_i)^{1-\sigma^s} + \ln E^s_j \left(P^s_j\right)^{\sigma^s - 1}. \tag{OA.79}$$

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

We can now extend Proposition 1.

Proposition 1 (multi-sector, multi-factor, import tariffs)**.** *Consider the monopolistic competition model with multiple factors, multiple sectors, input-output linkages and import tariffs described in the environment of Appendix Section A.2.2 under (OA.77). Then, for any origin i, destination j and sector s, the exporter firm share, n^s_{ij}, and the average firm exports, $\bar{x}^s_{ij}$, are given by equations (OA.78) and (OA.79), which are separable on country-level endogenous variables, exogenous bilateral shifters, and two elasticity functions of the exporter firm share $n \in [0,1]$, $\epsilon^s_{ij}(n)$ and $\rho^s_{ij}(n)$.*

Sufficient Statistics in General Equilibrium

We now describe the conditions establishing free entry, budget balance and factor market clearing.

Firms in sector s of country i can create a new variety by spending $\bar{F}^s_i$ units of the non-tradable sectoral input q^s_i. In equilibrium, free entry implies that N^s_i firms pay the fixed cost of entry in exchange for an ex-ante expected profit of zero, $\sum_j \mathbb{E}\left[\max\left\{\pi_{ij}^s(\omega);\ 0\right\}\right] = p_i^s \bar{F}_i^s$. Following the same steps described in Section 2.3, we can show that

$$\frac{1}{\sigma^s}\sum_j \frac{n_{ij}^s \bar{x}_{ij}^s}{1+t_{ij}^s} = p_i^s \bar{F}_i + p_i^s \sum_j \bar{f}_{ij}^s \int_0^{n_{ij}^s} \frac{\rho_{ij}^s(n)}{\epsilon_{ij}^s(n)}(1+\varrho_{ij}^s(n))dn. \tag{OA.80}$$

Thus, the free entry condition can be written as a function of the elasticity functions $\rho_{ij}^s(n)$ and $\epsilon_{ij}^s(n)$ (recall that we argued above that this is true also for $\bar{x}_{ij}^s$ and n_{ij}^s).

We now turn to the budget balance condition that determines the sectoral price index P_j^s in (OA.69). Using the expression for $p_{ij}^s(\omega)$ and (OA.69), we have that $(P_j^s)^{1-\sigma^s} = \sum_i \bar{r}_{ij}^s \left(p_i^s\right)^{1-\sigma^s} \int_{\Omega_{ij}^s} r_{ij}^s(\omega)\, d\omega$. Since $\int_{\Omega_{ij}^s} r_{ij}^s(\omega)\ d\omega = N_i^s Pr[\omega \in \Omega_{ij}^s]\mathbb{E}[r|\omega \in \Omega_{ij}^s] = N_i^s n_{ij}^s \rho_{ij}^s(n_{ij}^s)$, we can write P_j^s as

$$(P_j^s)^{1-\sigma^s} = \sum_i \bar{r}_{ij}^s \left(p_i^s\right)^{1-\sigma^s} \rho_{ij}^s(n_{ij}^s) n_{ij}^s N_i^s. \tag{OA.81}$$

We again 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

$$E_i^s = c_i^s \left(\sum_v w_i^v \bar{L}_i^v + \bar{T}_i + R_i^t\right) + \sum_k m_i^{sk} \sum_j \frac{N_i^k n_{ij}^k \bar{x}_{ij}^k}{1+t_{ij}^k}, \tag{OA.82}$$

with m_i^{sk} the intermediate spending share given by (OA.72), and R_i^t is the import tariff revenue that is given by

$$R_i^t = \sum_j \sum_s \frac{t_{ji}^s}{1+t_{ji}^s} N_j^s n_{ji}^s \bar{x}_{ji}^s.$$

Finally, the market clearing conditions for factor v in country i is

$$w_i^v \bar{L}_i^v = \sum_s l_i^{s,v} \sum_j \frac{N_i^s n_{ij}^s \bar{x}_{ij}^s}{1+t_{ij}^s}, \tag{OA.83}$$

with $l_i^{s,v}$ given by (OA.72).

Thus, because the conditions above only depend $\bar{x}_{ij}^s$ and n_{ij}^s, they can also be written as a function of the elasticity functions $\rho_{ij}^s(n)$ and $\epsilon_{ij}^s(n)$.

The following proposition summarizes the conditions that determine aggregate variables in general equilibrium.

Proposition 2 (multi-sector, multi-factor, import tariffs)**.** *Consider the monopolistic competition model with multiple factors, multiple sectors, input-output linkages and import tariffs described in the environment of Appendix Section A.2.2 under (OA.77). Assume knowledge of the exogenous fundamentals* $\{\bar{r}_{ij}^s, \bar{f}_{ij}^s, t_{ij}^s, \bar{F}_i^s, \gamma_i^s, \alpha_i^s, \theta_i^{ks}, \beta_i^{s,v}, \bar{L}_i^v, \bar{T}_i\}$*, the elasticity of substitution in consumption and production* $\{\sigma^s, \lambda_i, \mu_i^s, \eta_i^s, \kappa_i^s\}$*, and the elasticity functions* $\epsilon_{ij}^s(n)$ *and* $\rho_{ij}^s(n)$*. Then, the equilibrium vector* $\{N_i^s, P_i^s, E_i^s, w_i^v\}$ *solves the system of equations (OA.80)-(OA.83) with* n_{ij}^s *and* $\bar{x}_{ij}^s$ *given by (OA.78) and (OA.79), and the sectoral input price* p_i^s *given by (OA.70)-(OA.71).*

A.2.3 Allowing for Zero Bilateral Trade

In this section, we extend our baseline framework to allow for zero trade flows between two countries. We do so by considering a weaker version of the full support requirement for entry potentials in Assumption 1.

Environment

Consider the same environment described in Section 2.1.

Extensive and Intensive Margins of Firm Exports

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

$$r_{ij}(\omega) \sim H^r_{ij}\left(r|e\right), \qquad \text{and} \qquad e_{ij}(\omega) \sim H^e_{ij}(e). \tag{OA.84}$$

We now however consider a weaker version of Assumption 1.

Assumption 1': *$H^e_{ij}(e)$ is continuous and strictly increasing in $[0, \bar{e}^*_{ij}]$ with $\bar{e}_{ij} < \infty$.*

This assumption specifies that the distribution of entry potentials has full support in a bounded interval. This allows for zero trade flows, as in Helpman et al. (2008). We now use this assumption to derive the expressions for the extensive and intensive margins of firm-level exports.

Extensive margin of firm-level exports. Recall that $n_{ij} \equiv Pr[\omega \in \Omega_{ij}]$ where Ω_{ij} is given by (7). It implies that

$$n_{ij} = \begin{cases} 1 - H^e_{ij}(e^*_{ij}) & \text{if} \quad e^*_{ij} \leq \bar{e}_{ij} \\ 0 & \text{if} \quad e^*_{ij} > \bar{e}_{ij} \end{cases}$$

with e^*_{ij} defined in (8).

Let us now define

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

The definition of $\tilde{\epsilon}_{ij}(n)$ and the expression for n_{ij} above imply that that $\tilde{\epsilon}_{ij}(n_{ij}) = \min\left\{e^*_{ij}, \bar{e}_{ij}\right\}$. Thus, by defining $\epsilon_{ij}(n) \equiv \tilde{\epsilon}_{ij}(n_{ij})/\bar{e}_{ij}$, we get that

$$\ln \epsilon_{ij}(n_{ij}) = \min\left\{-\ln(\sigma \bar{f}_{ij}\bar{e}_{ij}/\bar{r}_{ij}) + \ln\left(w_i^\sigma\right) - \ln\left(E_j P_j^{\sigma-1}\right), 0\right\}. \tag{OA.85}$$

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

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

We again consider the transformation $n = 1 - H^e_{ij}(e)$ such that $e = \epsilon_{ij}(n)$ and $dH_{ij}(e) = -dn$. Thus,

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln \bar{r}_{ij} + \ln\left(w_i^{1-\sigma}\right) + \ln\left(E_j P_j^{\sigma-1}\right), \tag{OA.86}$$

where we normalize $\rho_{ij}(0) = 0$.

Proposition 1 (zero trade flows). *Consider the monopolistic competition model with CES demand described in the environment of Appendix Section A.2.3 under Assumption 1'. Then, for any origin i and destination j, the exporter firm share, n_{ij}, and the average firm exports, $\bar{x}_{ij}$, are given by equations (OA.85) and (OA.86), which are separable on country-level endogenous variables, exogenous bilateral shifters, and two elasticity functions of the exporter firm share $n \in [0,1]$, $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$.*

Sufficient Statistics in General Equilibrium

The modified assumption on the support of entry potentials does not affect any of the derivations for the conditions determining free entry, budget balance, and labor market clearing. Thus, we can immediately state the extension of Proposition 2 using the modified expression for the extensive margin of firm exports in (OA.85).

Proposition 2 (zero trade flows). *Consider the monopolistic competition model with CES demand in the environment of Appendix Section A.2.3 under Assumption 1'. Assume knowledge of the exogenous fundamentals $\{\bar{r}_{ij}, \bar{f}_{ij}, \bar{e}_{ij}, \bar{L}_i, \bar{T}_i, \bar{F}_i\}$, the demand elasticity of substitution σ, and the elasticity functions $\epsilon_{ij}(n)$ and $\rho_{ij}(n)$. Then, the equilibrium vector $\{P_i, N_i, w_i\}$ solves the system of equations (18)-(20) with n_{ij} and $\bar{x}_{ij}$ given by (OA.85) and (OA.86).*

A.2.4 Multi-product Firms

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

Environment

Preferences. 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

$$q_{ij}(\omega) = \left(\frac{p_{ij}(\omega)}{P_j}\right)^{-\sigma} \frac{E_j}{P_j}, \tag{OA.87}$$

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,

$$P_j^{1-\sigma} = \sum_i \int_{\Omega_{ij}^v} (p_{ij}(\omega))^{1-\sigma} \, d\omega, \tag{OA.88}$$

and Ω_{ij}^v is the set of varieties produced in country i that are sold in country j.

Technology. 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 $w_i \frac{\tau_{ij}(\eta)}{a_i(\eta)} \frac{\bar{\tau}_{ij}}{\bar{a}_i}$.

Entry and Revenue Potentials. 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

$$R_{ij}^N(\eta) = \bar{r}_{ij}^N r_{ij}^N(\eta) \left[\left(\frac{w_i}{P_j} \right)^{1-\sigma} E_j \right], \tag{OA.89}$$

where

$$r_{ij}^N(\eta) \equiv \left(\frac{\tau_{ij}(\eta)}{a_i(\eta)} \right)^{1-\sigma} \quad \text{and} \quad \bar{r}_{ij}^N \equiv \left(\frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij}}{\bar{a}_i} \right)^{1-\sigma}. \tag{OA.90}$$

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

$$N_{ij}(\eta) = \left(\frac{1}{\sigma} \frac{R_{ij}^N(\eta)}{w_i} \right)^{\alpha}. \tag{OA.91}$$

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

$$R_{ij}(\eta) = \bar{r}_{ij} r_{ij}(\eta) w_i^{1-(1+\alpha)\sigma} \left(E_j P_j^{\sigma-1} \right)^{1+\alpha} \tag{OA.92}$$

$$r_{ij}(\eta) \equiv \left(r_{ij}^N(\eta) \right)^{1+\alpha} \quad \text{and} \quad \bar{r}_{ij} \equiv \frac{1}{\sigma^\alpha} \left(\bar{r}_{ij}^N \right)^{1+\alpha}. \tag{OA.93}$$

Conditional on entering market j, the firm's profit in that market is

$$\begin{aligned} \pi_{ij}(\eta) \quad = \quad & N_{ij}(\eta) \tfrac{1}{\sigma} R_{ij}^N(\eta) - w_i \tfrac{1}{1+1/\alpha} N_{ij}(\eta)^{1+1/\alpha} - w_i \bar{f}_{ij} f_{ij}(\eta) \\ & \left(\tfrac{1}{\sigma} \tfrac{R_{ij}^N(\eta)}{w_i} \right)^{\alpha} \tfrac{1}{\sigma} R_{ij}^N(\eta) - w_i \tfrac{1}{1+1/\alpha} \left(\tfrac{1}{\sigma} \tfrac{R_{ij}^N(\eta)}{w_i} \right)^{1+\alpha} - w_i \bar{f}_{ij} f_{ij}(\eta) \\ & \tfrac{1}{(1+\alpha)\sigma} \tfrac{1}{\sigma^\alpha w_i^\alpha} \left(R_{ij}^N(\eta) \right)^{1+\alpha} - w_i \bar{f}_{ij} f_{ij}(\eta) \end{aligned}$$

and, therefore,

$$\pi_{ij}(\eta) = \frac{1}{(1+\alpha)\sigma} \bar{r}_{ij} r_{ij}(\eta) w_i^{1-(1+\alpha)\sigma} \left(E_j P_j^{\sigma-1} \right)^{1+\alpha} - w_i \bar{f}_{ij} f_{ij}(\eta). \tag{OA.94}$$

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, Ω_{ij}:

$$\Omega_{ij} = \{ \eta : e_{ij}(\eta) \geq e_{ij}^* \} \tag{OA.95}$$

where

$$e_{ij}(\eta) \equiv \frac{r_{ij}(\eta)}{f_{ij}(\eta)}, \quad \text{and} \quad e_{ij}^* \equiv \frac{\bar{r}_{ij}}{(1+\alpha)\sigma \bar{f}_{ij}} \left[\frac{w_i^{(1+\alpha)\sigma}}{\left(E_j P_j^{\sigma-1} \right)^{1+\alpha}} \right]. \tag{OA.96}$$

Extensive and Intensive Margins of Firm Exports

We use the definitions of entry and revenue potentials to characterize firm-level entry and sales in different markets in general equilibrium. We assume that

$$r_{ij}(\eta) \sim H_{ij}^{r}(r|e) \qquad \text{and} \qquad e_{ij}(\eta) \sim H_{ij}^{e}(e), \tag{OA.97}$$

where H_{ij}^{e} satisfies Assumption 1.

Extensive margin of firm-level exports. The share of firms of country i serving market j is $n_{ij} = Pr\left[\eta \in \Omega_{ij}\right]$. Defining $\epsilon_{ij}(n) \equiv \left(H_{ij}^{e}\right)^{-1}(1-n)$, equation (OA.95) yields

$$\ln \epsilon_{ij}(n_{ij}) = \ln\left((1+\alpha)\sigma \bar{f}_{ij}/\bar{r}_{ij}\right) + \ln\left(w_i^{(1+\alpha)\sigma}\right) - \ln\left(E_j P_j^{\sigma-1}\right)^{1+\alpha}. \tag{OA.98}$$

Intensive margin of firm-level exports. The average revenue of firms from country i in country j is $\bar{x}_{ij} \equiv \mathbb{E}\left[R_{ij}(\eta)|\eta \in \Omega_{ij}\right]$ where $R_{ij}(\eta)$ is given by (OA.92). Define the average revenue potential of exporters when $n\%$ of $i's$ firms in sector s export to j as $\rho_{ij}(n) \equiv \frac{1}{n}\int_0^n \mathbb{E}[r|e=\epsilon_{ij}(n')]\,dn'$ where $\mathbb{E}[r|e=\epsilon_{ij}(n)]$ 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

$$\ln \bar{x}_{ij} - \ln \rho_{ij}(n_{ij}) = \ln\left(\bar{r}_{ij}\right) + \ln\left(w_i^{1-(1+\alpha)\sigma}\right) + \ln\left(E_j P_j^{\sigma-1}\right)^{1+\alpha}. \tag{OA.99}$$

Extensive margin of products per firm.
The average number of products among firms from i operating in market j is $N_{ij}^{v} = \mathbb{E}\left[N_{ij}(\eta)|\eta \in \Omega_{ij}\right]$. The expression for $N_{ij}(\eta)$ in (OA.91) implies that

$$N_{ij}^{v} = \frac{1}{\sigma^{\alpha}}\left(\bar{r}_{ij}^{N}\right)^{\alpha} w_i^{-\alpha\sigma}\left(E_j P_j^{\sigma-1}\right)^{\alpha}\mathbb{E}\left[(r_{ij}^{N}(\eta))^{\alpha}|\eta \in \Omega_{ij}\right]$$

and, since $\bar{r}_{ij} \equiv \frac{1}{\sigma^{\alpha}}\left(\bar{r}_{ij}^{N}\right)^{1+\alpha}$,

$$N_{ij}^{v} = \sigma^{-\frac{\alpha}{1+\alpha}}\bar{r}_{ij}^{\frac{\alpha}{1+\alpha}} w_i^{-\alpha\sigma}\left(E_j P_j^{\sigma-1}\right)^{\alpha}\mathbb{E}\left[(r_{ij}(\eta))^{\frac{\alpha}{1+\alpha}}|\eta \in \Omega_{ij}\right].$$

We consider a similar transformation as the one used above. Define $\rho_{ij}^{v}(n) \equiv \frac{1}{n}\int_0^n \mathbb{E}[r^{\frac{\alpha}{1+\alpha}}|e=\epsilon_{ij}(n)]\,dn$. 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

$$\ln N_{ij}^{v} - \ln \rho_{ij}^{v}(n_{ij}) = \frac{\alpha}{1+\alpha}\ln(\bar{r}_{ij}/\sigma) + \ln w_i^{-\alpha\sigma} + \ln\left(E_j P_j^{\sigma-1}\right)^{\alpha}. \tag{OA.100}$$

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.

We can now extend Proposition 1 for the model with multi-product firms.

Proposition 1 (multi-product firms)**.** *Consider the monopolistic competition model with CES demand described in the environment of Appendix Section A.2.4 under (OA.97). Then, for any origin i and destination*

j, the exporter firm share, n_{ij}, the average firm exports, $\bar{x}_{ij}$, and the average products per firm, N^v_{ij}, are respectively given by equations (OA.98), (OA.99) and (OA.100), which are separable on country-level endogenous variables, exogenous bilateral shifters, and three elasticity functions of the exporter firm share $n \in [0, 1]$, $\epsilon_{ij}(n)$, $\rho_{ij}(n)$, and $\rho^v_{ij}(n)$.

Sufficient Statistics in General Equilibrium

We now turn to the conditions determining aggregate variables in general equilibrium. We consider the first the free entry condition for firms. As in the baseline, we assume that an entrant firm pays a fixed labor cost $\bar{F}_i$ to draw its type. In a free entry equilibrium, N_i firms pay the fixed cost of entry in exchange for an ex-ante expected profit of zero such that $\sum_j \mathbb{E}\left[\max\left\{\pi_{ij}(\eta);\ 0\right\}\right] = w_i \bar{F}_i$. The expected profit can be written as

$$\begin{aligned}
\mathbb{E}\left[\max\left\{\pi_{ij}(\omega);\ 0\right\}\right] &= \sum_j Pr[\eta \in \Omega_{ij}]\mathbb{E}\left[\tfrac{1}{(1+\alpha)\sigma} R_{ij}(\eta) - w_i \bar{f}_{ij} f_{ij}(\eta) | \eta \in \Omega_{ij}\right] \\
&= \sum_j n_{ij} \left(\tfrac{1}{(1+\alpha)\sigma} \bar{x}_{ij} - w_i \bar{f}_{ij} \mathbb{E}[r_{ij}(\eta)/e_{ij}(\eta) | \eta \in \Omega_{ij}]\right). \\
&= \sum_j n_{ij} \left(\tfrac{1}{(1+\alpha)\sigma} \bar{x}_{ij} - w_i \bar{f}_{ij} \int_{e^*_{ij}}^{\infty} \tfrac{1}{e} E[r|e] \tfrac{dH^e(e)}{1-H^e(e^*_{ij})}\right) \\
&= \sum_j \tfrac{1}{(1+\alpha)\sigma} n_{ij} \bar{x}_{ij} - w_i \bar{f}_{ij} \int_0^{n_{ij}} \tfrac{\rho_{ij}(n)}{\epsilon_{ij}(n)} (1 + \varrho_{ij}(n)) dn \\
&= \sum_j \tfrac{1}{(1+\alpha)\sigma} n_{ij} \bar{x}_{ij} - \tfrac{w_i}{(1+\alpha)\sigma} \tfrac{\bar{r}_{ij}}{\bar{e}_{ij}} \int_0^{n_{ij}} \tfrac{\rho_{ij}(n)}{\epsilon_{ij}(n)} (1 + \varrho_{ij}(n)) dn,
\end{aligned}$$

where the first row uses (OA.92) and (OA.94), the second row uses (OA.96), third row uses (OA.95), the fourth row uses the change of variables $n = 1 - H_{ij}(e)$ and the definition of $\rho_{ij}(.)$, and the last row uses the definition of $\bar{e}_{ij}$ in (OA.96).

We can then write the free entry condition as

$$\frac{1}{(1+\alpha)\sigma} \sum_j n_{ij} \bar{x}_{ij} = w_i \bar{F}_i + w_i \sum_j \bar{f}_{ij} \int_0^{n_{ij}} \frac{\rho_{ij}(n)}{\epsilon_{ij}(n)} (1 + \varrho_{ij}(n)) dn. \tag{OA.101}$$

We then turn to the budget balance condition that determines the CES price index. Here, we use the fact that $p_{ij}(\omega) = \frac{\sigma}{\sigma-1} \frac{\bar{\tau}_{ij} w_i}{\bar{a}_i} \frac{\tau_{ij}(\eta)}{a_i(\eta)}$ for every variety ω of firm η to write directly the CES price index as

$$P_j^{1-\sigma} = \sum_i \int_{\Omega_{ij}} N_{ij}(\eta) \left(p_{ij}(\eta)\right)^{1-\sigma}\ d\eta.$$

Using the expression for $N_{ij}(\eta)$ in (OA.91) and the definitions in (OA.93), this expression can be written as

$$P_j^{1-\sigma} = \sum_i w_i^{1-\sigma} \frac{\bar{r}_{ij}}{w_i^{\alpha}} \left[\left(\frac{w_i}{P_j}\right)^{1-\sigma} \frac{E_j}{w_i}\right]^{\alpha} \int_{\Omega_{ij}} r_{ij}(\eta)\ d\eta.$$

Notice that $\int_{\Omega_{ij}} r_{ij}(\eta)\ d\eta = N_i Pr[\eta \in \Omega_{ij}] \mathbb{E}[r | \eta \in \Omega_{ij}] = N_i n_{ij} \rho_{ij}(n_{ij})$. 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,

$$P_j^{1-\sigma} = \sum_i \bar{r}_{ij} w_i^{1-(1+\alpha)\sigma} \left(E_j P_j^{\sigma-1}\right)^{\alpha} \rho_{ij}(n_{ij}) n_{ij} N_i. \tag{OA.102}$$

Finally, we again follow Dekle et al. (2008) by introducing exogenous international transfers, so that spending is

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

Since labor is the only factor of production, labor income in i equals the total revenue of firms from i:

$$w_i \bar{L}_i = \sum_j N_i n_{ij} \bar{x}_{ij}. \tag{OA.103}$$

We can now extend Proposition 2 for the model with multi-product firms.

Proposition 2 (multi-product firms)**.** *Consider the monopolistic competition model with CES demand described in the environment of Appendix Section A.2.4 under (OA.97). Assume knowledge of the exogenous fundamentals* $\{\bar{r}_{ij}, \bar{f}_{ij}, \bar{L}_i, \bar{T}_i, \bar{F}_i\}$*, the elasticity of supplying new varieties in a firm* α*, the demand elasticity of substitution* σ*, and the elasticity functions* $\epsilon_{ij}(n)$ *and* $\rho_{ij}(n)$*. Then, the equilibrium vector* $\{P_i, N_i, w_i\}$ *solves the system of equations (OA.101)-(OA.103) with* n_{ij} *and* $\bar{x}_{ij}$ *given by (OA.98) and (OA.99).*

B Empirical Appendix

B.1 Additional Figures and Tables

Table OA.1: Estimation Data Sources

Country	Source for N_{ii}	Source for N_{ij}	Source for $\bar{x}_{ij}$	Sectoral Data Available
ALB	EDD/WBES	EDD	EDD	1
AUS	SDBS	AUS	AUS	
AUT	SDBS	TEC	TEC	
BEL	SDBS	TEC	TEC	
BFA	EDD/WBES	EDD	EDD	1
BGD	EDD/WBES	EDD	EDD	1
BGR	SDBS	EDD	EDD	
BOL	EDD/WBES	EDD	EDD	1
BRA	SDBS	EDD	EDD	
BWA	EDD/WBES	EDD	EDD	1
CAN	SDBS	TEC	TEC	
CHL	SDBS	EDD	EDD	1
CHN	CHN	CHN	CHN	
CIV	EDD/WBES	EDD	EDD	1
CMR	EDD/WBES	EDD	EDD	1
COL	EDD/WBES	EDD	EDD	1
CRI	EDD/WBES	EDD	EDD	1
CYP	SDBS	TEC	TEC	
CZE	SDBS	TEC	TEC	
DEU	SDBS	TEC	TEC	
DNK	SDBS	EDD	EDD	1
DOM	EDD/WBES	EDD	EDD	1
ECU	EDD/WBES	EDD	EDD	1
EGY	EDD/WBES	EDD	EDD	1
ESP	SDBS	EDD	EDD	1
EST	SDBS	TEC	TEC	
ETH	EDD/WBES	EDD	EDD	1
FIN	SDBS	TEC	TEC	
FRA	SDBS	TEC	TEC	
GAB	EDD/WBES	EDD	EDD	
GBR	SDBS	TEC	TEC	
GEO	EDD/WBES	EDD	EDD	1
GIN	EDD/WBES	EDD	EDD	1
GRC	SDBS	TEC	TEC	
GTM	EDD/WBES	EDD	EDD	1

continued

Table OA.1: Estimation Data Sources

Country	Source for N_{ii}	Source for N_{ij}	Source for $\bar{x}_{ij}$	Sectoral Data Available
HRV	SDBS	EDD	EDD	1
HUN	SDBS	TEC	TEC	
IRL	SDBS	TEC	TEC	
ITA	SDBS	TEC	TEC	
JOR	EDD/WBES	EDD	EDD	1
KEN	EDD/WBES	EDD	EDD	1
KGZ	EDD/WBES	EDD	EDD	1
KHM	EDD/WBES	EDD	EDD	
KOR	SDBS	TEC*	TEC	
LAO	EDD/WBES	EDD	EDD	
LBN	EDD/WBES	EDD	EDD	1
LKA	EDD/WBES	EDD	EDD	
LTU	SDBS	TEC	TEC	
LUX	SDBS	TEC	TEC	
LVA	SDBS	TEC	TEC	
MAR	EDD/WBES	EDD	EDD	1
MDG	EDD/WBES	EDD	EDD	1
MEX	EDD/WBES	EDD	EDD	1
MKD	EDD/WBES	EDD	EDD	
MLI	EDD/WBES	EDD	EDD	
MLT	SDBS	TEC	TEC	
MMR	EDD/WBES	EDD	EDD	1
MUS	EDD/WBES	EDD	EDD	1
MWI	EDD/WBES	EDD	EDD	1
NER	EDD/WBES	EDD	EDD	
NIC	EDD/WBES	EDD	EDD	1
NLD	SDBS	TEC	TEC	
NOR	SDBS	EDD	EDD	1
NPL	EDD/WBES	EDD	EDD	1
NZL	SDBS	TEC	TEC	
PAK	EDD/WBES	EDD	EDD	
PER	EDD/WBES	EDD	EDD	1
POL	SDBS	TEC	TEC	
PRT	SDBS	EDD	EDD	1
PRY	EDD/WBES	EDD	EDD	1
ROU	SDBS	TEC	TEC	
RWA	EDD/WBES	EDD	EDD	1
SEN	EDD/WBES	EDD	EDD	1
SLV	EDD/WBES	EDD	EDD	
SVK	SDBS	TEC	TEC	
SVN	SDBS	TEC	TEC	
SWE	SDBS	TEC	TEC	
SWZ	EDD/WBES	EDD	EDD	1
THA	EDD/WBES	EDD	EDD	1
TUR	SDBS	EDD	EDD	
TZA	EDD/WBES	EDD	EDD	1
UGA	EDD/WBES	EDD	EDD	
URY	EDD/WBES	EDD	EDD	1
USA	SDBS	TEC	TEC	
YEM	EDD/WBES	EDD	EDD	1
ZAF	EDD/WBES	EDD	EDD	1
ZMB	EDD/WBES	EDD	EDD	

Notes: This table shows the most frequent use, however in certain cases when N_{ij} is small, the EDD suppresses $\bar{x}$, but not N_{ij}, so we construct $\bar{x}_{ij} = X_{ij}/N_{ij}$, where X_{ij} comes from BACI and N_{ij} comes from the EDD. For countries from the EDD sample, we use the WBES (World Bank Enterprise Survey) to back out the number of N_{ii} from survey sampling probabilities.

Table OA.2: Estimation Data Summary

Country	Developed Dummy	Number of Destinations	Average across j: $\ln n_{ij}$	Average across j: $\ln \bar{x}_{ij}$	Standard deviation across j: $\ln n_{ij}$	Standard deviation across j: $\ln \bar{x}_{ij}$
ALB	0	73	-8.51	-2.82	1.67	2.28
AUS	1	144	-6.91	-1.23	1.9	1.39
AUT	1	58	-3.44	.11	.86	.98

continued

Table OA.2: Estimation Data Summary

Country	Developed	Number of	Average across j		Standard deviation across j	
	Dummy	Destinations	$\ln n_{ij}$	$\ln \bar{x}_{ij}$	$\ln n_{ij}$	$\ln \bar{x}_{ij}$
BEL	1	58	-3.21	.36	.74	1.08
BFA	0	53	-7.98	-1.54	1.46	2.24
BGD	0	127	-7	-2.02	2.08	1.42
BGR	0	145	-6.74	-2.26	1.98	1.98
BOL	0	61	-7.15	-1.73	1.45	1.72
BRA	0	156	-7.92	-.82	1.92	1.44
BWA	0	57	-9.25	-3.23	1.66	2.91
CAN	1	47	-4.78	-.18	1.29	1.31
CHL	0	135	-5.13	-1.3	1.9	1.59
CHN	0	155	-4.47	-1.15	1.73	.99
CIV	0	97	-8.41	-.95	1.52	1.72
CMR	0	87	-7.78	-2	1.36	1.5
COL	0	126	-8.64	-1.56	2.01	1.45
CRI	0	108	-8.36	-2.21	1.95	2.09
CYP	1	34	-5.28	-.97	1.13	1.14
CZE	0	58	-4.18	-.21	1.05	.86
DEU	1	35	-3.15	.47	.74	1.15
DNK	1	155	-4.69	-1.55	1.69	1.37
DOM	0	110	-8.7	-2.09	1.66	1.51
ECU	0	106	-8.69	-1.53	1.76	1.82
EGY	0	136	-7.61	-1.72	1.58	1.49
ESP	1	157	-5.78	-1.76	1.95	1.28
EST	0	70	-4.42	-1	1.29	1.29
ETH	0	74	-8.14	-2.68	1.26	2.34
FIN	1	35	-3.48	.3	.93	1.2
FRA	1	58	-3.99	.17	.9	.99
GAB	0	57	-6.48	-1.96	1.32	2.16
GBR	1	58	-3.96	-.2	.85	.99
GEO	0	84	-9.17	-1.43	1.6	1.77
GIN	0	60	-8.73	-2.31	1.31	3.02
GRC	1	58	-5.78	-.61	1.26	1.07
GTM	0	106	-8.21	-1.92	1.94	1.98
HRV	0	101	-7.29	-2.49	1.7	2.33
HUN	0	35	-4.58	.66	1.15	.91
IRL	1	35	-3.66	1.15	.95	1.4
ITA	1	58	-3.54	-.89	1.14	.82
JOR	0	118	-7.04	-2.07	1.72	1.71
KEN	0	112	-7.35	-2.74	1.69	1.7
KGZ	0	57	-7.58	-2.06	1.42	2.03
KHM	0	93	-7.25	-2.56	1.62	1.92
KOR	1	48	-5.24	-.04	1.32	.99
LAO	0	45	-6.55	-1.86	1.22	2.18
LBN	0	136	-6.86	-2.74	1.83	1.45
LKA	0	135	-8.34	-2.41	1.66	1.55
LTU	0	58	-4.73	-.6	1.31	1.3
LUX	1	33	-2.43	.61	.75	1.55
LVA	0	47	-4.25	-.75	1.18	1
MAR	0	117	-7.74	-1.3	1.67	1.77
MDG	0	81	-7.16	-2.34	1.57	1.71
MEX	0	149	-9.39	-1.5	2.29	1.61
MKD	0	77	-7.61	-2.7	1.85	1.96
MLI	0	41	-7.17	-2.14	1.31	2.77
MLT	1	57	-4.72	-.63	1.14	1.69
MMR	0	51	-8.12	-1.44	1.45	1.5
MUS	0	108	-8.14	-2.55	1.53	2.03
MWI	0	75	-7.95	-1.16	1.29	2.2
NER	0	27	-7.35	-2.19	.83	2.8
NIC	0	84	-8.33	-2.65	1.6	2.23
NLD	1	58	-3.09	.14	.71	1.07
NOR	1	148	-5.25	-1.62	1.9	1.93
NPL	0	78	-7.76	-4.19	1.68	1.48
NZL	1	39	-4.95	-.73	1.21	1.39
PAK	0	149	-7.52	-2.44	1.85	1.21
PER	0	122	-8.31	-1.65	1.98	1.65
POL	0	58	-4.72	-.2	1.16	.85
PRT	1	149	-6.37	-2	2.03	1.44
PRY	0	66	-7.07	-1.26	1.43	1.94
ROU	0	27	-4.78	.28	1.38	.93
RWA	0	52	-9.13	-3.98	1.3	2.38

continued

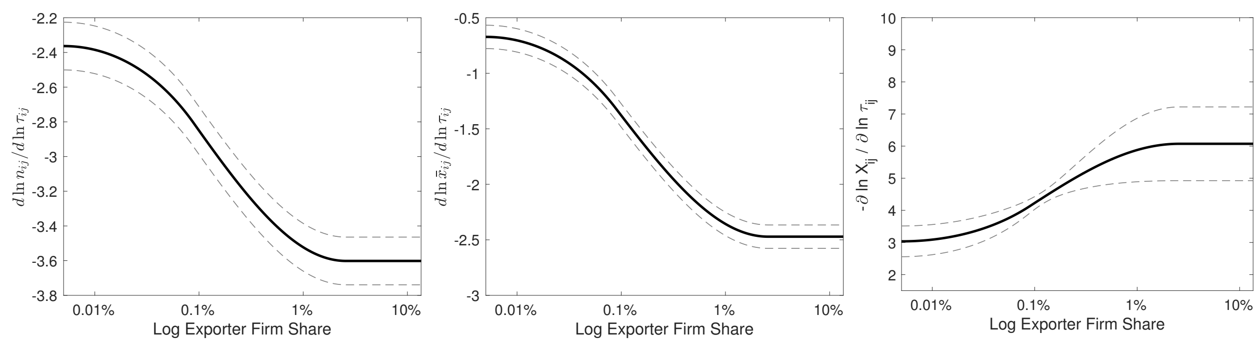
Table OA.2: Estimation Data Summary

Country	Developed Dummy	Number of Destinations	Average across j $\ln n_{ij}$	Average across j $\ln \bar{x}_{ij}$	Standard deviation across j $\ln n_{ij}$	Standard deviation across j $\ln \bar{x}_{ij}$
SEN	0	81	-7.18	-2.99	1.53	2.42
SLV	0	82	-7.49	-2.44	1.73	1.73
SVK	0	56	-4.28	-.16	1.54	.82
SVN	1	55	-4.16	-.43	1.11	.94
SWE	1	58	-3.48	.33	.9	.98
SWZ	0	58	-8.25	-1.53	1.39	2.83
THA	0	157	-8.3	-1.02	1.78	1.25
TUR	0	152	-6.82	-1.37	1.81	.92
TZA	0	101	-8.22	-1.89	1.52	1.8
UGA	0	81	-8.83	-2.92	1.45	2.97
URY	0	123	-8.02	-1.54	1.62	1.88
USA	1	47	-4.34	-.01	1.26	1.21
YEM	0	46	-6.95	-2.07	1.17	2.66
ZAF	0	153	-8.14	-1.62	1.86	1.35
ZMB	0	74	-8.77	-1.98	1.53	2.84

Notes: This table shows the summary statistics of the key variables used.

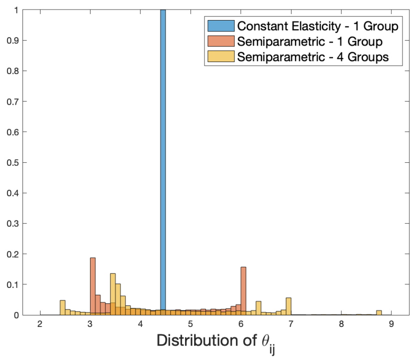
Figure OA.1: Semiparametric Gravity of Firm Exports – Single Group

(a) Exporter Firm Share Elasticity (b) Average Firm Exports Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). We report the elasticity of the exporter firm share to trade costs $\partial \ln n_{ij}/\partial \ln \tau_{ij} = (\sigma - 1)/\varepsilon(n)$ in panel (a), the elasticity of the average firm exports to trade costs $\partial \ln \bar{x}_{ij}/\partial \ln \tau_{ij} = (\sigma - 1)(1 - \varrho(n)/\varepsilon(n))$ in panel (b), and the (absolute) elasticity of bilateral trade with respect to trade costs $(\sigma - 1)\theta(n)$ in panel (c). Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

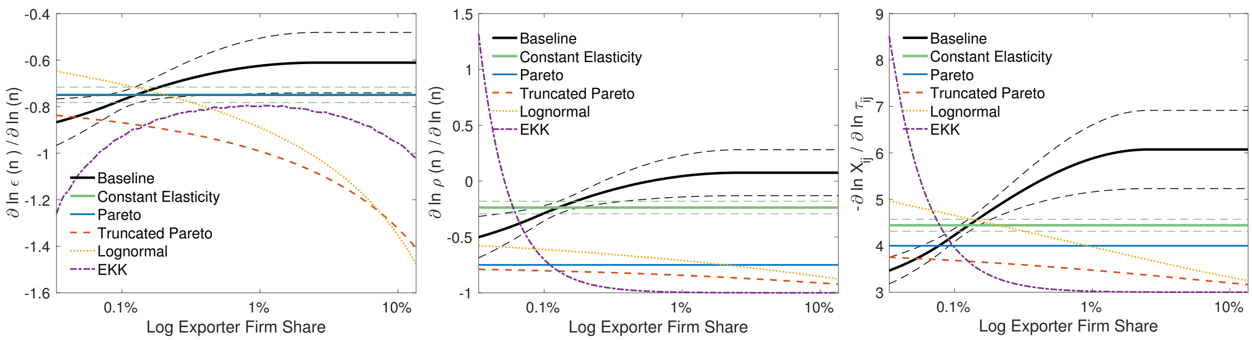
Figure OA.3: Empirical Distribution of Bilateral Trade Elasticities in 2012



Note. Figure displays the histogram of the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma-1)\theta(n_{ij}^0)$, in the 2012 sample of 7,386 origin-destination pairs. The blue bars denote the empirical distribution implied by the constant-elasticity benchmark obtained from the estimation of (37) under (24) (as reported in panel (c) of Figure OA.2). Yellow and orange bars denote empirical distributions implied by semiparametric estimates obtained with GMM estimator in (37) for a single group with all countries and for four groups based on the income level of the origin and destination (as reported in panel (c) of Figures 3 and 5, respectively).

Figure OA.2: Elasticity of Firm Exports and Distributional Assumptions – Single Group

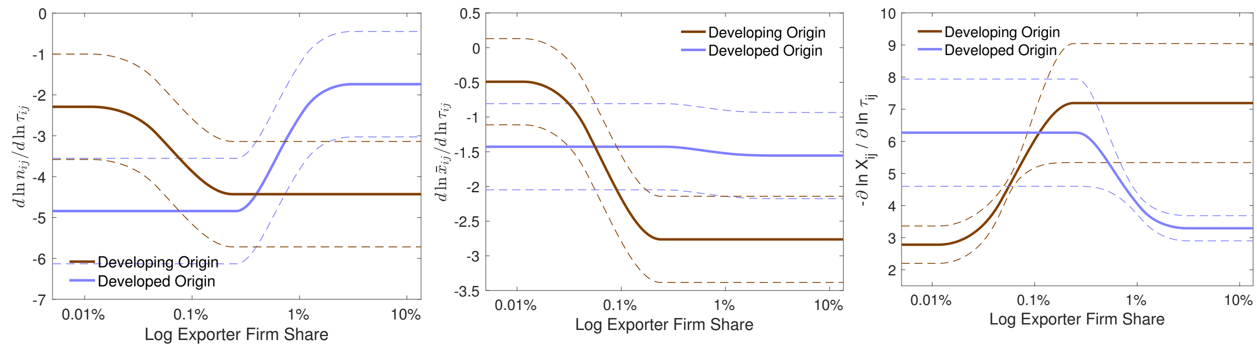
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma-1)\theta(n)$. Solid black lines and dashed black lines are the estimates and associated 90% confidence intervals of the semiparametric estimates reported in Figure 3. The green solid lines are the estimates of the constant-elasticity benchmark obtained from the estimation of (37) under (24) for our baseline sample of origin-destination pairs, with dashed green lines the associated 90% confidence intervals. Other lines correspond to the elasticity functions implied by alternative functional form assumptions, as described in Figure 1.

Figure OA.4: Semiparametric Gravity of Firm Exports – Developed and Developing Origins

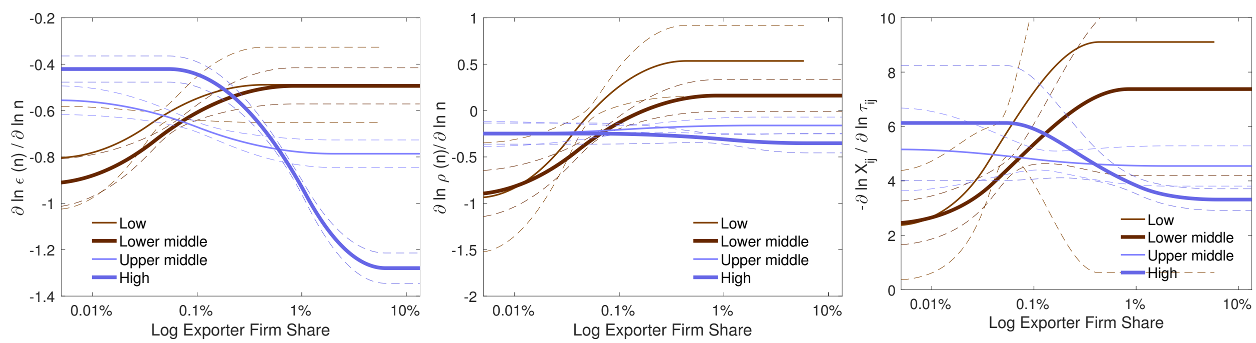
(a) Exporter Firm Share Elasticity (b) Average Firm Exports Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination. We assume that there are two groups of markets ($G = 2$) defined by whether the origin country is developed (light purple) or developing (dark brown), as defined in Table OA.2. We report the elasticity of the exporter firm share to trade costs $\partial \ln n_{ij} / \partial \ln \tau_{ij} = (\sigma - 1)/\varepsilon_g(n)$ in panel (a), the elasticity of the average firm exports to trade costs $\partial \ln \bar{x}_{ij} / \partial \ln \tau_{ij} = (\sigma - 1)(1 - \varrho_g(n)/\varepsilon_g(n))$ in panel (b), and the (absolute) elasticity of bilateral trade with respect to trade costs $(\sigma - 1)\theta_g(n)$ in panel (c). Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

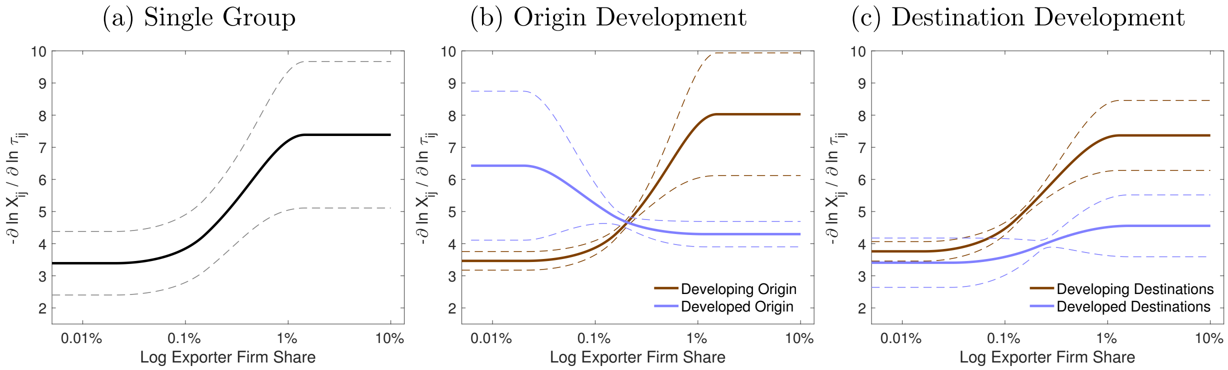
Figure OA.5: Semiparametric Gravity of Firm Exports – Origin's Income Level

(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



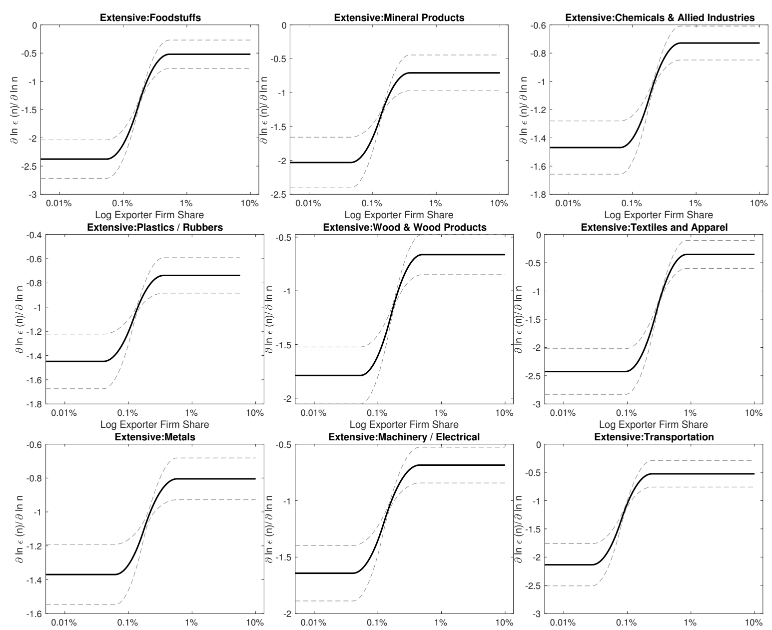
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination. We assume that there are four groups of markets ($G = 4$) defined by the origin's level of income (low, med-low, med-high, high) according to the World Bank classification in 2000. Panels (a) and (b) report respectively the elasticity of $\epsilon_g(n)$ and $\rho_g(n)$ with respect to the exporter firm share n, $\varepsilon_g(n)$ and $\varrho_g(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Figure OA.6: Semiparametric Gravity of Firm Exports – Within-Sector Estimation



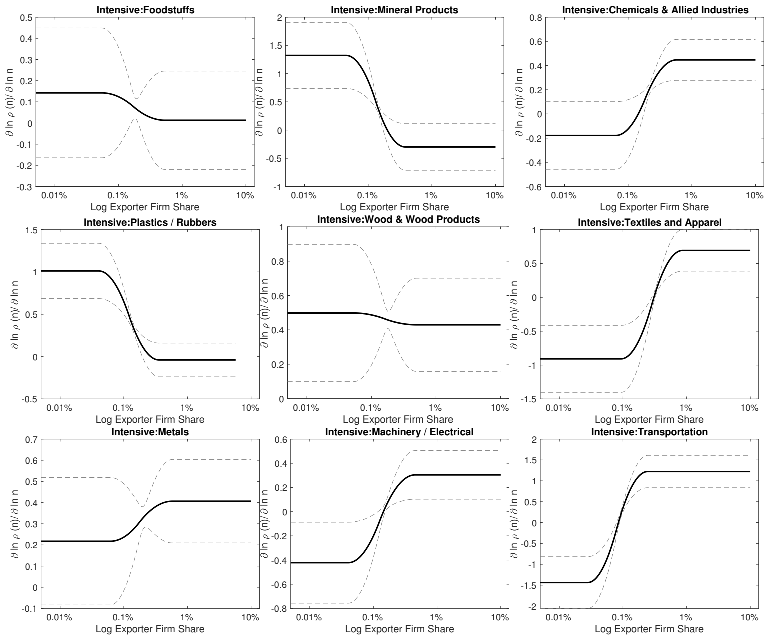
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 16,052 markets defined as origin-destination-sector triplets and fixed effects for sector-origin and sector-destination. We report estimates for a single group with all markets in panel (a), for two groups based on whether the origin country is developed in panel (b), and for two groups based on whether the destination country is developed in panel (c). Solid lines are estimates of the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.7: Semiparametric Gravity of Firm Exports – Extensive Margin Elasticity



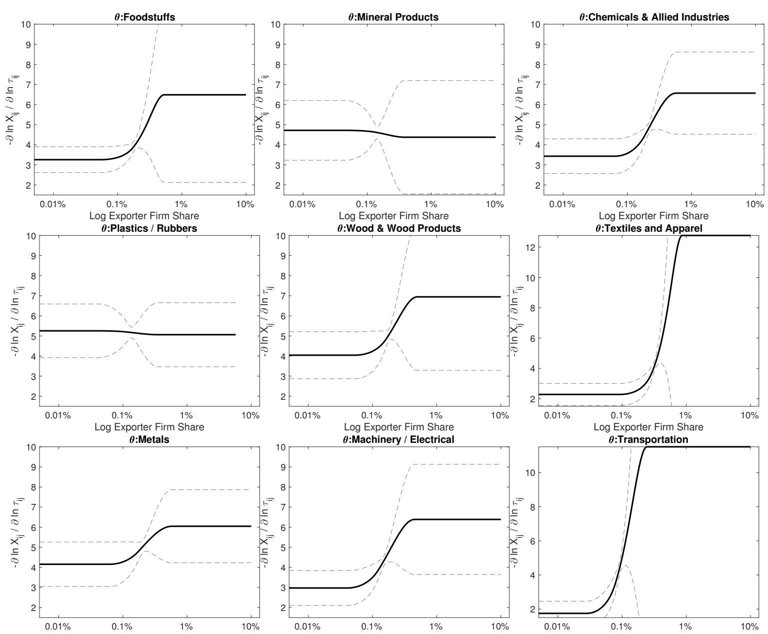
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of markets defined as origin-destination-sector triplets and fixed effects for sector-origin and sector-destination. We report estimates for a single group with all markets. Each panel reports estimates for the sector indicated in the title. Solid lines are the elasticity of $\epsilon_s(n)$ with respect to the exporter firm share n, $\varepsilon_s(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.8: Semiparametric Gravity of Firm Exports – Intensive Margin Elasticity



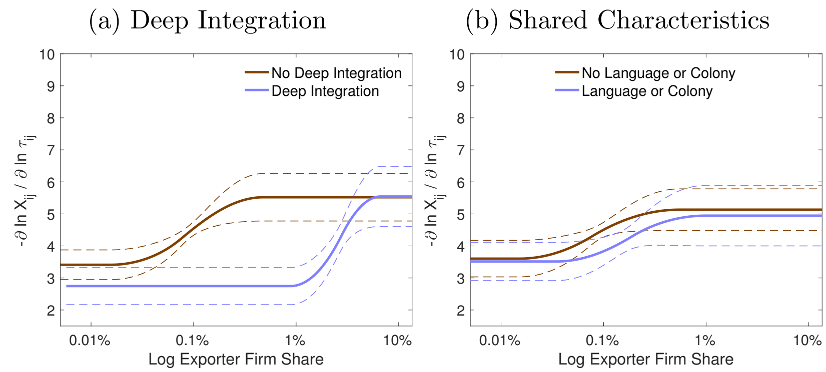
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of markets defined as origin-destination-sector triplets and fixed effects for sector-origin and sector-destination. We report estimates for a single group with all markets. Each panel reports estimates for the sector indicated in the title. Solid lines are the elasticity of $\rho_s(n)$ with respect to the exporter firm share n, $\varrho_s(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.9: Semiparametric Gravity of Firm Exports – Bilateral Trade Elasticity



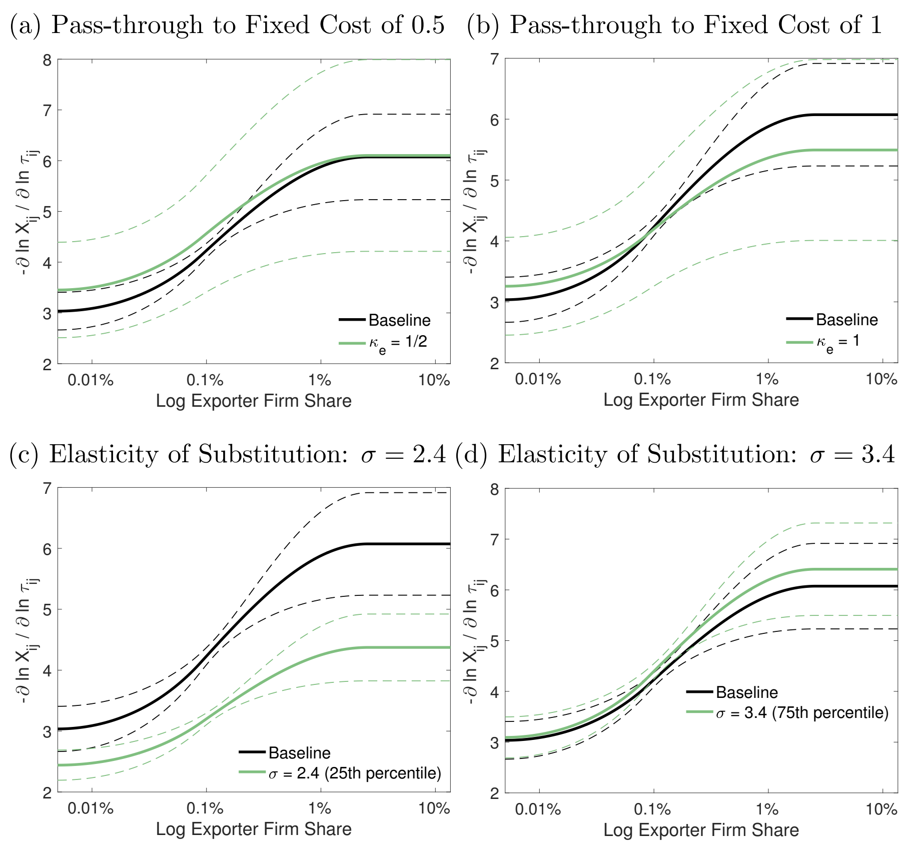
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of markets defined as origin-destination-sector triplets and fixed effects for sector-origin and sector-destination. We report estimates for a single group with all markets. Each panel reports estimates for the sector indicated in the title. Solid lines are estimates of the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_s(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.10: Semiparametric Gravity of Firm Exports – Determinants of Market Integration



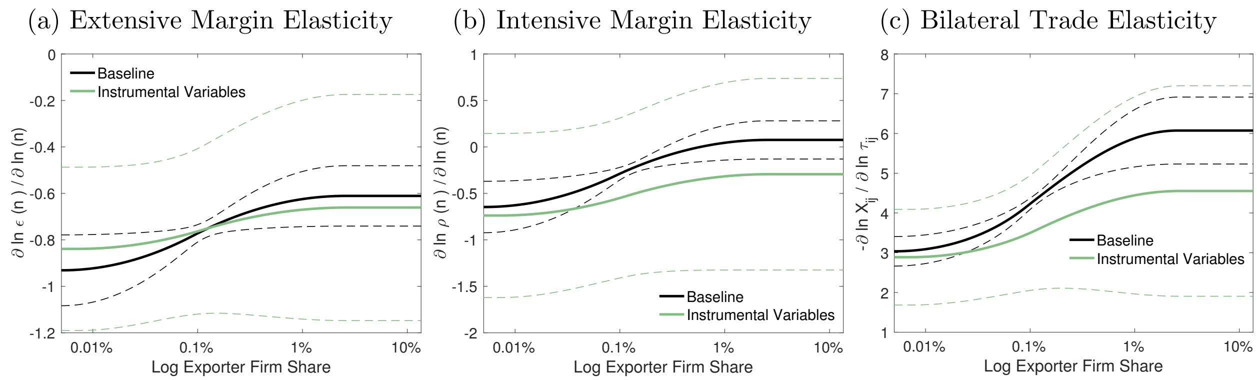
Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination. We assume that there are two groups of markets ($G = 2$) defined by whether the origin and destination have a free trade agreement and a common currency in panel (a), and the origin and destination have either a common language or colonial ties in panel (b). Solid lines are estimates of the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.11: Semiparametric Gravity of Firm Exports – Alternative Cost Pass-Through



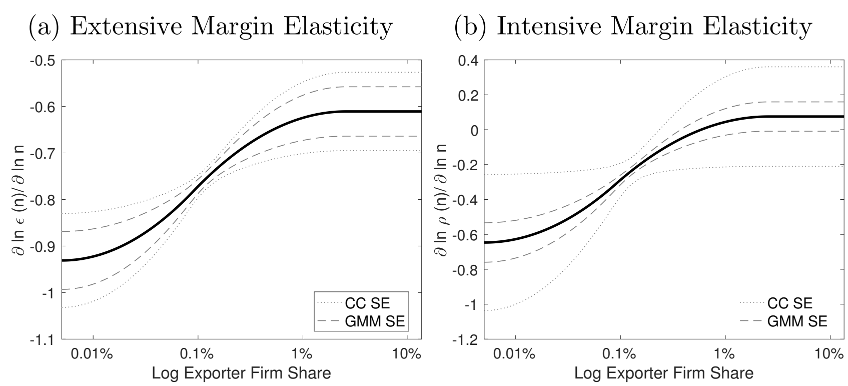
Note. Estimates obtained with GMM estimator in (37) the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). We report estimates assuming in panel (a) that the pass-through of tariffs to fixed costs is 0.5 ($\kappa^r = 1 - \sigma$ and $\kappa^e = \sigma - 0.5$), in panel (b) that pass-through of tariffs to fixed costs is 1 ($\kappa^r = 1 - \sigma$ and $\kappa^e = \sigma$), in panels (c) and d) that σ is respectively given by the 25th and 75th percentiles of the estimates in Redding and Weinstein (2024) ($\sigma = 2.4$ and $\sigma = 3.4$), and in panel (d) that $\sigma - 3.4$. Solid lines are estimates of the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta_g(n)$, and dashed lines are the associated 90% confidence intervals computed with robust standard errors.

Figure OA.12: Semiparametric Gravity of Firm Exports – Tariff IV



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). Here we use z_{ij}^{tariffIV} as instrument variable for bilateral import tariffs in z_{ij}. Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

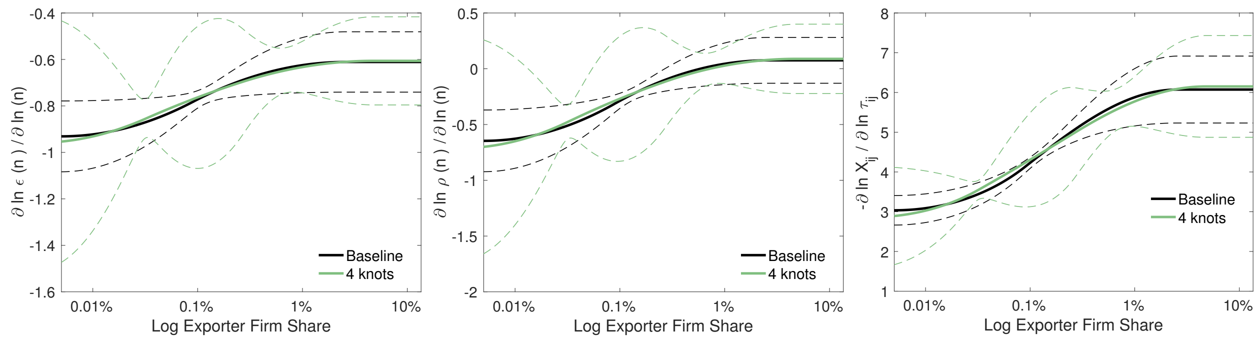
Figure OA.13: Semiparametric Gravity of Firm Exports – Alternative Inference



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with the nonparametric inference procedure of Chen and Christensen (2018).

Figure OA.14: Semiparametric Gravity of Firm Exports – Alternative Functional Form

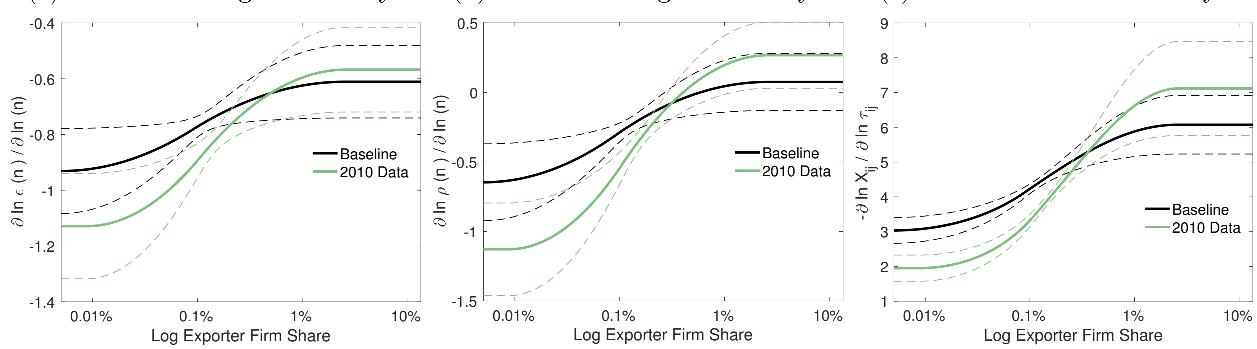
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for a single group pooling all pairs ($G = 1$). Estimates reported in green are based on Assumption 4 over five intervals ($M = 5$) instead of the three intervals imposed in the baseline specification shown in black ($M = 3$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Figure OA.15: Semiparametric Gravity of Firm Exports – Single Group, 2010

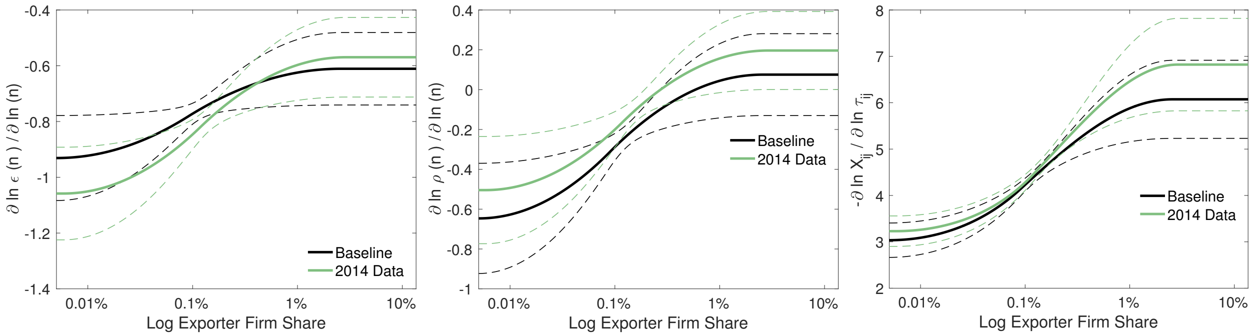
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2010 sample of 6,414 origin-destination pairs for a single group pooling all pairs ($G = 1$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Figure OA.16: Semiparametric Gravity of Firm Exports – Single Group, 2014

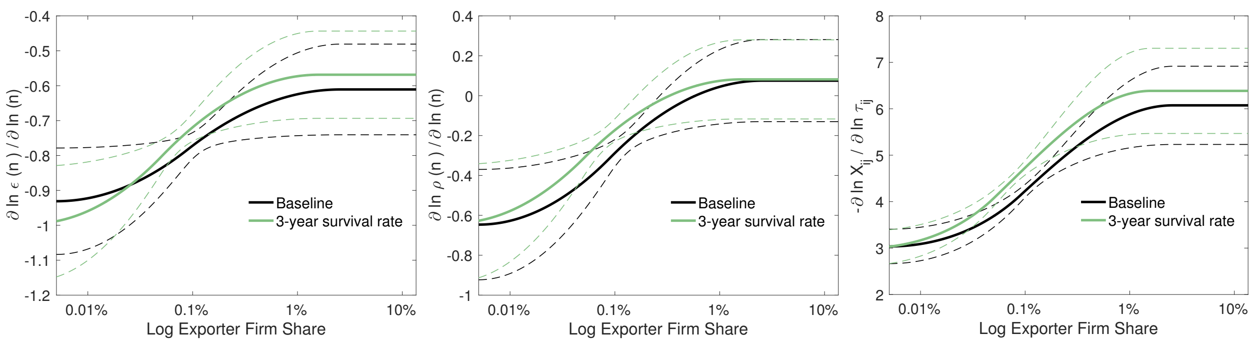
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2014 sample of 6,647 origin-destination pairs for a single group pooling all pairs ($G = 1$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Figure OA.17: Semiparametric Gravity of Firm Exports – 3-year Survival Rate n_{ii}

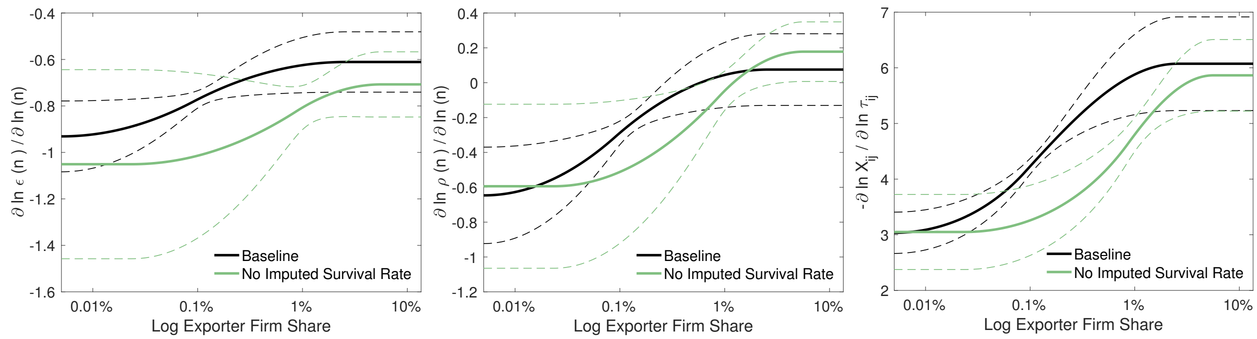
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the 2012 sample of 7,386 origin-destination pairs for single group pooling all pairs ($G = 1$). Measure of n_{ii} is the survival rate over three years instead of one year used in the baseline. Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

Figure OA.18: Semiparametric Gravity of Firm Exports – Dropping Observations with Imputed n_{ii}

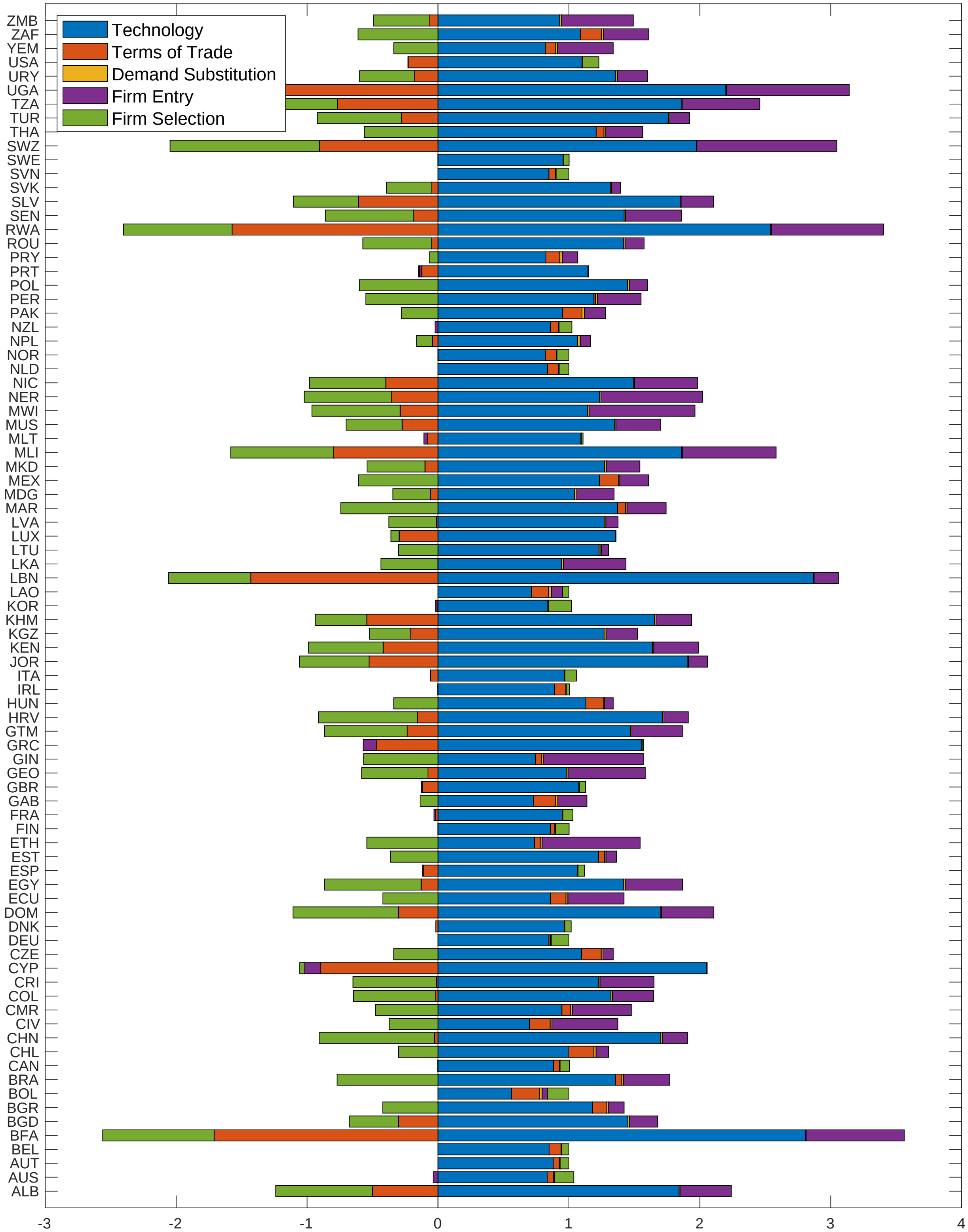
(a) Extensive Margin Elasticity (b) Intensive Margin Elasticity (c) Bilateral Trade Elasticity



Note. Estimates obtained with GMM estimator in (37) in the the 2012 subsample of 1,844 origin-destination pairs without imputed n_{ii} for single group pooling all pairs ($G = 1$). Panels (a) and (b) report respectively the elasticity of $\epsilon(n)$ and $\rho(n)$ with respect to the exporter firm share n, $\varepsilon(n)$ and $\varrho(n)$, and panel (c) reports the (absolute) elasticity of bilateral trade with respect to bilateral trade costs, $(\sigma - 1)\theta(n)$. Solid lines are the point estimates and dashed lines are the 90% confidence intervals computed with robust standard errors.

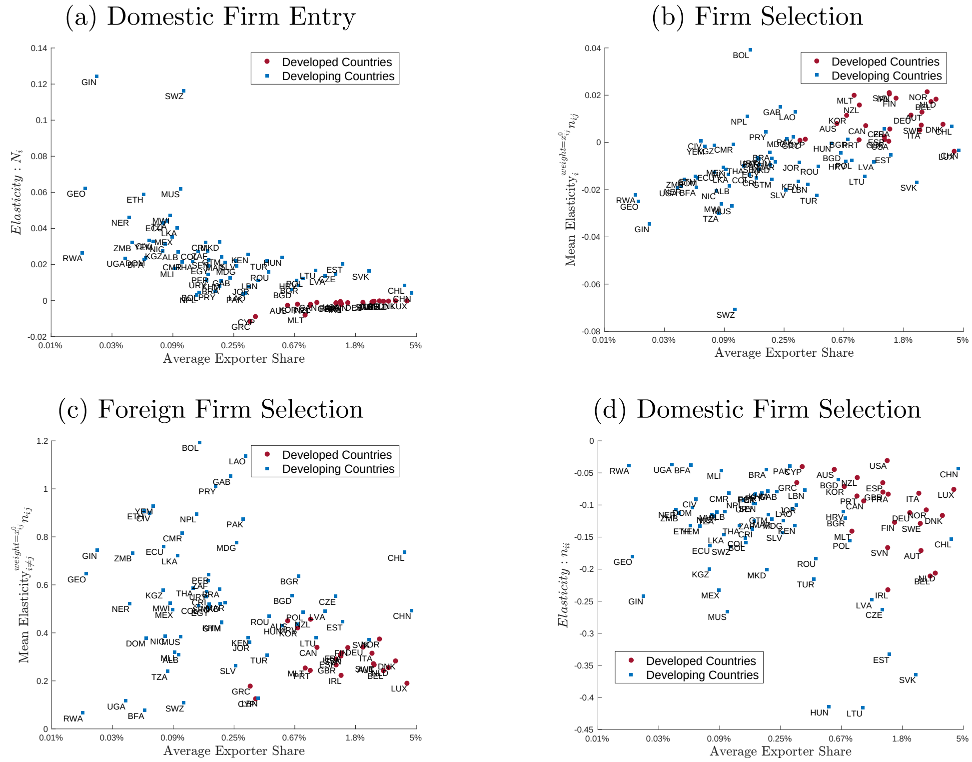
B.2 Counterfactual Analysis: Additional Results

Figure OA.19: Impact of a Uniform Reduction in Trade Costs on Welfare and its Components



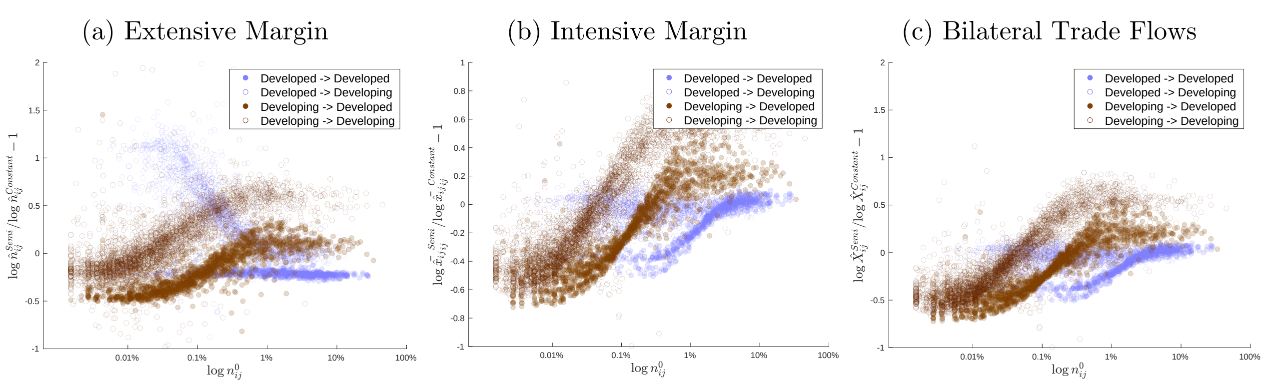
Note. Starting from the observed equilibrium in 2012, we compute the counterfactual equilibrium implied by a reduction of 1% in bilateral trade costs between all countries. For each group of countries, the size of the column denotes 100 times the log-change in real wage normalized by the shock size of 0.01. Each region of a row corresponds to a component of the welfare change in (26). Counterfactual predictions computed with Proposition 3.b, given the elasticity estimates reported in Figure 5.

Figure OA.20: Impact of a Uniform Reduction in Trade Costs on Firm Entry and Selection



Note. Starting from the observed equilibrium in 2012, we compute the counterfactual equilibrium implied by a reduction of 1% in bilateral trade costs between all countries. Each panel reports (100 times) changes in outcomes for a country against that country's log of the average firm exporter share in 2012. We report log-change of the mass of firms, $\ln \hat{N}_i$, in panel (a), of firm selection, $\sum_j x^0_{ij} \ln \hat{n}_{ij}$,in panel (b), of foreign firm selection, $\sum_{i \neq j} x^0_{ij} \ln \hat{n}_{ij} / \sum_{i \neq j} x^0_{ij}$, in panel (c), and of domestic firm selection, $\ln \hat{n}_{ii}$, in panel (d). Counterfactual predictions computed with Proposition 3.b, given the elasticity estimates reported in Figure 5.

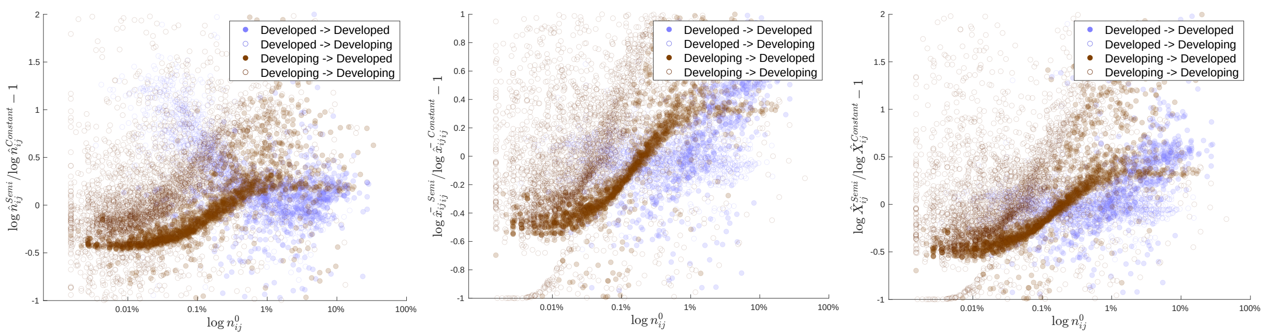
Figure OA.21: Impact of a Uniform Reduction in Trade Costs on Firm Export Margins: The Role of Parametric Assumptions



Note. For a reduction of 1% in bilateral trade costs between all countries starting from the observed equilibrium in 2012, the figure reports in the vertical axis is the ratio of the log-change of each margin of firm exports for an origin-destination pair predicted by the semiparametric and constant-elasticity specifications, and the horizontal axis is the log of the firm exporter share in 2012 for that origin-destination. Panel (a) does this for the extensive margin ($\ln \hat{n}_{ij}$), panel (b) for the intensive margin ($\ln \hat{\bar{x}}_{ij}$), and panel (c) for bilateral trade flows ($\ln \hat{X}_{ij}$). Semiparametric and constant-elasticity predictions use the elasticity estimates in Figures 5 and OA.2, respectively.

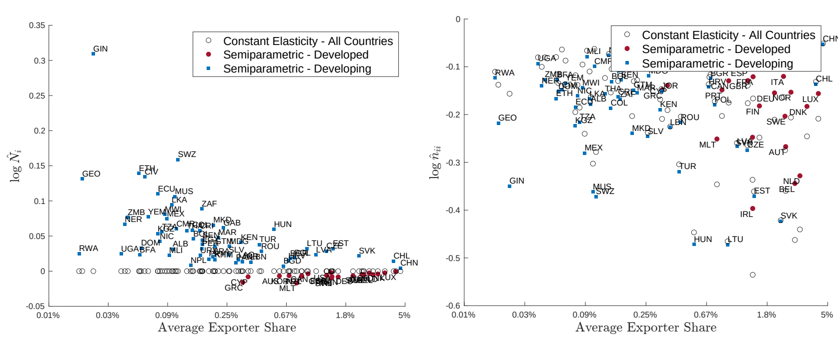
Figure OA.22: Impact of Reducing the Cost of Exporting from Developing to Developed Countries on Firm Export Margins: The Role of Parametric Assumptions

(a) Extensive Margin (b) Intensive Margin (c) Bilateral Trade Flows



Note. We consider the impact of a reduction of 1% in bilateral trade costs from developing countries in the GSP list to developed countries that concede preferential treatment to countries in the GSP list. The figure reports in the vertical axis is the ratio of the log-change of each margin of firm exports for an origin-destination pair predicted by the semiparametric and constant-elasticity specifications, and the horizontal axis is the log of the firm exporter share in 2012 for that origin-destination. Panel (a) does this for the extensive margin ($\ln \hat{n}_{ij}$), panel (b) for the intensive margin ($\ln \hat{\bar{x}}_{ij}$), and panel (c) for bilateral trade flows ($\ln \hat{X}_{ij}$). Semiparametric and constant-elasticity predictions use the elasticity estimates in Figures 5 and OA.2, respectively.

Figure OA.23: Gains from Trade: Entry and Selection of Domestic Firms

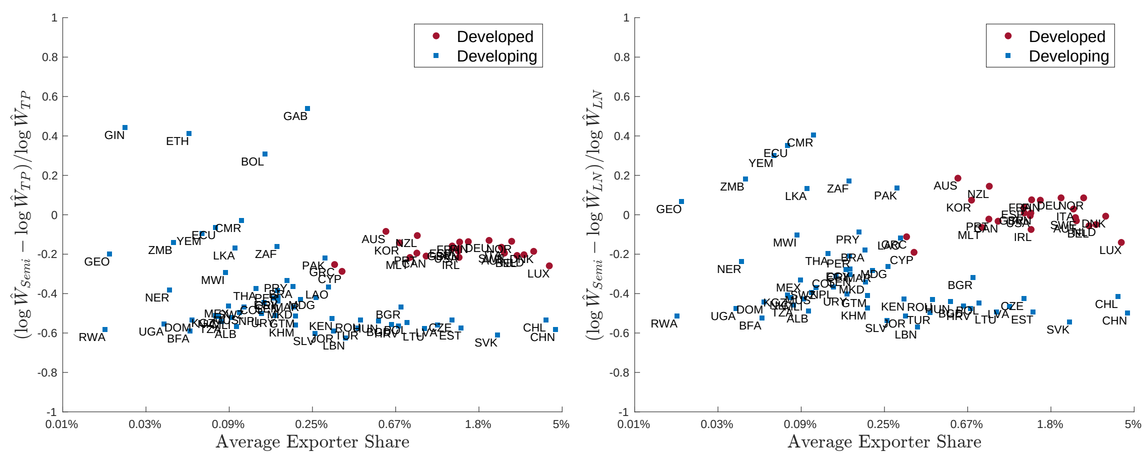


(a) Domestic Firm Entry (b) Domestic Firm Selection

Note. Figure reports the percentage change in the mass of firms $\hat{N}_i$ (panel a) and in the share of domestic firms $\hat{n}_{ii}$ (panel b) implied by moving from autarky to the observed equilibrium in 2012, computed with the formula in Corollary 1. White circles represent predictions obtained with the constant-elasticity estimates implied by (37) under (24). The blue dots and red diamonds represent predictions for developing and developed countries, respectively, that we obtain with the semiparametric estimates reported in Figure 5.

Figure OA.24: The Gains From Trade: Comparison to Other Parametric Forms

(a) Truncated Pareto (b) Log-Normal



Note. Figure reports the percentage change in the welfare gains from trade using truncated Pareto and log-normal productivity distributions. Panels (a) compares truncated Pareto to our semiparametric estimates. Panels (b) compares log-normal to our semiparametric estimatess. We report the functions obtained when the productivity distribution is truncated Pareto with cutoff parameter of $H = 2.85$, as in Melitz and Redding (2015), or log-normal, as in Head et al. (2014).