The Intensive Margin in Trade

Ana M. Fernandes
World Bank

Peter J. Klenow
Stanford and NBER

Sergii Meleshchuk
UC Berkeley

Martha Denisse Pierola
Inter-American Development Bank

Andrés Rodríguez-Clare *
UC Berkeley and NBER

May 4, 2017

Abstract

The Melitz model highlights the importance of the extensive margin (the number of firms exporting) for trade flows. Using the World Bank's *Exporter Dynamics Database* featuring firm-level exports from 50 countries, we find that around 50% of variation in exports does occur on the extensive margin — a quantitative victory for the Melitz framework. The remaining 50% on the intensive margin (exports per exporting firm) contradicts a special case of Melitz with Pareto-distributed firm productivity, which has become a tractable benchmark. This benchmark model predicts that, conditional on the fixed costs of exporting, *all* variation in exports across trading partners will occur on the extensive margin. Combining Melitz with lognormally-distributed firm productivity and firm-destination fixed trade costs can explain the intensive margin seen in the EDD data. In the EDD, the importance of the intensive margin rises steadily when going from the smallest to largest exporting firms across source countries, as is also predicted by the Melitz model with lognormally-distributed productivity.

*We are grateful to Caroline Freund, Keith Head, Thierry Mayer, Eduardo Morales and Jesse Perla for useful discussions, and to seminar participants at Berkeley, the Paris School of Economics, Stanford, the Minneapolis Fed, Princeton and the NBER ITI for comments. The World Bank provided access to the *Exporter Dynamics Database*. Research for this paper has in part been supported by the World Bank's Multidonor Trust Fund for Trade and Development and the Strategic Research Program on Economic Development and the Stanford Institute for Economic Policy Research. The findings in this paper are those of the authors and do not necessarily represent the views of the World Bank or its member countries.

1. Introduction

Across trading partners, exports can vary along the extensive margin (number of exporting firms) and the intensive margin (average exports per firm). The classic Krugman (1980) model predicts all export variation will be on the intensive margin because all firms export to every destination. Melitz (2003) brings the extensive margin to life with fixed costs of exporting, and emphasizes the importance of selection of firms into exporting. How important are the intensive and extensive margins empirically, and what does this tell us about the type of model that best fits the trade data?

Most firm-level empirical trade studies have only one or at most a few exporting countries. Bernard et al. (2007) decompose exports from the U.S. to other countries. Eaton et al. (2008) analyze firm-level exports for Colombia, Eaton et al. (2011) do so for France, Eaton et al. (2012a) for Denmark and France, Manova and Zhang (2012) for China, and Arkolakis and Muendler (2013) for Brazil, Chile, Denmark and Norway.

We use the World Bank's *Exporter Dynamics Database* (hereafter EDD) to systematically examine the importance of the extensive and intensive margins. The EDD covers firm-level exports from 59 (mostly developing) countries to all destination countries in most years from 2003 to 2013. For 49 of the countries, every exporting firm's exports to each destination in a given year can be broken down into products at the HS 6-digit level.[1] We add China to the EDD set of 49 countries to arrive at 50 countries for our analysis. Having many origin and destination countries in our dataset allows us to study the role of the intensive and extensive margins while allowing for origin-year and destination-year fixed effects that control for differences in population, wages, and other country characteristics that affect firm entry into exporting and exports per firm.

We find that between 40 and 60 percent of the variation in overall exports across origin-destination pairs is accounted for by the intensive margin, with the rest accounted for by the extensive margin. This breakdown into the intensive and extensive margin is robust to using different country samples or sets of fixed effects, excluding country pairs with few exporters or tiny exporters, and looking within industries. If we place exporting firms into percentiles for each trading pair and look across pairs, the importance of the intensive margin in explaining overall exports rises steadily from around 20 percent

[1] Fernandes et al. (2016) describe the dataset in detail.

for the smallest exporters to over 50 percent for the largest exporters.

We interpret the finding that up to 60 percent of the variation in bilateral trade flows are explained by the extensive margin as providing support for the Melitz (2003) model. But the finding that at least 40 percent of that variation is explained by the intensive margin even while allowing for origin-year and destination-year fixed effects contradicts an important special case of the Melitz model, the case with Pareto-distributed firm productivity and fixed trade costs that vary only because of separate origin and destination components. Melitz-plus-Pareto has a sharp prediction: conditional on the fixed costs of exporting, all variation in exports across trade partners should occur through the number of exporters (the extensive margin). Lower variable trade costs should stimulate sales of a given exporting firm, but draw in marginal exporting firms to the point that average exports per exporter (the intensive margin) is unchanged. This exact offset is a special property of the Pareto distribution. It is not so dependent on other aspects of the Melitz model.[2]

The upshot of our EDD facts could simply be that one needs to combine Melitz with a firm productivity distribution other than Pareto. But Melitz-Pareto has become a useful and tractable benchmark in international trade. It is consistent with many firm-level facts (Eaton et al., 2011), generates a gravity equation (Chaney, 2008), and yields a simple summary statistic for the welfare gains from trade (Arkolakis et al., 2012). We therefore explore whether it can be rescued before moving beyond Pareto.

We explore several potential explanations for the positive intensive margin elasticity (the tendency of exports per firm to rise along with overall exports) in the EDD data while retaining a Melitz-Pareto core. First, we consider the possibility that fixed trade costs vary by origin-destination pair. Higher fixed trade costs raise average exports per exporter, but also *lower* overall exports. For the intensive margin to be increasing in overall exports, one therefore needs variable trade costs to be very negatively correlated with fixed trade costs. A corollary is that, whereas variable trade costs rise decisively with distance between trade partners, fixed trade costs would need to fall with distance between trade partners. In this explanation, however, the intensive margin elasticity

[2]The full dominance of the extensive margin extends to some environments with firm-destination idiosyncratic demand and fixed costs (Eaton et al., 2011), convex marketing costs (Arkolakis, 2010), non-CES preferences (Arkolakis et al., 2015), non-monopolistic competition (Bernard et al., 2003), and multinational production (Arkolakis et al., 2014).

would be equally important for the smallest and largest exporting firms, contrary to what we see in the data, where the importance of the intensive margin rises steadily with exporting firm size.

Second, we explore the role of multi-product firms. If the typical firm exports more products to destinations with larger overall flows, this could account for the importance of the intensive margin for exports. We find that the number of HS 6-digit products per exporting firm does indeed account for about 12 percent of the variation in overall exports, or about one-fourth of the contribution of the intensive margin to overall exports. In the context of the multi-product Melitz-Pareto model developed by Bernard, Redding and Schott (2011), however, this explanation still requires a negative correlation between *firm-level* fixed costs of exporting and variable trade costs, and for *firm-level* fixed costs to fall with the distance between trading partners. Moreover, the significant intensive margin elasticity per firm-product implies that fixed costs of exporting *per product* also fall with distance.

A third hypothesis we investigate is granularity — a finite number of firms. With a finite number of firms, the intensive margin (and overall exports) can be high because of favorable productivity draws from the Pareto distribution within a country. We develop an estimator for the elasticity of fixed trade costs to distance that is valid under granularity as in Eaton et al. (2012a), and continue to find that fixed trade costs must fall with distance to explain a positive intensive margin elasticity. Using simulations of finite draws from a Pareto distribution, we find that granularity generates only a modest intensive margin elasticity, and — in contrast to what we observe in the data — almost entirely in the right tail of the exporter size distribution.

After these failed attempts to rescue the Melitz-Pareto model, we depart from the comforts of that model and consider a lognormal distribution of firm productivity. Head et al. (2014) analyze how the welfare gains from trade in the Melitz model differ with a lognormal instead of a Pareto distribution. Bas et al. (2015) show how the trade elasticity varies with a lognormal distribution. Both papers marshal evidence from firms in France and China pointing to the empirical relevance of the lognormal distribution.[3]

[3]This may seem surprising in light of the evidence in Axtell (2001) for a Pareto distribution of U.S. firm sizes. Moreover, Gabaix (2009) emphasizes that a Pareto distribution emerges naturally from random growth and some extensions. But Rossi-Hansberg and Wright (2007) find thinner-than-Pareto tails of the firm and (especially) plant size distributions in the U.S. And Luttmer (2011) shows that the largest U.S.

As in Eaton et al. (2011), we consider a Melitz model with firm-destination-specific demand and fixed cost shocks, but we assume that the firm productivity distribution is lognormal rather than Pareto. In particular, we assume that each firm is characterized by a productivity parameter as well as an idiosyncratic demand shifter and fixed cost for each destination market, all drawn from a multivariate lognormal distribution. We allow for a non-zero covariance between the demand shifter and the fixed cost in each destination, but set all other covariances to zero. One appealing feature of the model is that it is amenable to maximum-likelihood methods. As the likelihood is potentially not concave as a function of the parameters, and since we have a large number of parameters to estimate (means, variances, covariance, and trade costs), we rely on the estimation methodology proposed by (Chernozhukov and Hong, 2003).

Our estimation shows that a lognormal distribution for poductivity can indeed generate a sizable intensive margin elasticity. When variable trade costs fall and fixed costs are constant, the ratio of mean to minimum exports increases as the productivity cutoff falls under the lognormal distribution (while being constant under Pareto).[4] Shifting to lognormal productivity also changes our inference about fixed trade costs, rendering them positively correlated with variable trade costs and *rising* with distance. As in the data, the intensive margin elasticity rises steadily with the size percentile of exporters under a lognormal productivity distribution. We further show that lognormally distributed demand shocks and fixed cost shocks can contribute to fitting the intensive margin facts.

The rest of the paper is organized as follows. In Section 2 we describe the EDD data and document the empirical importance of the extensive and intensive margins in accounting for cross-country variation in exports. Section 3 contrasts the predictions of the Melitz-Pareto model (with a continuum of single product firms, multi-product firms or a finite number of firms) to the EDD facts. Section 4 shows how the implications of the Melitz model change when we drop the Pareto assumption and instead assume that the firm productivity distribution is lognormal. Section 5 concludes.

firms are far too young to emerge from random growth. A lognormal distribution, meanwhile, can arise from random growth innovations, albeit with exploding variance without mean reversion in levels.

[4]The result holds under other thin-tailed productivity distributions, such as bounded Pareto as in Feenstra (2014). But a bounded Pareto distribution loses all the analytical convenience of the unbounded Pareto while lacking the empirical convenience of the lognormal distribution.

2. The Intensive Margin in the Data

The Exporter Dynamics Database

We use the Exporter Dynamics Database (EDD) described in Fernandes et al. (2016) to study the intensive and extensive margins of trade. The EDD is based on firm-level customs data covering the universe of export transactions provided by customs agencies from 59 countries (53 developing and 6 developed countries). For each country, the raw firm-level customs data contains annual export flows (in values) disaggregated by firm, destination and Harmonized System (HS) 6-digit product. Oil exports are excluded from the customs data due to lack of accurate firm-level data for many of the oil-exporting countries. For most countries total non-oil exports in the EDD are close to total non-oil exports reported in COMTRADE/WITS. More than 100 statistics from the EDD are publicly available at the origin-year, origin-product-year, origin-destination-year, or origin-product-destination-year levels. These include average exports per firm as well as the number of exporting firms.

For the descriptive analysis in this section as well as for the regression and simulation work in the sections that follow we focus on a *core sample* that consists of 50 countries (49 from the EDD and China) for which we have the firm-level data.[5] However, to use the most comprehensive sample of countries available we rely for the motivating plots below on an *extended sample* that includes the 59 origin countries from the EDD plus China. Both samples cover a subset of years from 2003 and 2013 — see Table 1 and Table **??** in the Online Appendix.

We focus on EDD statistics based on products belonging only to the manufacturing sector. Specifically, using a concordance across the ISIC rev. 3 classification and the HS 6-digit classification, we consider only exports of HS 6-digit products that correspond to ISIC manufacturing sub-sectors 15-37. Using these data we calculate variants of average exports per firm, number of exporting firms, and total exports at the origin-destination-year level or at the origin-product-destination-year level. The product disaggregations that we use are HS 2-digit for the extended sample and HS 2-digit, HS 4-digit, or HS 6-digit for the core sample.

[5] China is not included in the EDD due to confidentiality concerns.

Importance of the intensive margin

Let X_{ij}, N_{ij} and $x_{ij} \equiv X_{ij}/N_{ij}$ denote total exports, total number of firms exporting, and average exports per firm from country i to country j, respectively. In Figure 1 we plot the intensive margin $(\ln x_{ij})$ and extensive margin $(\ln N_{ij})$ vs. total exports $(\ln X_{ij})$ for the extended sample of countries. We restrict the sample to the origin-destination pairs with more than 100 exporting firms (i.e., ij pairs for which $N_{ij} > 100$) to reduce noise associated with country pairs with few exporting firms.[6] All variables plotted are demeaned of origin-year and destination-year fixed effects. Each dot corresponds to $(\ln x_{ij}, \ln X_{ij})$ (Panel A) or $(\ln N_{ij}, \ln X_{ij})$ (Panel B). The red lines can be ignored for now.

A key statistic that we use to summarize the pattern observed in Figure 1 is the *intensive margin elasticity* (IME), which is the slope of the regression line in Panel A. In a given year, the IME can be obtained from an OLS regression of $\ln x_{ij}$ on $\ln X_{ij}$ with origin and destination fixed effects:

$$\ln x_{ij} = FE_i^o + FE_j^d + \alpha \ln X_{ij} + \varepsilon_{ij}. \tag{1}$$

The IME is the estimated regression coefficient

$$\hat{\alpha} = \frac{cov(\ln \tilde{x}_{ij}, \ln \tilde{X}_{ij})}{var\left(\ln \tilde{X}_{ij}\right)}, \tag{2}$$

where we write $\ln \tilde{z}_{ij}$ to denote variable $\ln z_{ij}$ demeaned by origin-year and destination-year fixed effects. The complement of the IME is the *extensive margin elasticity*, defined as EME $\equiv \frac{cov(\ln \tilde{N}_{ij}, \ln \tilde{X}_{ij})}{var\left(\ln \tilde{X}_{ij}\right)}$. The EME corresponds to the slope of the regression line in Panel B of Figure 1 and satisfies EME $= 1 -$ IME.

Figure 1 demonstrates that both the IME and the EME are positive and large. As shown in Panel A of Table 2, depending on the type of fixed effects included, the IME ranges from 0.4 to 0.46 in the core sample that we will use for the analysis in the next two sections. Our preferred estimate of the IME is 0.4 based on the inclusion of origin-year and destination-year fixed effects (as in Figure 1).[7] In this estimate, the intensive

[6] The core sample includes 1291 unique country pairs with $N_{ij} > 100$ while the extended sample includes 2075 unique country pairs with $N_{ij} > 100$.

[7] To be specific, the equation estimated in this case is $\ln x_{ijt} = FE_{it}^o + FE_{jt}^d + \alpha \ln X_{ijt} + \varepsilon_{ijt}$ using all years of available data for the country pairs included in the core sample.

margin accounts for approximately 40% of the variation in total exports across country pairs, while 60% is accounted for by the extensive margin. As the focus has so far been on accounting for the variation of bilateral trade flows while controlling for origin-year and destination-year fixed effects, it is natural to wonder how much of that variation is absorbed by the fixed effects alone. The results in Table 2 show that this is never more than 59 percent, implying that a large share of the variation in bilateral trade flows comes from the forces behind the estimated IME.[8]

Robustness

The finding of a positive and large IME is robust to considering different samples. In Panel B of Table 2 we estimate the IME including all country pairs — even those with less than 100 exporting firms. The IME in this case reaches 0.59 when origin-year and destination-year fixed effects are included. In the Online Appendix Table **??** we reproduce the regressions in Table 2 but now for the extended sample of countries. In the preferred specification with origin-year and destination-year fixed effects, the IME is 0.38 among origin-destination pairs with at least 100 exporting firms and 0.52 among all origin-destination pairs. To make sure the IME is not driven by small exporting firms, we re-estimate it after excluding firms whose annual exports fell below $1,000 in any year. The corresponding IME estimates in Table 3 (core sample) and Online Appendix Table **??** (extended sample) change only slightly.

A separate concern is measurement error. Since total exports is the sum of firm-level exports, classical measurement error in exports per exporter x would bias the IME upward, but classical measurement error in the number of exporters N would bias the IME downward. Depending on their relative importance compared to the true IME, classical measurement error could bias the IME upward or downward. If the measurement error is serially uncorrelated, then instrumenting with leads and/or lags should yield an unbiased estimate of the IME. As shown in Online Appendix Table **??**, the instrumented IME's are very close to the OLS IME, both economically and statistically.

Our results for the IME could be coming from country differences in industry composition of exports combined with industry differences in average exports per firm. In

[8]This percentage comes from the R-squared of an OLS regression of bilateral total exports in logs ($\ln X_{ijt}$) on origin-year and destination-year fixed effects.

Figure 2 we plot the (demeaned) intensive and extensive margins against total exports at the origin-industry-destination-year level using HS 2-digit industries. The pattern here is similar to that in Figure 1. Table 4 shows that the IME actually increases when moving to industry-level data. At the lowest level of aggregation available (HS 6-digit), for the core sample of countries the IME is 0.51 with origin-year-industry and destination-year-industry fixed effects. The results also hold in the extended sample, for which we calculated IME disaggregated at HS2 product level. As reported in the Online Appendix Table **??**, this IME is also close to 0.52.

IME by percetiles

A positive IME could be due to the presence of export superstars that increase both average exports per firm and total exports for some country pairs, as discussed in Freund and Pierola (2015). We study this possibility by considering separate IME regressions for each exporter size percentile. For each origin-destination-year combination we distribute the exporting firms into percentiles based on their exports. Denoting average exports per firm in percentile pct as x_{ij}^{pct}, we run the regressions:

$$\ln x_{ij}^{pct} = FE_i^o + FE_j^d + \alpha^{pct} \ln X_{ij} + \epsilon_{ij}.$$

We define the IME for each percentile as $\text{IME}^{pct} \equiv \hat{\alpha}^{pct}$.

We plot the IME^{pct} for each percentile (with confidence intervals) in Figure 3 along with the red line at the overall IME of 0.4.[9] The IME is 0.5 for the highest percentile. But the positive overall IME is not coming exclusively from the export superstars: the IME^{pct} rises steadily from 0.2 at the 50th percentile to 0.3 at the 80th percentile.

IME for multi-product firms

We can dig deeper and study whether average exports per firm can be explained by the number of products exported per firm or by exports per product per firm. Let m_{ij} be the average number of products exported from i to j by firms exporting from i to j, and let $x_{ij}^p \equiv x_{ij}/m_{ij}$ be the average exports per product per firm exporting from i to j. We define the IME *at the product level* as $\text{IME}^p \equiv cov(\ln \tilde{x}_{ij}^p, \ln \tilde{X}_{ij})/var(\ln \tilde{X}_{ij})$. Since $x_{ij} = x_{ij}^p m_{ij}$,

[9]For exporter percentiles to be well-defined we focus on country pairs for which $N_{ij} > 100$.

the IME is equal to the IME^p plus the extensive product margin elasticity,

$$\text{IME} = \text{IME}^p + \frac{cov(\ln \tilde{m}_{ij}, \ln \tilde{X}_{ij})}{var(\ln \tilde{X}_{ij})}.$$

Table 5 reports the results for the IME^p for the core sample.[10] Most of the IME is explained by the systematic variation in average exports per product per firm, rather than in the average number of products exported by firm.

Taking stock: the IME in the EDD

Summarizing the results so far, we find the intensive margin elasticity to be positive and significant, both statistically and economically. This finding is robust to the inclusion of a variety of fixed effects, various samples, exclusion of small firms, and disaggregation by industry. The IME is positive and monotonically increasing across the whole distribution of exporter size. The systematic cross-country-pair variation of average exports per firm comes primarily from the behavior of average exports per product per firm.

Correlation between intensive and extensive margin, and relation with distance

We now move beyond the intensive margin elasticities and report additional stylized facts on the correlations between the intensive margin, the extensive margin, and distance. There is a positive and significant correlation between average exports per firm and the number of exporting firms (0.25, standard error 0.01) after taking out origin-year and destination-year effects. Table 6 shows how these margins vary with log distance with alternative sets of fixed effects. The elasticities are all negative and significant when controlling for origin-year and destination-year fixed effects: average exports per firm, the number of firms, average number of products exported per firm, and average exports per product per firm all decline with distance between trade partners.

[10] Bernard et al. (2009) present a similar decomposition for U.S. exports. We compare their results to ours below.

Relation to previous empirical results

We finish this section by relating our stylized facts to those of EKK, EKS, Bernard et al. (2007) and Bernard et al. (2009). EKK use firm-level export data for a single origin (France) and show that average exports per firm increase with market size of the destination (measured as manufacturing absorption) with an elasticity of 1/3. In Figure 4 we plot market size (horizontal axis) against our estimated destination fixed effects (vertical axis) from a regression of average exports per firm on origin-year and destination-year fixed effects based on the extended sample and country pairs for which $N_{ij} > 100$. A regression line through the points in the plot implies that average exports per firm increase with destination market size with an elasticity of 0.19, a bit lower than the result in EKK.[11]

EKK also show that firms exporting to more destinations exhibit higher sales in the domestic (French) market. Our data does not include domestic sales, but we can instead look at sales in the most popular destination market for each origin. Let $x_{il|j}$ denote average exports to destination l computed across firms from i that sell in markets l and j and let $l^*(i) \equiv \arg\max_k N_{ik}$ be the largest destination market for each origin country i (e.g., the United States for Mexico). In Figure 5 we plot $\log \frac{x_{il^*(i)|j}}{x_{il^*(i)|l^*(i)}}$ (vertical axis) against $\log \frac{N_{ij}}{N_{il^*(i)}}$ (horizontal axis) for all i and j for the core sample.[12] It is very clear that the results derived by EKK for French firms remains valid for our data with many origin countries: firms that sell in more markets are more productive as proxied by their sales in their origin country's most popular destination market.

EKS find that average exports per firm are very similar across four origin countries (Brazil, Denmark, France and Uruguay) for which they have customs data. They regress average exports per firm on origin and destination fixed effects and find that the origin fixed effects differ little across their four origins. Running the same regression in our dataset (but pooling across years and including year fixed effects), we find that origin fixed effects do vary significantly across countries (the coefficient of variation in the estimated origin fixed effects ranges from 0.81 to 2.56, depending on the sample used) and

[11] Similar findings are obtained in unreported plots where the destination fixed effects are based on the extended sample and all country pairs or based on the core sample and either country pairs for which $N_{ij} > 100$ or all country pairs.

[12] The EKK estimating sample includes only firms with sales in France. To implement an approach comparable to theirs, we drop all firms from country i that do not sell to $l^*(i)$, so the sample includes only $N_{il^*(i)}$ firms for country i. This implies that all firms that make up N_{ij} are also selling to $l^*(i)$.

are higher for countries with higher GDP per capita and higher total exports.[13] Moreover, origin-year and destination-year fixed effects are not enough to capture the variation in $\ln x_{ij}$: a regression of $\ln x_{ij}$ on origin-year and destination-year fixed effects yields an R-squared of 0.65 when only country pairs with $N_{ij} > 100$ are considered and only 0.37 when all country pairs are considered.

Using firm-level export data for the United States, Bernard et al. (2009) present a similar decomposition to the one we present above for multi-product firms, except that they cannot allow for destination fixed effects because their data is for a single origin. They find that IME^p is around 0.23, which is not far from our finding of around 0.29. On the other hand, contrary to our results, Bernard et al. (2007) find that average exports per product per firm increase with distance. We believe that the difference arises from the fact that, by having data for multiple origins, we are able to control for destination fixed effects. In fact, Table 6 shows that regressing $\ln x_{ij}^p$ on $\ln dist_{ij}$ with only origin and year fixed effects but without destination fixed effects yields a positive and significant coefficient as in Bernard et al. (2007), whereas the coefficient becomes negative and significant when destination fixed effects are added. The same happens when regressing $\ln x_{ij}$ on $\ln dist_{ij}$.

3. The Intensive Margin in the Melitz-Pareto Model

In this section we ask how the Melitz model with Pareto distributed productivity stacks up relative to the findings of the previous section. We focus on the implications of this model for the intensive margin elasticity. We start with the simplest model, which entails a continuum of single-product firms with a Pareto distribution for productivity as in Chaney (2008) and Arkolakis et al. (2008). We derive a series of properties of this model, and then explore their robustness to allowing for destination-specific demand and fixed trade cost shocks at the firm level as in EKK, for multi-product firms, and for granularity.

3.1. The Basic Melitz-Pareto Model

[13] For this purpose, we run regressions of the estimated origin fixed effects on population, GDP, GDP per capita, and total exports, jointly and separately.

Theory

As this is a well-known model, we will be brief in the presentation of the main assumptions. There are many countries indexed by i, j. Labor is the only factor of production available in fixed supply L_i in country i and the wage is w_i. Preferences are constant elasticity of substitution (CES) with elasticity of substitution across varieties $\sigma > 1$. Each firm produces one variety under monopolistic competition. In each country i there is a large pool of firms of measure N_i with productivity φ distributed Pareto with shape parameter $\theta > \sigma - 1$ and scale parameter b_i, $\Pr(\varphi \leq \varphi_0) = G_i(\varphi_0) = 1 - (\varphi_0/b_i)^{-\theta}$. Firms from country i also incur fixed trade costs F_{ij} as well as iceberg trade costs τ_{ij} to sell in country j.[14]

Sales in destination j by a firm from origin i with productivity φ are

$$x_{ij}(\varphi) = A_j \left(\bar{\sigma}\frac{w_i \tau_{ij}}{\varphi}\right)^{1-\sigma}, \tag{3}$$

where $A_j \equiv P_j^{1-\sigma} w_j L_j$, $P_j^{1-\sigma} = \sum_i N_i \int_{\varphi \geq \varphi_{ij}^*} \left(\bar{\sigma}\frac{w_i \tau_{ij}}{\varphi}\right)^{1-\sigma} dG_i(\varphi)$ is the price index in j, $\bar{\sigma} \equiv \sigma/(\sigma-1)$ is the markup, and φ_{ij}^* is the productivity cutoff for exports from i to j, which is determined implicitly by

$$x_{ij}(\varphi_{ij}^*) = \sigma F_{ij}. \tag{4}$$

The value of overall exports and the number of firms that export from i to j are then $X_{ij} = N_i \int_{\varphi \geq \varphi_{ij}^*} x_{ij}(\varphi) dG_i(\varphi)$ and $N_{ij} = N_i \int_{\varphi \geq \varphi_{ij}^*} dG_i(\varphi)$, respectively. Using again the fact that $G_i(\varphi)$ is Pareto and assuming that $\varphi_{ij}^* > b_i$ for all i, j, we get that

$$X_{ij} = \left(\frac{\theta}{\theta - (\sigma - 1)}\right) A_j (w_i \tau_{ij})^{1-\sigma} b_i^\theta N_i \left(\varphi_{ij}^*\right)^{\sigma-\theta-1} \tag{5}$$

and

$$N_{ij} = b_i^\theta N_i \left(\varphi_{ij}^*\right)^{-\theta}. \tag{6}$$

[14] F_{ij} is in units of the numeraire. Since we focus on cross-section properties of the equilibrium, we do not need to specify whether the fixed trade cost entails hiring labor in the origin or the destination country.

Combining (4) , (5) and (6), the extensive margin is

$$N_{ij} = N_i \left(\frac{w_i}{b_i}\right)^{-\theta} \left(\frac{\sigma}{A_j}\right)^{-\theta/(\sigma-1)} \tau_{ij}^{-\theta} F_{ij}^{-\theta/(\sigma-1)}, \tag{7}$$

while the intensive margin is

$$x_{ij} \equiv \frac{X_{ij}}{N_{ij}} = \left(\frac{\theta\sigma}{\theta - (\sigma - 1)}\right) F_{ij}. \tag{8}$$

We can always decompose variable and fixed trade costs as follows: $\tau_{ij} = \tau_i^o \tau_j^d \widetilde{\tau}_{ij}$ and $F_{ij} = F_i^o F_j^d \widetilde{F}_{ij}$. Taking logs in (7) and (8), and defining variables appropriately, we have

$$\ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \widetilde{\tau}_{ij} - \bar{\theta} \ln \widetilde{F}_{ij} \tag{9}$$

and

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \widetilde{F}_{ij}, \tag{10}$$

where $\bar{\theta} \equiv \frac{\theta}{\sigma-1}$. These are the two key equations that we use to derive the results in the rest of this section.

Combining the definition of the intensive margin elasticity given in the previous section (i.e., IME $= \frac{cov(\ln \tilde{x}_{ij}, \ln \tilde{X}_{ij})}{var\left(\ln \tilde{X}_{ij}\right)}$) with equations (9) and (10), the model implies that

$$\text{IME} = \frac{-\left(\bar{\theta} - 1\right) var(\ln \widetilde{F}_{ij}) - \theta cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})}{var\left(-\theta \ln \widetilde{\tau}_{ij} - \left(\bar{\theta} - 1\right) \ln \widetilde{F}_{ij}\right)}. \tag{11}$$

This result can be used to extract several implications of the model, which we present in the form of four observations in the rest of this section.

Our first observation says that if all variation in fixed trade costs comes from origin and destination fixed effects with no country-pair component, for example because $F_{ij} \propto w_i^{\gamma} w_j^{1-\gamma}$ (as in Arkolakis (2010)), then the model implies that the intensive margin elasticity is zero.

Observation 1: If $var\left(\ln \widetilde{F}_{ij}\right) = 0$ then IME $= 0$.

Since this is a key result, it is worth understanding it in more detail. Using equations (3) and (4) together with the definition of x_{ij}, taking logs and differentiating w.r.t. $\ln \tau_{ij}$

we get

$$\frac{d\ln x_{ij}}{d\ln\tau_{ij}} = 1-\sigma - \frac{d\ln\left(1-G_i(\varphi_{ij}^*)\right)}{d\ln\varphi_{ij}^*}\left(1-\frac{x_{ij}(\varphi_{ij}^*)}{x_{ij}}\right).$$

The first term is the direct effect on incumbent firms, while the second term captures selection. In turn, selection is the product of $-\frac{d\ln\left(1-G_i(\varphi_{ij}^*)\right)}{d\ln\varphi_{ij}^*}$, which captures the effect of τ_{ij} (and hence φ_{ij}^*) on average exports per firm through its impact on the share of firms that export, and $\left(1-\frac{x_{ij}(\varphi_{ij}^*)}{x_{ij}}\right)$, which captures how much less firms export at the cutoff relative to the average. Obviously, if $x_{ij} = x_{ij}(\varphi_{ij}^*)$ then there is no selection, while the effect of selection is maximized if $x_{ij}(\varphi_{ij}^*)/x_{ij} = 0$. With a Pareto distribution for productivity we have $-\frac{d\ln\left(1-G_i(\varphi_{ij}^*)\right)}{d\ln\varphi_{ij}^*} = \theta$ and $\frac{x_{ij}}{x_{ij}(\varphi_{ij}^*)} = \frac{\theta}{\theta-(\sigma-1)}$, therefore $\frac{d\ln x_{ij}}{d\ln\tau_{ij}} = 0$.

Combined with the assumption that $\bar{\theta} > 1$, the result in equation (11) also implies that if the intensive margin elasticity is positive then there must be a negative correlation between the variable and fixed trade costs (ignoring origin and destination fixed costs).

Observation 2: If IME > 0 then $corr(\ln\widetilde{F}_{ij}, \ln\widetilde{\tau}_{ij}) < 0$.

Ignoring origin and destination fixed effects, equation (10) implies that

$$cov(\ln\widetilde{F}_{ij}, \ln\widetilde{dist}_{ij}) = cov(\ln\widetilde{x}_{ij}, \ln\widetilde{dist}_{ij}).$$

Thus, if average exports per firm fall with distance then fixed trade costs must also fall with distance. This is captured formally by our third observation which is related to the fixed trade costs elasticity with respect to distance.

Observation 3: $\frac{cov(\ln\tilde{x}_{ij},\ln\widetilde{dist}_{ij})}{var(\ln\widetilde{dist}_{ij})} = \frac{cov(\ln\widetilde{F}_{ij},\ln\widetilde{dist}_{ij})}{var(\ln\widetilde{dist}_{ij})}$.

We can go beyond the previous qualitative observations and derive the fixed and variable trade costs implied by the model so as to compute actual values for $corr(\ln\widetilde{F}_{ij}, \ln\widetilde{\tau}_{ij}) < 0$ and $\frac{cov(\ln\widetilde{F}_{ij},\ln\widetilde{dist}_{ij})}{var(\ln\widetilde{dist}_{ij})}$. Combining equations (9) and (10) to solve for $\ln\widetilde{F}_{ij}$ and $\ln\widetilde{\tau}_{ij}$ in terms of $\ln x_{ij}$ and $\ln N_{ij}$ yields

$$\ln\widetilde{F}_{ij} = \delta_i^{F,o} + \delta_j^{F,d} + \ln x_{ij} \tag{12}$$

and

$$\theta\ln\widetilde{\tau}_{ij} = \delta_i^{\tau,o} + \delta_j^{\tau,d} - \bar{\theta}\ln x_{ij} - \ln N_{ij}. \tag{13}$$

Model-implied values for $\ln \widetilde{F}_{ij}$ are (ignoring origin and destination fixed effects) directly given by $\ln x_{ij}$, but for $\ln \widetilde{\tau}_{ij}$ a value for $\bar{\theta}$ is required to go from $\ln x_{ij}$ and $\ln N_{ij}$ in the data to model-implied values for $\theta \ln \tilde{\tau}_{ij}$.

Exports of a firm in the p^{th} percentile of the exporter size distribution are $\sigma F_{ij} \left(\varphi^p/\varphi^*_{ij}\right)^{\sigma-1}$, where φ^p is such that $\Pr\left[\varphi < \varphi^p | \varphi > \varphi^*_{ij}\right] = p$. Since productivity is distributed Pareto, the ratio φ^p/φ^*_{ij} and thus average exports per firm in each percentile should be the same for all ij pairs. This implies that the intensive margin elasticity calculated separately for each exporter size percentile is the same as the overall intensive margin elasticity.

Observation 4: $\text{IME}^{pct} = \text{IME}$, for all pct.

Data

We now use Observations 1 – 4 above to relate the simple Melitz-Pareto model to the data as described in Section 2.

Observation 1 indicates that if fixed trade costs vary by origin and destination but not across country pairs, i.e., $var(\widetilde{F}_{ij}) = 0$, then the IME should be equal to zero while the EME should be equal to one. This is captured in Figures 1 and 2 by the horizontal line for the model-implied intensive margin (panel a) and the line with unit slope for the model-implied extensive margin (panel b). These implications of the model stand in sharp contrast to what is seen in the data, both in Figures 1 and 2 and in Tables 2, 3, and 4, which reveal an IME of 0.4 or higher.

For the simple Melitz-Pareto model to be consistent with the data, we need to move away from $var(\widetilde{F}_{ij}) = 0$. As per Observation 2, however, the positive IME seen in the data implies a *negative* correlation between model-implied fixed and variable trade costs. Moreover, Observation 3 combined with the result in Table 6 of a negative distance elasticity of average exports per firm implies that model-implied fixed trade costs *fall* with distance.

We explore these results further by using equations (12) and (13) to compute model-implied fixed and variable trade costs.[15] The correlation between the resulting fixed and variable trade costs is -0.786 (with a standard error of 0.007). Figure 6 plots these trade costs against distance. The Figure shows that model-implied fixed trade costs are

[15]To compute model-implied variable trade costs as in equation (13), values for θ and $\bar{\theta}$ are required. We set $\theta = 5$ from Head (2014) and $\sigma = 5$ from Bas et al. (2015), which jointly imply $\bar{\theta} = 1.25$.

decreasing with distance, while model-implied variable trade costs are increasing with distance.[16] The distance elasticities corresponding to Figure 6 are reported in Table 7. For fixed trade costs this elasticity is -0.285 (as per Observation 3, this is equal to the distance elasticity of average exports reported in Table 6) while for variable trade costs the distance elasticity is 0.272, both statistically significant.

Finally, according to Observation 4, the simple Melitz-Pareto model implies that $\text{IME}^{pct} = \text{IME}$ for all pct. This theoretical prediction of a common elasticity across percentiles is captured by the horizontal line red in Figure 3. This is at odds with the data.

To conclude, the simple version of the Melitz-Pareto model with fixed trade costs varying only because of origin and destination fixed effects is clearly at odds with the data. One can of course allow a richer pattern of variation in fixed trade costs across country pairs to make the model perfectly consistent with the data, but then the positive IME has further puzzling implications for fixed trade costs, which should fall with distance and be very negatively correlated with variable trade costs. To the best of our knowledge, there are no models that would microfound such a strong and negative correlation between the two types of trade costs and a negative fixed trade costs elasticity with respect to distance.[17] The data is also at odds with the implication from the Melitz-Pareto model of a constant IME across exporter size percentiles.

3.2. Multi-Product Extension of Melitz-Pareto

In this section we explore whether the puzzling implications for trade costs arising from the Melitz-Pareto model can be avoided by extending the model to multi-product firms. The idea would be that average exports per firm may fall along with total exports (thereby creating a positive IME) as firms facing higher product-level fixed trade costs export fewer products (even though they export more per product). Roughly speaking, allowing for multi-product firms implies that part of the extensive margin in the basic Melitz-

[16]Variable trade costs must increase with distance so that total exports fall with distance, as implied by the results in 6.

[17]Allowing for tariffs in addition to iceberg trade costs would naturally lead to a *positive* correlation between model-implied variable and fixed trade costs. This is because a tariff affects trade flows both by increasing the price of the affected good, as with iceberg trade costs, and by decreasing the net profits conditional on the quantity sold, as with fixed trade costs. See the online appendix of Costinot and Rodríguez-Clare (2014), Felbermayr et al. (2015), and Caliendo et al. (2015).

Pareto model now operates inside the firm and appears as an intensive margin. We will see, however, that under the Pareto assumption the effect of higher product-level fixed trade costs on the number of products exported per firm is exactly offset by higher average exports per product, so that Observation 1 in the basic model will remain valid in this extension.

Theory

We consider an extension of the Melitz-Pareto model due to Bernard, Redding and Schott (2011). Each firm can produce a differentiated variety of each of a continuum of products in the interval [0,1] with productivity $\varphi\lambda$, where φ is common across products and λ is product-specific. The firm component φ is drawn from a Pareto distribution $G^f(\varphi)$ with shape parameter θ^f, while the firm-product component λ is drawn from a Pareto distribution $G^p(\lambda)$ with shape parameter θ^p. To have well-defined terms given a continuum of firms, we impose $\theta^f > \theta^p > \sigma - 1$. To sell any products in market j, firms from country i have to pay a fixed cost F_{ij}, and to sell each individual product requires an additional fixed cost of f_{ij}. Variable trade costs are still τ_{ij}.

The cutoff λ for a firm from country i with productivity φ that wants to export to market j, $\lambda^*_{ij}(\varphi)$, is given implicitly by

$$A_j \left(\frac{w_i \tau_{ij}}{\varphi \lambda^*_{ij}(\varphi)} \right)^{1-\sigma} = \sigma f_{ij}. \tag{14}$$

We can then write the profits in market j for a firm from country i with productivity φ as

$$\pi_{ij}(\varphi) \equiv \int_{\lambda^*_{ij}(\varphi)}^{\infty} \left[\left(\frac{\lambda}{\lambda^*_{ij}(\varphi)} \right)^{\sigma-1} - 1 \right] f_{ij} dG^p(\lambda). \tag{15}$$

The cutoff productivity for firms from i to sell in j is given implicitly by $\pi_{ij}(\varphi^*_{ij}) = F_{ij}$. As in the canonical model, the number of firms from country i that export to market j is $N_{ij} = \left[1 - G^f(\varphi^*_{ij})\right] N_i$, while the number of products sold by firms from i in j is $M_{ij} = N_i \int_{\varphi^*_{ij}}^{\infty} \left[1 - G^p\left(\lambda^*_{ij}(\varphi)\right)\right] dG^f(\varphi)$. Combining the previous expressions, using the fact that $G^p(\lambda)$ and $G^f(\varphi)$ are Pareto, writing $f_{ij} = f^o_i f^d_j \tilde{f}_{ij}$, $F_{ij} = F^o_i F^d_j \tilde{F}_{ij}$, and $\tau_{ij} = \tau^o_i \tau^d_j \tilde{\tau}_{ij}$, and defining variables appropriately we get

$$\ln X_{ij} = \mu_i^{X,o} + \mu_j^{X,d} - \theta^f \ln \widetilde{\tau}_{ij} - \left(\frac{\theta^f}{\sigma - 1} - \frac{\theta^f}{\theta^p} \right) \ln \widetilde{f}_{ij} - \left(\frac{\theta^f}{\theta^p} - 1 \right) \ln \widetilde{F}_{ij}, \tag{16}$$

$$\ln x_{ij}^p \equiv \ln X_{ij} - \ln M_{ij} = \mu_i^{x^p,o} + \mu_j^{x^p,d} + \ln \widetilde{f}_{ij}, \tag{17}$$

and

$$\ln x_{ij} \equiv \ln X_{ij} - \ln N_{ij} = \mu_i^{x^f,d} + \mu_j^{x^f,d} + \ln \widetilde{F}_{ij}. \tag{18}$$

It is easy to verify that if $f_{ij} = 0$ for all i, j then this model collapses to the canonical model with single-product firms.

Recalling our definition of the intensive margin elasticity at the firm and product level introduced in Section 2 and letting $\bar{\theta} \equiv \theta^f / (\sigma - 1)$ and $\chi \equiv \theta^f / \theta^p$, then from equations (16) to (18) we have

$$\text{IME} = -\frac{(\chi - 1)\, var\left(\ln \widetilde{F}_{ij}\right) + \left(\bar{\theta} - \chi\right) cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) + \theta cov(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij})}{var(\ln \widetilde{X}_{ij})} \tag{19}$$

and

$$\text{IME}^p = -\frac{\left(\bar{\theta} - \chi\right) var\left(\ln \widetilde{f}_{ij}\right) + (\chi - 1)\, cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) + \theta cov(\ln \widetilde{f}_{ij}, \ln \widetilde{\tau}_{ij})}{var(\ln \widetilde{X}_{ij})}. \tag{20}$$

Observation 1 in the single-product firm model remains valid in the multi-product firm model, while we now have an analogous observation for the product-level intensive margin elasticity:

Observation 5: If $var\left(\ln \widetilde{f}_{ij}\right) = 0$ then $\text{IME}^p = 0$.

The assumption $\theta^f > \theta^p > \sigma - 1$ implies that $\chi > 1$ and $\bar{\theta} > \chi > 1$ and in turn this leads to the following extensions of observation 2:

Observation 6: If $\text{IME} > 0$ then either $cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) < 0$ or $cov(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij}) < 0$ (or both).

Observation 7: If $\text{IME}^p > 0$ then either $cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) < 0$ or $cov(\ln \widetilde{f}_{ij}, \ln \widetilde{\tau}_{ij}) < 0$ (or both).

Observation 3 remains valid in the multi-product firm model, and we now also have an analogous observation for product-level fixed trade costs:

Observation 8: $\frac{cov(\ln \widetilde{x}^{p}_{ij}, \ln \widetilde{dist}_{ij}))}{var(\ln \widetilde{dist}_{ij})} = \frac{cov(\ln \widetilde{f}_{ij}, \ln \widetilde{dist}_{ij}))}{var(\ln \widetilde{dist}_{ij}))}$.

As in the single-product case, we can use the model to back out the implied trade costs. Equation (18) can be used to obtain a model-implied $\widetilde{F}_{ij}$ (which would be the same as the one derived in the single-product model) while Equation (17) can be used to obtain a model-implied $\widetilde{f}_{ij}$, and Equation (16) can then be used to obtain a model-implied $\widetilde{\tau}_{ij}$.

Data

Since Observations 1 and 3 remain valid when the basic model is extended to allow for multi-product firms, the conclusions regarding the necessity of having fixed trade costs decrease with distance remain valid. Turning to the implications for product-level fixed trade costs, the finding in Section 2 of a positive IME at the product level, IME$^p > 0$ in Table 5, combined with Observation 5 implies that, to be consistent with the data, the multi-product version of the Melitz-Pareto model presented above requires $var\left(\ln \widetilde{f}_{ij}\right) > 0$. However, observations 6 and 7 imply that the two types of fixed trade costs would need to be negatively correlated, or that the covariances between those fixed trade costs and variable trade costs would have to be negative. Moreover, Observation 8 combined with the results in Table 6 implies that model-implied product-level fixed trade costs decrease with distance with an elasticity of -0.060, as shown in the third column of Table 7 and illustrated in Figure 7. We conclude that the puzzling implications of the Melitz-Pareto model remain valid when the model extended to allow for multi-product firms.

3.3. Firm-Level Demand and Fixed-Cost Shocks

EKK extend the basic Melitz-Pareto model presented in Section 3.1 to allow for (log-normally distributed) firm-level destination-specific demand and fixed-cost shocks. Except for constants that capture the net effects of these shocks, our equations (7) and (8) remain valid in the EKK environment, and hence so do observations 1-3.[18]

It is important to note, however, that if productivity is distributed Pareto then the presence of log-normally distributed demand or fixed-cost shocks would imply that

[18] This can be confirmed by simple manipulation of equations (20) and (28) in EKK.

equations (7) and (8) no longer hold. The critical assumption in EKK that allows their model to be consistent with our equations (7) and (8) is that, loosely speaking, they consider the limit as the scale parameter of the Pareto distribution converges to zero.[19]

To formally establish this result, recall that to get equations (7) and (8) we assumed that $\varphi^*_{ij} > b_i$. If instead $\varphi^*_{ij} \leq b_i$ then $N_{ij} = N_i$ and $x_{ij} = \left(\frac{\theta}{\theta-(\sigma-1)}\right) A_j \left(\frac{w_i \tau_{ij}}{b_i}\right)^{1-\sigma}$. In the extreme, if $\varphi^*_{ij} \leq b_i$ holds for all i, j pairs, then we would have IME $= 1$ rather than IME $= 0$. Now think about the case with firm-specific demand and fixed-cost shocks. Specifically, assume that each firm is characterized by a productivity level φ as well as a demand shock α_j and a fixed cost shock f_j in each destination j, with φ drawn from a Pareto distribution (with scale parameter b_i and shape parameter θ) and α_j and f_j drawn iid from some distribution. Let $x_{ij}(\varphi, \alpha_j) \equiv A_j \alpha_j (\bar{\sigma} \frac{w_i \tau_{ij}}{\varphi})^{(1-\sigma)}$ and let $\varphi^*_{ij}(\alpha_j, f_j)$ be implicitly defined by $x_{ij}(\varphi^*_{ij}, \alpha_j) = \sigma f_j$. By the same argument we used in Section 3.1, if for all i, j and all possible (α_j, f_j) we have $\varphi^*_{ij}(\alpha_j, f_j) > b_i$, we can easily show that we still have IME $= 0$.[20] However, if α_j and f_j are lognormally distributed, then for $b_i > 0$ for all i there must be a positive mass of firms for which $\varphi^*_{ij}(\alpha_j, f_j) < b_i$, and for those firms there would be a positive intensive margin elasticity. EKK essentially avoid this by taking the limit with $b_i \to 0$ for all i.

In principle, one could use this result to argue that a Melitz model with Pareto distributed productivity but extended to allow for log-normally distributed demand and fixed-cost shocks could match the positive IME that we see in the data. However, such a model would not exhibit any of the convenient features of the canonical Melitz-Pareto model: the sales distributions is not distributed Pareto, the trade elasticity is not common and fixed, and the gains from trade are not given by the ACR formula. Given that, our approach in this paper is to move all the way to a model where productivity as well as destination-specific demand and fixed shocks are lognormally distributed. Such a model at least has the advantage that it is computationally tractable, and amenable to Maximum Likelihood Estimation, as we show in Section 4.

[19] More exactly, EKK specify a function for the measure of firms with productivity above some level, with that measure going to infinity as productivity goes to zero. This is equivalent to taking a limit with the (exogenous) measure of firms going to infinity and the scale parameter of the Pareto distribution going to zero. Although equations (7) and (8) do not hold anywhere in this sequence, they do hold in the limit.

[20] Consider the group of firms from country i that have some given draw $\{(\alpha_j, f_j), j = 1, ..., n\}$. The exact same argument used in Section 3.1 can be used to show that the sample of firms obtained by combining such firms across all origins i satisfies IME $= 0$. One can then simply integrate across all possible draws $\{(\alpha_j, f_j), j = 1, ..., n\}$ to show that IME $= 0$ for the whole set of firms.

3.4. Granularity

The previous sections have considered a model with a continuum of firms. With a discrete and finite number of firms it may be possible to generate a positive covariance between the intensive margin and total exports that could in principle explain our empirical findings. We explore this possibility in this section.

Theory

Eaton et al. (2012b) extend the Melitz-Pareto model above to allow for granularity. Equations (9) and (10) then become

$$\ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \widetilde{\tau}_{ij} - \bar{\theta} \ln \widetilde{F}_{ij} + \xi_{ij} \tag{21}$$

and

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \widetilde{F}_{ij} + \varepsilon_{ij}, \tag{22}$$

where ξ_{ij} and ε_{ij} are error terms arising from the fact that now the number of firms is discrete and random. Using the same definition for the intensive margin elasticity as in Section 3, the previous equations imply that

$$\text{IME} = \frac{-\left(\bar{\theta}-1\right) var(\ln \widetilde{F}_{ij}) - \theta cov\left(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij}\right) + var(\varepsilon_{ij}) + \text{COV}}{var\left(-\theta \ln \widetilde{\tau}_{ij} - \left(\bar{\theta}-1\right) \ln \widetilde{F}_{ij} + \varepsilon_{ij} + \xi_{ij}\right)}, \tag{23}$$

where $\text{COV} \equiv cov(\ln \widetilde{F}_{ij} + \varepsilon_{ij}, \xi_{ij}) + cov(\ln \widetilde{F}_{ij} + \ln \widetilde{\tau}_{ij}, \varepsilon_{ij})$. If $var(\varepsilon_{ij})$ is large relative to -COV, this could explain IME > 0 even with $cov\left(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij}\right) > 0$. Thus, in theory, granularity could explain the positive intensive margin elasticity that we find in the data without relying on implausible patterns for fixed trade costs.

To check whether granularity is a plausible explanation for the positive IME in the data we will conduct two tests. First, we will estimate the fixed trade cost elasticity with respect to distance taking into account granularity and the possible biases it may induce. Second, we will simulate firm-level exports under granularity and the assumption of fixed trade costs that vary by origin and destination only and estimate the implied IME. We describe each of these tests in turn.

Fixed Trade Costs and Distance with Granularity

In the Melitz-Pareto model with a continuum of firms, average exports per firm can be expressed as $x_{ij} = \kappa F_{ij}$, where $\kappa \equiv \frac{\sigma\bar{\theta}}{\theta-1}$. If we relax the continuum assumption to allow for granularity, then average exports per firm can be expressed as $x_{ij} = \kappa F_{ij} + \varepsilon_{ij}$, where ε_{ij} is an error term that arises from random realizations of productivity draws, and the first moment of which is independent of any variables that determine bilateral fixed trade costs. If we further assume that $F_{ij} = F_i^o F_j^d e^{\zeta \ln dist_{ij}} + v_{ij}/\kappa$, where v_{ij} satisfies $\mathbb{E}(v_{ij}|dist_{ij}) = 0$, we can then write

$$x_{ij} = \kappa F_i^o F_j^d e^{\zeta \ln dist_{ij}} + u_{ij}, \tag{24}$$

where $u_{ij} \equiv v_{ij} + \varepsilon_{ij}$ is an error term that captures both the deviation of F_{ij} from its mean as well as the granularity error term ε_{ij}. Since both $\mathbb{E}(v_{ij}|\ln dist_{ij})$ and $\mathbb{E}(\varepsilon_{ij}|\ln dist_{ij})$ are equal to zero, it follows that $\mathbb{E}(u_{ij}|dist_{ij}) = 0$. The challenge in estimating the fixed trade costs elasticity with respect to distance, ζ, from this equation is that we cannot simply take logs to obtain a log-linear equation to be estimated by OLS, because the error term that comes from granularity is not log-additive.

To take advantage of the time dimension of our data, we extend (24) to allow for an origin-time and destination-time specific components in the expression of fixed trade costs,

$$x_{ijt} = \kappa F_{it}^o F_{jt}^d e^{\zeta \ln dist_{ij}} + u_{ijt}, \tag{25}$$

where again $\mathbb{E}(u_{ijt}|dist_{ij}) = 0$. We estimate 25 using Poisson pseudo maximum likelihood method as in Silva and Tenreyro (2011) in the next subsection.

The IME under Granularity: Simulation

To assess how well granularity can explain a positive IME, we simulate exports of N_{ij} firms for each of the country pairs in the sample. We add demand shocks to allow for a less than perfect correlation between exports of different firms across different destinations. In the standard Melitz model with demand shocks, exports from i to j of a firm with productivity φ and destination-specific demand shock α_j can be calculated as

$$x_{ij}\left(\varphi,\alpha_j\right)=\sigma F_{ij}\left(\frac{\alpha_j\varphi}{\alpha_{ij}^{*}\varphi_{ij}^{*}}\right)^{\sigma-1}, \tag{26}$$

where $\alpha_{ij}^{*}\varphi_{ij}^{*}$ is a combination of productivity and demand shocks of the smallest exporter from i selling to j . To estimate the IME in simulations we perform the following steps:

1. Draw φ and α_j from some distribution. The number of draws is equal to N_{ij}, the number of exporters in the EDD dataset for each origin-destination pair in 2009. To be more precise, we draw the product $\alpha_j\varphi$ for each firm-destination pair assuming either that, as in the standard Melitz model, there are no demand shocks and hence the product $\alpha_j\varphi$ is perfectly correlated across destinations or that, at the other extreme, there is no correlation in the product $\alpha_j\varphi$ across destinations (pure demand shocks case). In both cases, we draw $\alpha_j\varphi$ from a Pareto distribution with a shape parameter to be specified below.

2. Assume that $var\left(\widetilde{F}_{ij}\right)=0$, so that $F_{ij}=F_i^oF_j^d$. This will allow us to study the IME generated by granularity by itself.

3. Use equation (26) to simulate the exports for each firm and to calculate average exports per firm (in total and in each percentile) for each origin-destination pair.

4. Run the IME regression 1 on the simulated export data, with $\ln x_{ij}$ being either the intensive margin for all firms exporting from i to j, or for each percentile in the size distribution of exporters from i to j .

Data

We now discuss the evidence obtained first for the fixed trade costs elasticity with respect to distance and second for the IME with simulated data.

We use equation 25 to estimate firm-level as well as product-level fixed trade cost elasticities with respect to distance (ζ). Table 8 shows that both of these elasticities are negative and statistically significant, so both firm-level and product-level model-implied fixed trade costs are decreasing with distance, although with a much smaller elasticity than when not accounting for granularity (compare results of Tables 7 and 8). Hence granularity does not help to eliminate one of the puzzles emerging from the comparison between the Melitz-Pareto model and the data.

Table 9 reports the estimated IME using simulated data for alternative values of $\bar{\theta}$ and for either zero or perfect correlation between the product of demand and productivity shocks across destinations. We consider 4 values of $\bar{\theta}$: our estimate $\bar{\theta} = 2.4$, the value that can be inferred from standard estimates of θ and σ in the literature (i.e., $\theta = 5$, the central estimate of the trade elasticity in Head and Mayer, 2014, and $\sigma = 5$ from Bas et al. (2015), so $\bar{\theta} = 1.25$), as well as $\bar{\theta} = 1.75$ from Eaton et al. (2011) (which they estimate using the procedure outlined in the Appendix) and $\bar{\theta} = 1$ (as in Zipf's Law).

Two broad patterns emerge from the table. First, the simulated IME decreases with $\bar{\theta}$. This is because the effect of granularity on the IME is stronger when there is more dispersion in productivity levels. Second, the simulated IME is highest when productivity is less correlated across destinations, again because this gives granularity more room to generate a covariance between average exports per firm and total exports.

For our estimate of $\bar{\theta}$ ($\bar{\theta} = 2.4$) and with no demand shocks (so there is perfect correlation in firm-level productivity across destinations), the simulated IME of 0.001 is quite low. The highest simulated IME occurs for the case in which $\bar{\theta} = 1$ and there is no correlation between the product of demand shocks and productivity across destinations. In this case the simulated IME is 0.33, not too far from our preferred estimate based on the data of 0.4. But we think of this as an extreme case because $\bar{\theta} = 1$ is far from the estimates that come out of trade data, and because of the implausible assumption that firm-level exports are completely uncorrelated across destinations.

To explore this further, we examine the implications for the IME across percentiles. We calculate average simulated exports per firm in each percentile and use those to estimate an IME per percentile. We plot the resulting 100 IME estimates in Figure 8 along with the corresponding IME estimates based on the actual data. The IME based on the actual data is increasing with a spike at the top percentile. Granularity and the Pareto distribution fail to reproduce this pattern in the simulated data, since the corresponding IME is much smaller than in the data for most percentiles. The IME in the simulated data is almost zero for small percentiles and is relatively high for a small number of top percentiles. We conclude that granularity does not offer a plausible explanation for the positive estimated IME in the data.

4. The Intensive Margin in the Melitz-Lognormal Model

In this section we depart from the assumption of a common Pareto distribution of firm-level productivity and instead assume a lognormal distribution.[21] In the theory section we start by showing how this can lead to a positive IME in a simple Melitz model, and then propose a maximum-likelihood estimation procedure for a richer Melitz model with heterogeneous fixed costs and demand shocks. The data section presents the results from the estimation and the implications for the IME as well as for the model-implied trade costs.

4.1. Theory

A simple Melitz model with a lognormal distribution

Consider a model exactly as that presented in Section 2, but with productivity distributed lognormal. We will show here the implication of this for the IME. The ratio of average to minimum exports for each country pair can be written as

$$x_{ij}/x_{ij}(\varphi_{ij}^*) = H(\varphi_{ij}^*), \tag{27}$$

where

$$H\left(\varphi_{ij}^*\right) \equiv \frac{1}{\left(\varphi_{ij}^*\right)^{\sigma-1}} \int_{\varphi_{ij}^*}^{\infty} \varphi^{\sigma-1} \frac{g(\varphi)}{1-G(\varphi_{ij}^*)} d\varphi.$$

If firm productivity is distributed Pareto with parameter $\theta > \sigma - 1$ (as in Section 2) then $H(\varphi) = \frac{\theta}{\theta-(\sigma-1)}$ (see equation 8), so that the average to minimum ratio does not depend on selection. This property only holds with a Pareto distribution of productivity. Assume instead that in each origin country i firm productivities are drawn from a lognormal

[21]One could consider combining a lognormal distribution with a Pareto distribution on the right tail, as in Nigai (2017). We have used Nigai's Matblab code on our data to estimate the point of truncation (percentile) where the lognormal ends and the Pareto begins. We find that for 75% of country pairs with more than a hundred exporters the point of truncation occurs after the 99th percentile, and for the median country pair the truncation point is at the 99.9%. In light of these results, in the rest of the paper we focus on the case in which productivity is described by a fully lognormal distribution.

distribution with location parameter $\mu_{\varphi,i}$ and scale parameter σ_φ. Letting $\Phi()$ be the CDF of the standard normal distribution, this implies that

$$G_i(\varphi) = \Phi\left(\frac{\ln\varphi - \mu_{\varphi,i}}{\sigma_\varphi}\right). \tag{28}$$

Letting $h(x) \equiv \Phi'(x)/\Phi(x)$ be the ratio of the PDF to the CDF of the standard normal, Bas et al. (2015) (henceforth BMT), show that

$$H\left(\varphi_{ij}^*\right) = \frac{h\left[-(\ln\varphi_{ij}^* - \mu_{\varphi,i})/\sigma_\varphi\right]}{h\left[-(\ln\varphi_{ij}^* - \mu_{\varphi,i})/\sigma_\varphi + \bar{\sigma}_\varphi\right]}, \tag{29}$$

where $\bar{\sigma}_\varphi \equiv (\sigma - 1)\,\sigma_\varphi$. Combined with $1 - G(\varphi_{ij}^*) = N_{ij}/N_i$, we have

$$\frac{x_{ij}}{x_{ij}(\varphi_{ij}^*)} = \Omega\left(\frac{N_{ij}}{N_i}\right) \equiv \frac{h\left(\Phi^{-1}\left(\frac{N_{ij}}{N_i}\right)\right)}{h\left(\Phi^{-1}\left(\frac{N_{ij}}{N_i}\right) + \bar{\sigma}_\varphi\right)}. \tag{30}$$

Thus, the average to minimum ratio of exports for country pair ij only depends on the share of total firms in i that export to j, with the relationship given by the function $\Omega()$.

As argued by BMT, $\Omega()$ is an increasing function. To understand the implication of this property, consider a decline in τ_{ij}, so that φ_{ij}^* decreases with no effect on minimum sales (which remain at σF_{ij}). The decline in τ_{ij} leads to an increase in exports of incumbent firms (which increases average exports per firm) and entry of low productivity firms (which decreases average exports per firm). Under Pareto these two effects exactly offset each other so there is no change in average exports per firm. If productivity is distributed lognormal the second effect does not fully offset the first, and hence average exports per firm increase with a decline in τ_{ij}. Since this also increases the number of firms that export (and hence total exports), this will naturally generate a positive IME.

Given values of $\bar{\sigma}_\varphi$ as well as N_i for every country, we can use our data on N_{ij} to compute $\Omega\left(N_{ij}/N_i\right)$ for all country pairs. Combined with $x_{ij}(\varphi_{ij}^*) = \sigma F_{ij}$ and imposing $F_{ij} = F_i^o F_j^d$, we can use equation (30) to get the model-implied average exports per firm (in logs),

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln\Omega\left(N_{ij}/N_i\right). \tag{31}$$

In contrast to Observation 1 for the Melitz-Pareto model, under under lognormality we will have a positive IME even with $var(\tilde{F}_{ij}) = 0$.

We can also compute model-implied fixed and variable trade costs similarly to what we did under the assumption of Pareto-distributed productivity. First, we obtain $\tilde{F}_{ij}$ from

$$\ln \widetilde{F}_{ij} = \delta_i^{F,o} + \delta_j^{F,d} + \ln x_{ij} - \ln \Omega \left(\frac{N_{ij}}{N_i} \right). \tag{32}$$

Second, to compute $\widetilde{\tau}_{ij}$, we combine equations (3), (4), (28) and (30) to get (with appropriately defined fixed effects)

$$(\sigma - 1) \ln \widetilde{\tau}_{ij} = \delta_i^{\tau,o} + \delta_j^{\tau,d} - \ln x_{ij} + \ln \Omega_i \left(\frac{N_{ij}}{M_i} \right) + \bar{\sigma}_\varphi \Phi^{-1} \left(1 - \frac{N_{ij}}{N_i} \right). \tag{33}$$

Armed with estimates of $\widetilde{F}_{ij}$ and $(\sigma - 1)\widetilde{\tau}_{ij}$, we can compute their correlation and check whether $\widetilde{F}_{ij}$ increases or decreases with distance (demeaned by origin and destination fixed effects).

These empirical exercises require estimates for $\bar{\sigma}_\varphi$ as well as N_i for every country. We use Bento and Restuccia (2015) (henceforth BR) data to estimate a value for N_i for all the countries in our sample.[22] We acknowledge slippage between theory and data in that we obviously do not have a measure of the entry level N_i, but (at best) only for the number of existing firms, which in theory would correspond to $(1 - G_i(\varphi_{ii}^*))\, N_i$ (our approach in the next subsection avoids this problem). We use the QQ-estimation proposed by Head et al. (2014) (henceforth HMT) to obtain estimates of σ_φ and $\mu_{\varphi,i}$ for every i (see the Appendix for a detailed description).

Full Melitz-lognormal model

The previous section has shown that a model with a lognormal distribution of firm productivity is capable of generating a positive intensive margin elasticity conditional on fixed costs. However, the model we considered had two very stark predictions. First, fixed trade costs that are common across firms lead to the prediction that sales of the

[22]Using census data as well as numerous surveys and registry data, BR compiled a dataset with the number of manufacturing firms for a set of countries. Unfortunately, the sample in BR has missing observations for a number of countries in the EDD. We impute missing values projecting the log number of firms on log population. There is a tight positive relationship between log number of firms in the BR dataset and log population with an elasticity of 0.945, as reported in Table 10 and in Figure 9.

least productive exporter from i to j are equal to σF_{ij}. In the data we observe many firms with very small export sales (sometimes as low as \$1) which implies unrealistic fixed trade costs. Second, as shown by Eaton et al. (2011), the model implies a perfect hierarchy of destination markets (i.e., destinations can be ranked according to profitability, with all firms that sell to a destination also selling to more profitable destinations) and perfect correlation of sales across firms that sell to multiple markets from one origin. None of these predictions holds in the data.

In this section we consider a richer model with firm-specific fixed trade costs and demand shocks that vary by destination. This is similar to the setup in Eaton et al. (2011). We assume that firm productivity, demand shocks (denoted by α_j) and fixed trade costs (denoted by f_j) are distributed jointly lognormal, i.e., for each origin i:

$$\begin{bmatrix} \ln\varphi \\ \ln\alpha_1 \\ \vdots \\ \ln\alpha_J \\ \ln f_1 \\ \vdots \\ \ln f_J \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{\varphi,i} \\ \mu_\alpha \\ \vdots \\ \mu_\alpha \\ \mu_{f,i1} \\ \vdots \\ \mu_{f,iJ} \end{bmatrix}, \begin{bmatrix} \sigma_\varphi^2 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_\alpha^2 & \dots & 0 & \sigma_{\alpha f} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_\alpha^2 & 0 & \dots & \sigma_{\alpha f} \\ 0 & \sigma_{\alpha f} & \dots & 0 & \sigma_f^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{\alpha f} & 0 & \dots & \sigma_f^2 \end{bmatrix} \right). \tag{34}$$

Note that we allow mean log productivity to be origin-specific while imposing that the mean of demand shocks be the same across origin-destination pairs (however, we cannot separately identify these parameters). Mean fixed costs are allowed to vary across origin-destination pairs and can be correlated with demand shocks within destinations. In our empirical estimation we will not be able to separately identify mean productivity from wages and variable trade costs – they will all be absorbed into an origin-destination fixed effect. Also, we restrict the dispersion of log productivity to be the same across all origins, and we restrict the dispersion of log demand shocks and log fixed trade costs to be the same across all origin-destination pairs.

Without risk of confusion, we change notation in this section and use $X_i \equiv (X_{i1}, ..., X_{iJ})$ to denote the random variable representing log sales of a firm from i in each of the J destinations, with $x_i \equiv (x_{i1}, ..., x_{iJ})$ being a realization of X_i, and $g_{X_i}(x_i)$ being the associated probability density function. According to the model, a firm does not export to destination j if it has a large fixed trade cost draw f_j relative to its productivity and its demand shock for that destination. Let $D_{ij} \equiv \ln\left[A_j (w_i \tau_{ij})^{1-\sigma}\right]$ and let $Z_{ij} \equiv D_{ij} + \ln\alpha_j + (\sigma - 1)\ln\varphi$ be sales in destination j by a firm from i with productivity φ and demand shock α_j. This is a latent variable that we observe only if a firm actually exports,

$$X_{ij} = \begin{cases} Z_{ij} & if \;\; \ln\sigma + \ln f_{ij} \leq Z_{ij} \\ \emptyset & otherwise \end{cases},$$

with $Z_i \equiv (Z_{i1}, ..., Z_{iJ})$ distributed according to

$$\begin{bmatrix} Z_{i1} \\ \vdots \\ Z_{iJ} \end{bmatrix} \sim N\left(\begin{bmatrix} d_{i1} \\ \\ d_{iJ} \end{bmatrix}, \begin{bmatrix} \bar{\sigma}_\varphi^2 + \sigma_\alpha^2 & \cdots & \bar{\sigma}_\varphi^2 \\ \vdots & \ddots & \vdots \\ \bar{\sigma}_\varphi^2 & \cdots & \bar{\sigma}_\varphi^2 + \sigma_\alpha^2 \end{bmatrix} \right), \tag{35}$$

where $d_{ij} \equiv D_{ij} + \mu_\alpha + (\sigma - 1)\mu_{\varphi,i}$ and $\bar{\sigma}_\varphi \equiv (\sigma - 1)\sigma_\varphi$.

Using firm-level data from the EDD across different origins and destinations, we can estimate the parameters in (35) as well as mean log fixed trade costs (up to a constant) and their dispersion using maximum likelihood methods. The Appendix shows how to derive the density function $g_{X_{i1},..,X_{iJ}}(x_{1i}, ..., x_{iJ})$ for the case when we observe sales to J destinations. For now, we simplify the analysis by considering only data for three destinations (USA, Germany and Japan), which we label $j = 1, 2, 3$ for the year 2007. We compute $g_{X_{i1},..,X_{i3}}(x_{1i}, ..., x_{i3})$ for each observation in our dataset (which is a realization of $\{X_{i1}, X_{i2}, X_{i3}\}$ that we observe). Since all random variables are independent across firms, we can compute the log-likelihood function as a sum of log-densities,

$$\ln L\left(\theta | \left\{x_{i1}\left(k_i\right), x_{i2}\left(k_i\right), x_{i3}\left(k_i\right)\right\}_{i,k_i}\right) = \sum_i \sum_{k_i=1}^{\widetilde{N}_i} \ln\left[g_{(X_{i1},X_{i2},X_{i3})}\left(x_{i1}\left(k_i\right), x_{i2}\left(k_i\right), x_{i3}\left(k_i\right)\right)\right], \tag{36}$$

where $\widetilde{N}_i$ is the number of firms from i that sell to either of the three destinations we consider, and where k_i is an index for a particular observation in our dataset (for origin i it takes values in $1, ..., N_i$) and θ is a vector of parameters that we want to estimate,

$$\theta = \left\{\left\{d_{ij}, \bar{\mu}_{f,ij}\right\}_{i,j}, \bar{\sigma}_\varphi, \sigma_\alpha, \sigma_f, \rho\right\}$$

where $\bar{\mu}_{f,ij} = \ln\sigma + \mu_{f,ij}$ and $\rho = \frac{\sigma_{\alpha f}}{\sigma_\alpha \sigma_f}$. As the likelihood is potentially not concave in θ and because there are 238 parameters to estimate, we rely on the estimation methodology proposed by Chernozhukov and Hong (2003). We use the Metropolis-Hastings MCMC algorithm to construct a chain of estimates $\theta^{(n)}$, dropping the first 30000k runs ("burn in" period) and then continuing until $n = 750,000$. Chernozhukov and Hong (2003) show that $\bar{\theta} \equiv \frac{1}{N}\sum_{n=1}^{N}\theta^{(n)}$ is a consistent estimator of θ, while the covariance matrix of $\bar{\theta}$ is given by the variance of $\theta^{(n)}$, so we use this to construct confidence intervals for $\bar{\theta}$.

Loosely speaking, identification works as follows. First, data on export flows and the number of exporters across country pairs helps in identifying d_{ij} and $\bar{\mu}_{f,ij}$. Second, the variance of firm sales within each ij pair helps in identifying the sum of the dispersion parameters for productivity and demand shocks, $\bar{\sigma}_\varphi + \sigma_\alpha$. Third, the extent of correlation of firm sales from a particular origin across different destinations helps in identifying σ_φ separately from σ_α: the more correlated firm sales are across destinations, the larger is σ_φ relative to σ_α. Fourth, the correlation between fixed costs and demand shocks can be inferred from the distribution of sales of small firms. Intuitively, if correlation is negative, then a firm with a bad demand shock would also likely draw a high fixed cost shock and thus will not export, hence, we would not see a lot of small firms in the data. Finally, to understand how σ_f is identified, imagine for simplicity that there is only one destination. We then have

$$g_{X_{i1}}\left(x_{i1}\right) = \frac{g_{Z_{i1}}\left(x_{i1}\right) \times \Pr\left\{\ln\sigma + \ln f_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\right\}}{C}$$

where $C \equiv \Pr\{\ln \sigma + \ln f_{i1} \leq Z_{i1}\}$ and where $g_{Z_{i1}}()$ is the probability density function of the latent sales Z_{i1}. This implies that we can get the density of X_{i1} by applying weights $\frac{\Pr\{\ln \sigma + \ln f_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\}}{C}$ to the density of Z_{i1}. The parameter σ_f regulates how these weights behave with x_{i1}. In the extreme case in which $\sigma_f = 0$ then the weights are 0 for $x_{i1} \leq \mu_{f_{i1}}$ and $1/C$ for $x_{i1} > \mu_{f_{i1}}$, while in the other extreme with $\sigma_f = \infty$ the weights are all equal to 1. For intermediate cases the density of X_{i1} will be somewhere in the middle, with the left tail becoming fatter and the right tail becoming thinner as σ_f increases. This suggests that we can identify σ_f from the shape of the density of sales.

We will use the results of the estimation to conduct similar exercises to those in the previous sections. First, we will compute the IME for all firms and for each percentile using the estimated model. Second, after removing origin and destination fixed effects, we will compute the correlation across the estimated values of d_{ij} and $\bar{\mu}_{f,ij}$, and between them and distance.

4.2. Data

Simple Melitz model with lognormal distribution

Table 11 reports the QQ-estimate of $\bar{\sigma}_\varphi$. We report three sets of estimates: for the full sample, the largest 50% of firms and the largest 25% of firms for each origin-destination pair in each year. These estimates are on the high side relative to the estimate obtained by HMT, so we will use the minimum among these estimates, $\bar{\sigma}_\varphi = 4.02$, which corresponds to the subsample with the largest 25% of firms.[23] First, a lognormal distribution allows the intensive margin elasticity to be positive even under the assumption of a continuum of firms. Second, for our estimate of the shape parameter, the implied IME is 0.28, which is close to that from the data.[24] Third, most of the action comes from the right tail of the exporter size distribution, as seen in Figure 10.

We use equations (32) and (33) to compute the model-implied fixed and variable trade costs. The correlations between those costs and distance are reported in Table 12 and plotted in Figure 11. In contrast to our results under Pareto, now under lognormal both the model-implied variable and fixed trade costs are increasing with distance.

[23] See the section in the Appendix titled QQ-Estimation of $\bar{\sigma}_\varphi$ for a discussion of these estimates and their relation to the estimate in HMT.

[24] Using Head et al. (2014) estimate of $\bar{\sigma}_\varphi = 2.4$ we get IME of around 0.12

Overall, the model does much better in fitting the data when we assume that firm productivity is distributed lognormal than when we assume that it is distributed Pareto. However, the IME for each percentile is not a perfect match to the data, and the there is still a negative correlation between the model implied variable and fixed trade costs, although it is much closer to zero than with Pareto (-0.3 rather than -0.8). In any case, this is just a ”proof of concept” that lognormally-distributed productivity can by itself improve the performance of the model relative to the data. In the next subsection we present the results obtained with the estimated full Melitz-lognormal model.

Full Melitz-lognormal model

To estimate the parameters of the full Melitz-lognormal model using firm-level data from the EDD for the year 2007 across 38 different origins and three destinations: Japan, the US and Germany,.

Before presenting the results of the estimation and discussing their implications for the IME, we show three figures revealing the fit of the estimated model with the data. Figure 12 shows a plot of the CDF for firm-level sales from one origin (name undisclosed for reasons of confidentiality) to the United States.[25] The estimated and empirical CDFs almost overlap. Other origin-destination pairs exhibit mostly similar fit for the CDF of firm sales, with a couple of exceptions associated with country pairs with low N_{ij}.

We next look at deviations from the strict hierarchy of firms sales across destinations (for each origin) in the data and in the estimated model. If there were no demand and fixed cost shocks across firms, then all firms from a given origin that export to less popular destinations would also export to the most popular destination. The share of firms that only sell in the less popular destinations is then a measure of the extent to which this strict hierarchy predicted by the simplest model is violated. According to Figure 13, the share predicted by the estimated model is quite close to the one in the data.

Finally, Figure 14 shows the correlation in sales across the U.S. and Germany for firms from a given origin that sell in both destinations. The estimated model implies that this correlation is narrowly clustered around 0.2 across our 38 origin countries, while in the data this correlation exhibits more dispersion.

The results of the estimation for the dispersion parameters $(\bar{\sigma}_\varphi, \sigma_\alpha, \sigma_f)$ are shown in

[25]With CDF $G_i(x)$, we have $G_i(x) = p$, hence the plot is of the function $\ln G_i^{-1}(p)$.

Table 13. The estimated values for $\bar{\sigma}_\varphi$ and σ_α are close to 3.2 and 2.9 respectively, while the estimate for σ_f is close to 2.6, all with very tight 95% confidence intervals. The estimated value for ρ is 0.32. The estimate of $\bar{\sigma}_\varphi$ is in between the estimated value of 2.4 in BMT and the value of 4.5 that we estimated in the simple lognormal model. To put these comparisons in context, note that in contrast to BMT and the simple Melitz model above, here we also have firm-specific demand shocks. Thus, for a particular origin-destination pair the standard deviation of latent sales is $(\bar{\sigma}_\varphi^2 + \sigma_\alpha^2)^{1/2} = 4.3$, although selection due to fixed trade costs brings the implied standard deviation for actual sales down to around 3, which is what we observe in the data.

Table (14) and Figure (15) show the implications of the estimated model for the IME. We compute the IME implied by the estimated model by drawing 1MM firms for each origin (this implies one million latent log sales and log fixed costs for each destination), computing average sales (taking into account selection), and then multiplying average sales by N_{ij} in the data to compute total exports. We pick one million because at this point we are not interested in granularity – this is just a numerical approximation to the case with a continuum of firms. The IME implied by the model is 0.55. This is actually higher than our preferred IME estimate of 0.4 in Section 2, but the gap comes in large part from the different sample of origin-destination pairs used here. Using the same sample of 38 origins and 3 destinations for the year 2007 we estimate IME of 0.55 (with a standard error of 0.038) that is statistically indistinguishable from the one implied by our estimated lognormal model.[26] We plot the associated IME for each percentile in Figure (15) – the pattern of the IME across percentiles is remarkably close to what we see in the data.

A natural question is how much of the high IME implied by the estimated model comes from the lognormality of productivity as opposed to the existence and shape of demand and fixed trade cost shocks. Table (14) reports the IME implied by the estimated model when we suppress those shocks by setting $\sigma_\alpha = \sigma_f = 0$. The IME drops only slightly in this case, suggesting that the key feature allowing the model to generate a high IME is indeed the shape of the productivity distribution.

Table 15 shows the elasticity of variable and fixed trade costs with respect to distance

[26]The confidence interval in Table (14) comes from the fact that we are running a regression to compute the IME, as in the data – it does not come from computing the IME for different values of the parameters along the Markov chain, although this is something we plan to do in the near future.

(controlling for origin and destination fixed effects). Now both types of trade costs are strongly increasing in distance. Surprisingly, however, we still get a negative correlation between fixed and variable trade costs.

Overall, our estimated full lognormal-Melitz model does a very good job in fitting the EDD data and in solving the puzzles associated with the Pareto model. The lognormal model generates an IME that is close to the one we see in the EDD and implies fixed trade costs that are positively correlated with variable trade costs and distance. The implied pattern for the IME across different percentile is also very similar to what we see in the data.

5. Conclusion

The canonical Melitz model of trade with Pareto-distributed firm productivities has a stark prediction: conditional on the level of the fixed costs of exporting, all variation in exports across partners should be due to the number of exporting firms (the extensive margin). There should be no variation in the intensive margin (exports per exporting firm), again conditional on fixed costs.

We use the World Bank's *Exporter Dynamics Database* plus China to test this prediction. Compared to existing studies, this data allows one to look for systematic variation in the intensive and extensive margins of trade, allowing for year, origin, and destination components of fixed trading costs.

We find that at least 40% of the variation in exports occurs along the intensive margin. That is, when exports from a given origin to a given destination are high, exports per exporting firm are responsible for at least 40% of the high exports. This finding is robust to looking at all destinations or only the largest destinations, including all firms or ignoring very small firms, including all country pairs or only ones for which more than 100 firms export, and disaggregating across industries. When we look at average exports by percentile of exporting firms (rather than the average), we find the intensive margin is more important the higher the percentile.

Although variation in fixed trade costs across country pairs can make the Melitz-Pareto model fit the intensive margin in the data, such fixed trade costs would need to be negatively correlated with variable trade costs and with distance. Moreover, variation in fixed trade costs does not reproduce the pattern of a steadily rising intensive margin by exporter percentiles. Allowing firms to export multiple products or taking into account granularity (a finite number of exporting firms) does not reverse these implications.

In contrast, moving away from a Pareto distribution and assuming that the productivity distribution is lognormal resolves the puzzles. Specifically, we estimate a Melitz model with lognormally distributed firm productivity and idiosyncratic firm-destination demand shifters and fixed costs. We estimate this model using maximum likelihood methods on the EDD firm-level data. The estimated Melitz-lognormal model is consistent with the positive intensive margin overall and with the intensive margin rising by exporting firm percentile. This specification also implies fixed trade costs that are positively correlated with variable trade costs and distance.

Whether the underlying distribution of firm productivity is Pareto or lognormal may matter for the gains from trade. First, Head et al. (2014) argue that the static gains from trade are typically larger in a calibrated Melitz-lognormal model than under Melitz-Pareto (e.g., the equivalent of 5% vs. 2% of GDP). Second, recent models of the dynamic gains from trade have emphasized how domestic firms can learn from firms selling or producing in the domestic market. In Alvarez et al. (2014), Buera and Oberfield (2015), and Perla et al. (2015), trade liberalization boosts the level or growth rate of technology through this channel. The size of this dynamic gain should depend on whether the distribution of firm productivity is Pareto vs. lognormal, as it interacts with how trade alters the distribution of producer and seller productivity. For example, trade liberalization induces more entry of marginal exporters under Pareto than under lognormal, as seen by no change in exports per exporter under Pareto (zero intensive margin elasticity, unit extensive margin elasticity) vs. a sizable intensive margin and weaker extensive margin under lognormal.

References

Alvarez, Fernando, Francisco J. Buera, and Robert E. Lucas Jr., "Idea Flows, Economic Growth, and Trade," 2014.

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

_ and Marc-Andreas Muendler, "Exporters and their products: a collection of empirical regularities," *CESifo Economic Studies*, 2013, *59* (2), 223–248.

_ , Arnaud Costinot, and Andrés 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," *Unpublished, MIT*, 2015.

_ , Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and Production in the Global Economy," 2014.

_ , Svetlana Demidova, Peter J Klenow, and Andres Rodriguez-Clare, "Endogenous Variety and the Gains from Trade," *American Economic Review*, 2008, *98* (2), 444–50.

Axtell, Robert L, "Zipf distribution of US firm sizes," *Science*, 2001, *293* (5536), 1818–1820.

Bas, Maria, Thierry Mayer, and Mathias Thoenig, "From Micro to Macro: Demand, Supply, and Heterogeneity in the Trade Elasticity," 2015.

Bento, Pedro and Diego Restuccia, "Misallocation, Establishment Size, and Productivity," Technical Report 2015.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "Firms in international trade," *The Journal of Economic Perspectives*, 2007, *21* (3), 105–130.

_ , _ , _ , and _ , "The margins of US trade," *The American Economic Review*, 2009, *99* (2), 487–493.

_ , Jonathan Eaton, J Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American economic review*, 2003, *93* (4), 1268–1290.

Buera, Francisco J. and Ezra Oberfield, "The Global Diffusion of Ideas," 2015.

Caliendo, Lorenzo, Robert C Feenstra, John Romalis, and Alan M Taylor, "Tariff Reductions, Entry, and Welfare: Theory and Evidence for the Last Two Decades," Technical Report, National Bureau of Economic Research 2015.

Chaney, Thomas, "Distorted gravity: the intensive and extensive margins of international trade," *The American Economic Review*, 2008, *98* (4), 1707–1721.

Chernozhukov, Victor and Han Hong, "An MCMC approach to classical estimation," *Journal of Econometrics*, 2003, *115* (2), 293–346.

Costinot, Arnaud and Andrés Rodríguez-Clare, "Trade Theory with Numbers: Quantifying the Consequences of Globalization," *Handbook of International Economics*, 2014, *4*, 197.

Eaton, Jonathan, Marcela Eslava, Maurice Kugler, and James Tybout, *The margins of entry into export markets: evidence from Colombia*, The Organization of Firms in a Global Economy, Cambridge, MA: Harvard University Press, 2008.

_ , Samuel Kortum, and Francis Kramarz, "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 2011, *79* (5), 1453–1498.

_ , Samuel S Kortum, and Sebastian Sotelo, "International trade: Linking micro and macro," 2012.

_ , _ , and _ , "International trade: Linking micro and macro," Technical Report, National bureau of economic research 2012.

Feenstra, Robert C, "Restoring the product variety and pro-competitive gains from trade with heterogeneous firms and bounded productivity," 2014.

Felbermayr, Gabriel, Benjamin Jung, and Mario Larch, "The welfare consequences of import tariffs: A quantitative perspective," *Journal of International Economics*, 2015, *97* (2), 295–309.

Fernandes, Ana M., Caroline Freund, and Martha Denisse Pierola, "Exporter behavior, country size and stage of development: Evidence from the exporter dynamics database," *Journal of Development Economics*, 2016, *119*, 121 – 137.

Freund, Caroline and Martha Denisse Pierola, "Export superstars," *Review of Economics and Statistics*, 2015, *97* (5), 1023–1032.

Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 2009, *1* (1), 255–294.

Head, Keith, "Gravity Equations: Workhorse, Toolkit, and Cookbook," *Handbook of International Economics*, 2014, *4*, 131.

_ , Thierry Mayer, and Mathias Thoenig, "Welfare and Trade without Pareto," *American Economic Review*, 2014, *104* (5), 310–16.

Luttmer, Erzo GJ, "On the mechanics of firm growth," *The Review of Economic Studies*, 2011, *78* (3), 1042–1068.

Manova, Kalina and Zhiwei Zhang, "Export Prices Across Firms and Destinations," *The Quarterly Journal of Economics*, 2012, *127* (1), 379–436.

Mayer, Thierry and Soledad Zignago, Working Papers 2011-25, CEPII 2011.

Melitz, Marc J, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 2003, *71* (6), 1695–1725.

Nigai, Sergey, "A tale of two tails: Productivity distribution and the gains from trade," *Journal of International Economics*, 2017, *104* (C), 44–62.

Perla, Jesse, Christopher Tonetti, and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," 2015.

Rossi-Hansberg, Esteban and Mark LJ Wright, "Establishment size dynamics in the aggregate economy," *The American Economic Review*, 2007, *97* (5), 1639–1666.

Silva, JMC Santos and Silvana Tenreyro, "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator," *Economics Letters*, 2011, *112* (2), 220–222.

Appendix

Estimation of $\bar{\theta}$

An estimate of $\bar{\theta}$ is required to compute model-implied $\ln \widetilde{F}_{ij}$ and $\ln \widetilde{\tau}_{ij}$ as functions of $\ln x_{ij}$, $\ln N_{ij}$, and estimated fixed effects. We follow Eaton et al. (2011) and derive the following expression

$$\frac{x_{il|j}}{x_{il|l}} = \left(\frac{N_{ij}}{N_{il}}\right)^{-1/\bar{\theta}} \tag{37}$$

where $x_{il|j}$ are average exports per firm for firms from i that sell in market l but restricted to those firms that sell in markets l and j. EKK have information on domestic sales for each firm, so they use $l = i$. We do not have such information, so we use $l^*(i) \equiv \arg\max_k N_{ik}$, that is, the largest destination market for each origin country i (e.g., the United States for Mexico). Letting

$$z_{ij} \equiv \frac{x_{il^*(i)|j}}{x_{il^*(i)|l^*(i)}} \tag{38}$$

and

$$m_{ij} \equiv \frac{N_{ij}}{N_{il^*(i)}} \tag{39}$$

then we have

$$\ln z_{ij} = -\frac{1}{\bar{\theta}} \ln N_{ij}. \tag{40}$$

This suggests an OLS regression to recover an estimate for $\bar{\theta}$.

Eaton et al. (2011) estimate this regression for French firm-level data (including information on sales in France) and obtain a coefficient of -0.57, which implies $\bar{\theta} = 1.75$. In their case, they keep in their estimating sample only firms with positive sales in France, so the variables $x_{FF|j}$ and N_{Fj} are calculated based on the same set of firms. To implement an approach comparable to theirs, we drop all firms from country i that do not sell to $l^*(i)$, so the sample includes only $N_{il^*(i)}$ firms for country i. This implies that all firms that make up N_{ij} are also selling to $l^*(i)$. Figure 5 reproduces Figure 3 from Eaton et al. (2011) by plotting the variables in equation (40). The slope in the graph is equal to $1/\bar{\theta}$, and the corresponding estimated values are reported in Table A1. Based on all observations in the core sample of countries and using no weighting, the estimated

$\bar{\theta}$ is over 19. But in Figure 5 for small values of m_{ij}, which correspond to small values of N_{ij}, there is a lot of dispersion in z_{ij}. To minimize the effect of that noise we weight observations by $\sqrt{N}_{ij}$ and this lowers the estimate of $\bar{\theta}$ to 4.8. Finally, when we drop all observations with $N_{ij} < 100$ (remember that here N_{ij} measures the number of firms from country i that sell to country j *and* also to $l^*(i)$) we obtain $\bar{\theta} = 2.3$, which is still higher than in Eaton et al. (2011). We will use this estimate in our simulations of the intensive margin elasticity.

QQ-Estimation of σ_φ

Exports from country i to country j of a firm with productivity φ in the model with CES preferences and monopolistic competition is given by $x_{ij}(\varphi) = \sigma F_{ij}\left(\varphi/\varphi^*_{ij}\right)^{\sigma-1}$. Since $\ln\varphi \sim N(\mu_{\varphi,i}, \sigma_\varphi)$ then $\ln x_{ij}(\varphi) \sim N_{trunc}(\bar{\mu}_{\varphi,ij}, \bar{\sigma}_\varphi; \ln(\sigma F_{ij}))$, where $\bar{\sigma}_\varphi = \sigma_\varphi(\sigma-1)$, $\bar{\mu}_{\varphi,ij} = \mu_{\varphi,i}(\sigma-1) + \ln(\sigma F_{ij}) + (1-\sigma)\ln\left(\varphi^*_{ij}\right)$, and the truncation point is $\ln(\sigma F_{ij})$.

As in HMT, we estimate $\bar{\sigma}_\varphi$ using a quantile-quantile regression, which minimizes the distance between the theoretical and empirical quantiles of log exports. Empirical quantiles are given by:

$$Q^E_{ij,n} = \ln x_{ij,n} \tag{41}$$

where n is the rank of the firm among exporters from i to j. We calculate theoretical quantiles of exports from i to j as

$$Q^T_{ij,n} = \bar{\mu}_{\varphi,ij} + \bar{\sigma}_\varphi \Phi^{-1}\left(\hat{\Phi}_{ij,n}\right), \tag{42}$$

where $\hat{\Phi}_{ij,n} = \frac{N_i-(n-1)}{N_i}$ is the empirical CDF and N_i is the imputed number of firms from the BR data. Following HMT we adjust the empirical CDF so that $\hat{\Phi}_{ij,n} = \frac{N_i-(n-1)-0.3}{N_i+0.4}$ since otherwise we would get $\Phi^{-1}\left(\hat{\Phi}_{ij,1}\right) = \infty$ when $n = 1$. The QQ-estimator of $\bar{\sigma}_\varphi$ is the coefficient β obtained from the regression

$$\ln x_{ij,n} = \alpha_{ij} + \beta\Phi^{-1}\left(\hat{\Phi}_{ij,n}\right) + \varepsilon_{ij,n}. \tag{43}$$

Table 11 reports the QQ-estimate of $\bar{\sigma}_\varphi$. We report three sets of estimates: for the full sample, the largest 50% of firms and the largest 25% of firms for each origin-destination pair in each year. According to the model, the estimates of the slope should not change when we consider different sub-samples, but this is not the case in Table 11. This comes from a not very surprising empirical failure of the simple Melitz-lognormal model outlined in the first part of the previous section: whereas this model implies that the sales distribution for any country pair should be distributed as a truncated lognormal (with the truncation at sales of σF_{ij}), no such truncation exists in the data (i.e., we observe exporters with very small sales).

A related issue is that our estimates for either of the sub-samples are significantly larger than the HMT estimate of 2.4. The difference comes from the fact that HMT assume that the sales distribution for any ij pair is lognormal, whereas we stick close to the simple model and assume that it is a truncated lognormal, and then use data for N_{ij} and our estimated values N_i to derive implicit truncation points. These truncation points tend to be on the right tail of the distribution, since N_{ij}/N_i tends to be quite low, hence the small $\bar{\sigma}_\varphi$ estimated by HMT would not be able to match the observed dispersion in the sales of exporters. In general, the higher the N_i one takes as an input in the QQ regression, the higher the estimate of the shape parameter one obtains.

In private correspondence, the authors of HMT pointed out that their approach would be consistent with the Melitz-lognormal model if one allows for heterogeneous fixed costs and lets the variance of these costs go to infinity, whereas our approach would be right if the variance goes to zero. This is part of our motivation in allowing for heterogeneous fixed costs and then in using MLE to estimate the full Melitz-lognormal model.

Quasi-Bayesian Estimation for the full Melitz-lognormal model

The likelihood function is a product of density functions of individual firms that sell or do not sell to multiple destinations. In this section we will use the notation from Section 4 of the paper. Let $\bar{\varphi}_i \equiv (\sigma - 1)[\ln \varphi - \mu_{\varphi,i}]$ be a random variable that denotes deviations from mean productivity for country i. Individual firm density of export sales $(x_{i1}, ..., x_{iJ})$ can be written as:

$$f_{X_{i1},..,X_{iJ}}(x_{i1},...,x_{iJ}) = \int_{\omega} f_{X_{i1},...,X_{iJ}|\bar{\varphi}_i}(x_{i1},...,x_{iJ}|\omega) f_{\bar{\varphi}_i}(\omega) d\omega \tag{44}$$

$$= \int_{\omega} \prod_j f_{X_{ij}|\bar{\varphi}_i}(x_{ij}|\omega) f_{\bar{\varphi}_i}(\omega) d\omega \tag{45}$$

where the second equality comes from the fact that conditional on productivity, sales are independent across markets (as well as the probability of selling to those markets). We now need to characterize $g_{X_{ij}|\bar{\varphi}_i}(g_{ij}|\omega)$ to calculate the likelihood function. In general we have:

$$f_{X_{ij}|\bar{\varphi}_i}(x_{ij}|\omega) = \left[f_{Z_{ij}|\bar{\varphi}_i}(x_{ij}|\omega)\Pr\{Z_{ij} \geq \ln\sigma + \ln f_{ij}|\bar{\varphi}_i = \omega, Z_{ij} = x_{ij}\}\right]^{\mathbb{I}(x_{ij}\neq\emptyset)} \times$$
$$\times \left[\Pr\{\ln\sigma + \ln f_{ij} \geq Z_{ij}|\bar{\varphi}_i = \omega\}\right]^{\mathbb{I}(x_{ij}=\emptyset)} \tag{46}$$

The term on the first line of 46 corresponds to the density function for the cases when we observe exports, while the second line corresponds to the mass at the point $x_{ij} = \emptyset$.

For the case when sales are not zero $X_{ij} = Z_{ij}$ and

$$Z_{ij}|\left[\bar{\varphi}_i = \omega\right] = \omega + d_{ij} + \ln\alpha - \mu_\alpha \tag{47}$$

$$Z_{ij}|\left[\bar{\varphi}_i = \omega\right] \sim N(d_{ij} + \omega, \sigma_\alpha^2) \tag{48}$$

In addition

$$\Pr[Z_{ij} \geq \ln\sigma + \ln f_{ij}|\bar{\varphi}_i = \omega, Z_{ij} = x_{ij}] = \Pr[\ln\sigma + \ln f_{ij} \leq x_{ij}|\ln\alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega] \tag{49}$$

$$\ln\sigma + \ln f_{ij}|\left[\ln\alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega\right] \sim N\left(\mu_1, \sigma_1^2\right) \tag{50}$$

$$\mu_1 = \bar{\mu}_{f,ij} + \frac{\sigma_{\alpha f}}{\sigma_\alpha^2}(x_{ij} - d_{ij} - \omega)$$

$$\sigma_1^2 = \sigma_f^2(1 - \rho^2)$$

Finally :

$$\Pr[Z_{ij} \leq \ln\sigma + \ln f_{ij} | \bar{\varphi}_i = \omega] = \Pr[-\ln\sigma - \ln f_{ij} + (\ln\alpha - \mu_\alpha) + d_{ij} \leq -\omega] \tag{51}$$

$$-\ln\sigma - \ln f_{ij} + (\ln\alpha - \mu_\alpha) + d_{ij} \sim N(-\bar{\mu}_{f,ij} + d_{ij}, \sigma_2^2) \tag{52}$$

$$\sigma_2^2 = \sigma_f^2 + \sigma_\alpha^2 - 2\sigma_{\alpha f}$$

Let ϕ and Φ denote PDF and CDF of standard normal. Plugging functional forms into 46 we can get the object of interest:

$$f_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ}) = \int_\omega \prod_j \left\{ \left[\frac{1}{\sigma_\alpha} \phi\left(\frac{x_{ij} - d_{ij} - \omega}{\sigma_\alpha} \right) \Phi\left(\frac{x_{ij} - \left[\bar{\mu}_{f,ij} + \frac{\sigma_{\alpha f}}{\sigma_\alpha^2}(x_{ij} - d_{ij} - \omega) \right]}{\sqrt{\sigma_f^2(1-\rho^2)}} \right) \right]^{\mathbb{I}(x_{ij} \neq \emptyset)} \times \right.$$
$$\left. \times \left[\Phi\left(\frac{-\omega + \bar{\mu}_{f,ij} - d_{ij}}{\sqrt{\sigma_f^2 + \sigma_\alpha^2 - 2\sigma_{\alpha f}}} \right) \right]^{\mathbb{I}(x_{ij} = \emptyset)} \right\} \frac{1}{\sigma_{\bar{\varphi}}} \phi\left(\frac{\omega}{\sigma_{\bar{\varphi}}} \right) d\omega \tag{53}$$

However, since we only have a truncated sample of $X'_{ij}s$ (as we don't observe sales of firms that do not export), we need to normalize the density by the inverse of probability that a firm is selling to at least one destination, and so we are interested in the object:

$$g_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ}) = f_{X_{i1},..,X_{iJ} \cap \text{Is an exporter}}(x_{1i},...,x_{iJ} \cap \text{Is an exporter}) \tag{54}$$

$$\begin{aligned} g_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ}) &= \\ \frac{f_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ})}{\Pr_i[\text{observe sales to at least 1 destination}]} &= \\ \frac{f_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ})}{1 - \Pr_i[\text{observe sales to no destinations}]} &= \\ \frac{f_{X_{i1},..,X_{iJ}}(x_{1i},...,x_{iJ})}{1 - \int_\omega \prod_j \left[\Phi\left(\frac{-\omega + \bar{\mu}_{f,ij} - d_{ij}}{\sqrt{\sigma_f^2 + \sigma_\alpha^2 - 2\sigma_{\alpha f}}} \right) \right] \frac{1}{\sigma_{\bar{\varphi}}} \phi\left(\frac{\omega}{\sigma_{\bar{\varphi}}} \right) d\omega} & \end{aligned} \tag{55}$$

The likelihood function is a product of density functions 55. Parameters to estimate are

$$\theta = \left\{ \{d_{ij}, \bar{\mu}_{f,ij}\}_{i,j}, \bar{\sigma}_{\varphi}, \sigma_{\alpha}, \sigma_f, \rho \right\}$$

To compute density 53, which is in the numerator of 55. We can think of 53 in the following general form:

$$\begin{aligned} f_{X_{i1},..,X_{iJ}}(x_{1i}, ..., x_{iJ}) &= \int_{\omega} G(\omega)\phi\left(\frac{\omega}{\sigma_{\bar{\varphi}}}\right) d\omega \\ &= \int_{\omega} G(\omega)\frac{1}{\sqrt{2\pi}} exp\left(-\left[\frac{\omega}{\sqrt{2}\sigma_{\bar{\varphi}}}\right]^2\right) d\omega \end{aligned} \tag{56}$$

where $G(\omega)$ is a known function of ω. Using change of variables $\widetilde{\omega} = \frac{\omega}{\sqrt{2}\sigma_{\bar{\varphi}}}$ and $d\omega = \sqrt{2}\sigma_{\bar{\varphi}} d\widetilde{\omega}$ we can write:

$$f_{X_{i1},..,X_{iJ}}(x_{1i}, ..., x_{iJ}) = \int_{\widetilde{\omega}} G(\sqrt{2}\sigma_{\bar{\varphi}}\widetilde{\omega})\frac{\sigma_{\bar{\varphi}}}{\sqrt{\pi}} exp\left(-\widetilde{\omega}^2\right) d\widetilde{\omega} \tag{57}$$

We can speed up the process to calculate object in 57 by applying a Gauss-Hermite quadrature. In general:

$$\int_x g(x)exp(-x^2)dx \approx \sum_i (g(x_i)w_i) \tag{58}$$

we calculate 33 values of the variable x_i as well as weights w_i using the Gauss-Hermite method.

Tables and Figures

Table 1: Core Sample of EDD countries+China, years firm-level data is available

ISO3	Country name	1st year	Last year	ISO3	Country name	1st year	Last year
ALB	Albania	2004	2012	KHM	Cambodia	2003	2009
BFA	Burkina Faso	2005	2012	LAO	Laos	2006	2010
BGD	Bangladesh	2005	2013	LBN	Lebanon	2008	2012
BGR	Bulgaria	2003	2006	MAR	Morocco	2003	2013
BOL	Bolivia	2006	2012	MDG	Madagascar	2007	2012
BWA	Botswana	2003	2013	MEX	Mexico	2003	2012
CHL	Chile	2003	2012	MKD	Macedonia	2003	2010
CHN	China	2003	2008	MMR	Myanmar	2011	2013
CIV	Cote d'Ivoire	2009	2012	MUS	Mauritius	2003	2012
CMR	Cameroon	2003	2013	MWI	Malawi	2009	2012
COL	Colombia	2007	2013	NIC	Nicaragua	2003	2013
CRI	Costa Rica	2003	2012	NPL	Nepal	2011	2013
DOM	Dominican Republic	2003	2013	PAK	Pakistan	2003	2010
ECU	Ecuador	2003	2013	PRY	Paraguay	2007	2012
EGY	Egypt	2006	2012	PER	Peru	2003	2013
ETH	Ethiopia	2008	2012	QOS	Kosovo	2011	2013
GAB	Gabon	2003	2008	ROU	Romania	2005	2011
GEO	Georgia	2003	2012	RWA	Rwanda	2003	2012
GIN	Guinea	2009	2012	THA	Thailand	2012	2013
GTM	Guatemala	2005	2013	TZA	Tanzania	2003	2012
HRV	Croatia	2007	2012	UGA	Uganda*	2003	2010
IRN	Iran	2006	2010	URY	Uruguay	2003	2012
JOR	Jordan	2003	2012	YEM	Yemen	2008	2012
KEN	Kenya	2006	2013	ZAF	South Africa	2003	2012
KGZ	Krygyztan	2006	2012	ZMB	Zambia	2003	2011

* indicates that Uganda does not have data for 2006

Table 2: IME regressions, core sample

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
Panel a: country pairs with $N_{ij} \geq 100$			
IM elasticity	0.438***	0.459***	0.400***
Standard error	[0.0058]	[0.0041]	[0.0049]
R^2	0.55	0.74	0.85
Variation in $\ln X_{ij}$ explained by FE,%	0.01	0.20	0.59
Observations	7,781	7,768	7,324
Panel b: all country pairs			
IM elasticity	0.503***	0.530***	0.579***
Standard error	[0.0018]	[0.0017]	[0.0022]
R^2	0.77	0.81	0.85
Variation in $\ln X_{ij}$ explained by FE, %	0.00	0.20	0.50
Observations	47,129	47,129	47,037
Year FE	Yes		
Origin $\times$ year FE		Yes	Yes
Destination $\times$ year FE			Yes

Note: robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: IME regression, small firms excluded, core sample

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
Panel a: country pairs with $N_{ij} \geq 100$			
IM elasticity	0.437***	0.459***	0.398***
Standard error	[0.0058]	[0.0042]	[0.0050]
R^2	0.54	0.74	0.85
Variation in $\ln X_{ij}$ explained by FE,%	0.01	0.19	0.59
Observations	7,698	7,684	7,234
Panel b: all country pairs			
IM elasticity	0.497***	0.525***	0.573***
Standard error	[0.0013]	[0.0013]	[0.0015]
R^2	0.77	0.81	0.84
Variation in $\ln X_{ij}$ explained by FE, %	0.00	0.19	0.50
Observations	46,925	46,925	46,832
Year FE	Yes		
Origin $\times$ year FE		Yes	Yes
Destination $\times$ year FE			Yes

Note: firms with annual exports lower than $1000 excluded

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: IME regression, disaggregated within manufacturing, core sample

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
Panel a: HS 2-digit			
IM elasticity	0.569***	0.510***	0.467***
Standard error	[0.0022]	[0.0017]	[0.0049]
Observations	37,321	35,621	10,732
Panel b: HS 4-digit			
IM elasticity	0.651***	0.569***	0.515***
Standard error	[0.0019]	[0.0013]	[0.0069]
Observations	62,776	58,516	4,640
Panel c: HS 6-digit			
IM elasticity	0.664***	0.593***	0.508***
Standard error	[0.0020]	[0.0014]	[0.0094]
Observations	67,967	61,501	2,972
Year $\times$ HS FE	Yes		
Origin $\times$ Year $\times$ HS FE		Yes	Yes
Destination $\times$ Year $\times$ HS FE			Yes

Note: $N_{ij} > 100$

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Product-level IME regression, core sample

	Coefficient from $\ln x_{ij}^{p}$ on $\ln X_{ij}$		
IM elasticity	0.380***	0.397***	0.288***
Standard error	[0.0070]	[0.0054]	[0.0073]
R^2	0.35	0.62	0.78
Variation in $\ln X_{ij}$ explained by FE,%	0.01	0.20	0.59
Observations	7781	7,768	7,324
Year FE	Yes		
Origin $\times$ year FE		Yes	Yes
Destination $\times$ year FE			Yes

Note: $N_{ij} > 100$

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Margins of trade and distance

Elasticity with respect to distance		
x_{ij}	0.123***	-0.280***
Standard error	[0.0150]	[0.0130]
N_{ij}	-0.416***	-1.010***
Standard error	[0.0134]	[0.0128]
x^p_{ij}	0.288***	-0.071***
Standard error	[0.0158]	[0.0146]
m_{ij}	-0.165***	-0.209***
Standard error	[0.0059]	[0.0051]
Observations	7,725	7,320
Origin $\times$ year FE	Yes	Yes
Destination $\times$ year FE		Yes

Note: $N_{ij} > 100$

Table 7: Trade costs and distance

	$\ln \widetilde{F}_{ij}$	$\ln \widetilde{\tau}_{ij}$	$\ln \widetilde{f}_{ij}$
$\ln dist_{ij}$	-0.280***	0.272***	-0.071***
Standard error	[0.0140]	[0.0046]	[0.0146]
Observations	7,320	7,320	7,320

Note: $N_{ij} > 100$

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Fixed trade costs distance elasticity and granularity

	Fixed trade costs elasticity	
	Firm level	Product level
ζ	-0.022***	-0.007***
Standard error	[0.0029]	[0.0026]
Observations	7,320	7,320

Note: $N_{ij} > 100$

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Product-level IME is calculated for the core sample

Table 9: IME under granularity

	$corr(\alpha_j\varphi, \alpha_k\varphi)$	
	0	1
$\widetilde{\theta} = 2.3$	0.005	0.001
$\widetilde{\theta} = 1.75$	0.020	0.005
$\widetilde{\theta} = 1.25$	0.133	0.036
$\widetilde{\theta} = 1$	0.333	0.103

Note: $N_{ij} > 100$

N_{ij} data as of 2007, core sample, 867 obs.

Table 10: Number of firms and population

	log number of firms	
log population	0.945***	0.944***
Standard error	[0.0136]	[0.0139]
Observations	468	468
Year FE		Yes

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: QQ estimates of $\bar{\sigma}_{\varphi}$

	All firms	Top 50%	Top 25%
$\bar{\sigma}_{\varphi}$	6.829***	4.676***	4.020***
	[0.0010]	[0.0006]	[0.0008]
Observations	11,902,823	5,917,685	2,949,514
R^2	0.81	0.93	0.94
Bilateral FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Robust standard errors in brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Trade costs and distance, lognormal

	log fixed costs	log variable costs
$\ln dist$	0.156***	0.299***
Standard error	[0.0155]	[0.0051]
Observations	7738	7738

$N_{ij} > 100$

Robust standard errors in brackets

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Estimates of dispersion, full lognormal model

	Estimate	95% CI
$\bar{\sigma}_{\varphi}$	3.23	[3.17, 3.30]
σ_{α}	2.93	[2.90, 2.96]
σ_{f}	2.63	[2.60, 2.66]
ρ	0.32	[0.28, 0.36]

Table 14: Implied IME in full lognormal model

	IME	95% CI
Unrestricted model	0.546	[0.545, 0.548]
Setting $\sigma_f = \sigma_{\alpha} = 0$	0.533	[0.527, 0.538]

Table 15: Implied trade costs in full lognormal model

	Estimate	95% CI
$corr\left(\widetilde{F}_{ij}, \widetilde{\tau}_{ij}\right)$	-0.22	[-0.04, -0.39]
	Distance elasticity	
Fixed costs	1.28	[0.98, 1.57]
Variable costs	0.26	[0.04, 0.49]

Table A1: Estimates of $\bar{\theta}$

	$\bar{\theta}$	s. e.	Observations
All observations, no weights	18.61***	[0.787]	39,712
Weights $\sqrt{N_{ij}}$	4.481***	[0.0360]	39,712
Dropping $N_{ij} < 100$	2.657***	[0.0175]	7,781
Dropping $M_{ij} < 100$	2.360***	[0.0147]	5,267

Robust standard errors in brackets

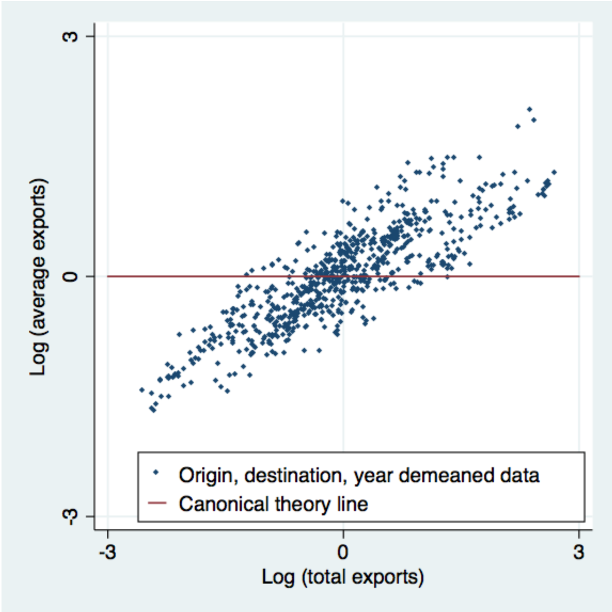
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

N_{ij} denotes the number of exporters from i to j

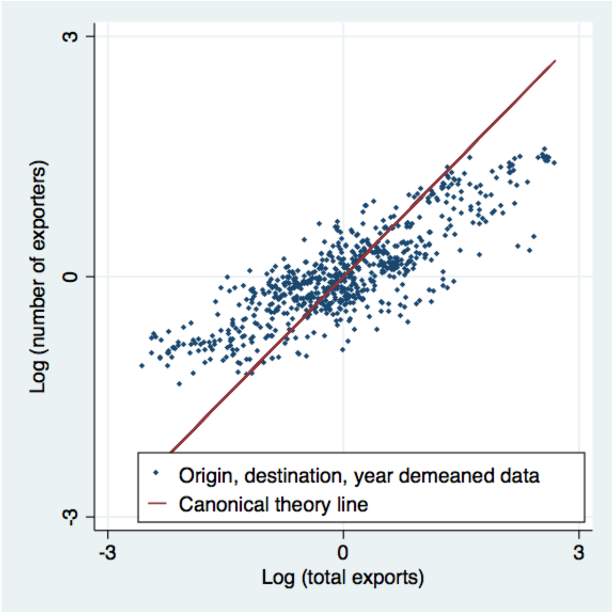
M_{ij} denotes the number of exporters from i to j that also export to i's largest destination

Figure 1: Intensive and Extensive margins of exporting

Panel A: Average size of exporters (intensive margin) and total exports



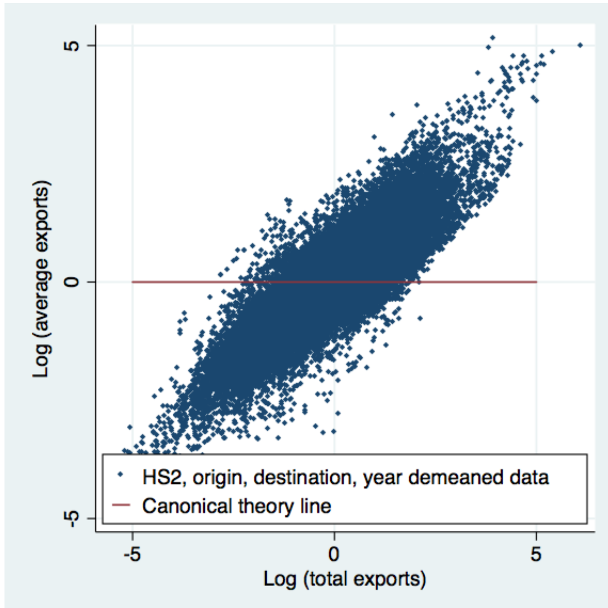
Panel B: Number of exporters (extensive margin) and total exports



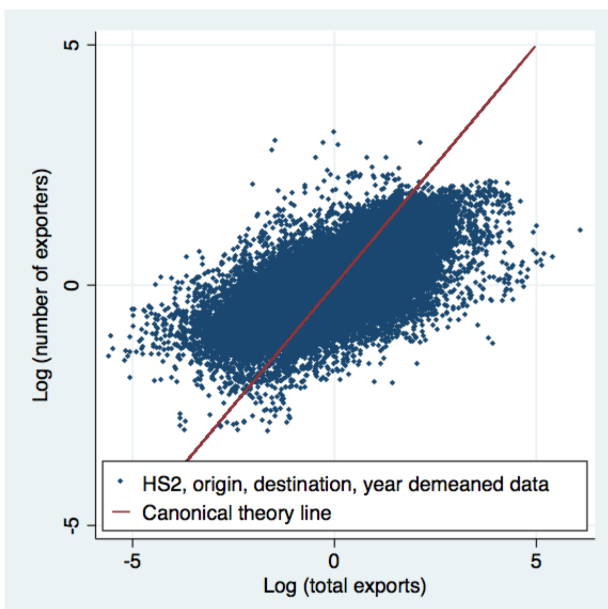
Note: the source are the statistics in the Exporter Dynamics Database for the extended sample. The x-axis represents log total exports at the origin country-destination country-year level demeaned by origin, destination, and year fixed effects. Only origin-destination pairs with more than 100 exporting firms considered. The dots represent the raw measures. The line is the slope predicted by the Melitz-Pareto model.

Figure 2: Intensive and Extensive margins of exporting, by industry

Panel a: Average size of exporters (intensive margin) and total exports

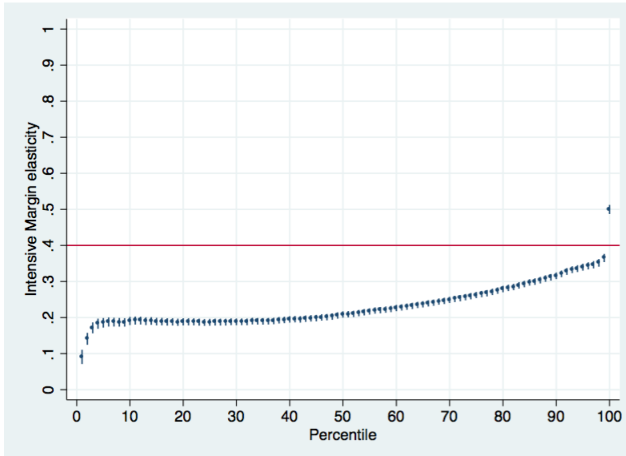


Panel b: Number of exporters (extensive margin) and total exports



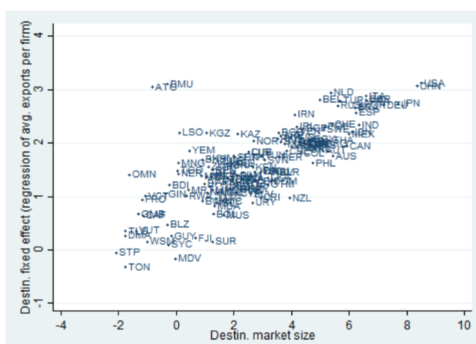
Note: the source are the statistics in the Exporter Dynamics Database for the extended sample. The x-axis represents log total exports at the origin country-HS 2-digit product-destination country-year level demeaned by origin, destination, HS 2-digit, and year fixed effects. Only origin-HS 2-digit-destination triplets with more than 100 exporting firms are considered. The line is the slope predicted by the Melitz-Pareto model.

Figure 3: IME for each percentile, data



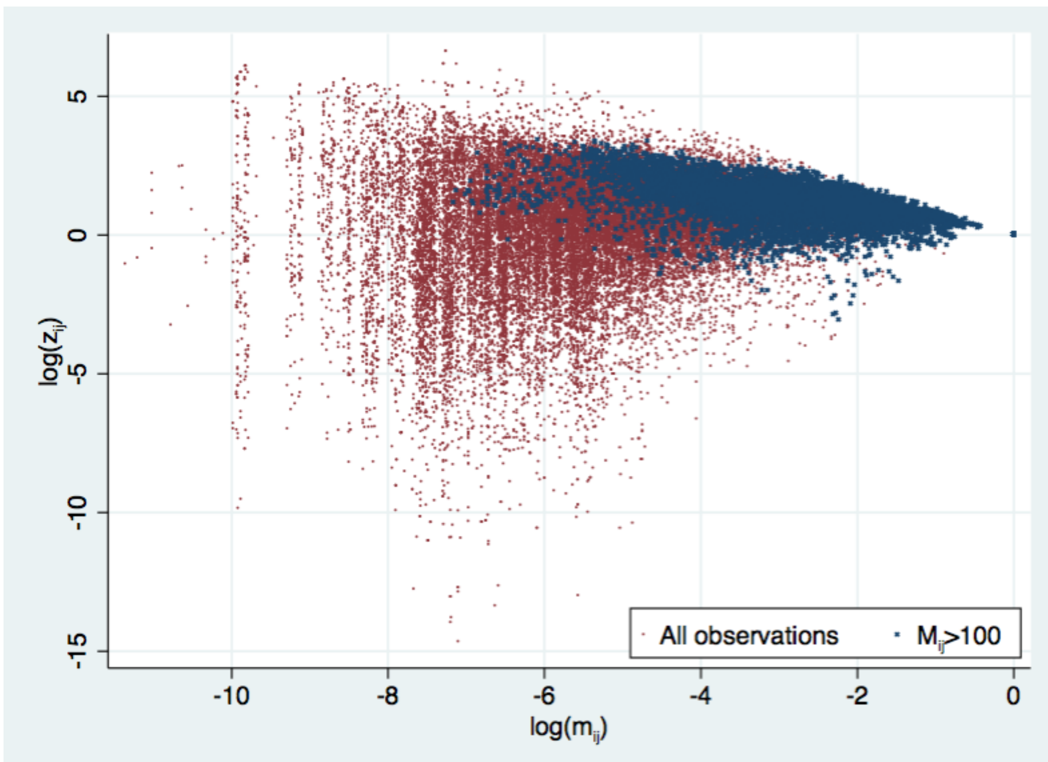
Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents percentiles of the average exporter size distribution. Each dot represents the coefficient from the regression of log average exports per firm in an exporter size percentile on log total exports. The data is demeaned by origin, destination, and year fixed effects.

Figure 4: Manufacturing absorption and averaged exports per firm (destination fixed effects)



Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents the log of manufacturing absorption in each destination country measured by manufacturing gross production plus manufacturing imports minus manufacturing exports (measured in billions of USD). The y-axis represents the estimated destination fixed effects obtained from a regression of log average exports per firm on origin, destination, and year fixed effects based on the core sample considering origin-destination pairs with more than 100 exporting firms. Manufacturing gross production is calculated as manufacturing value-added from the World Development Indicators divided by 0.418 (the factor used by EKK). Manufacturing imports and exports are obtained from COMTRADE/WITS.

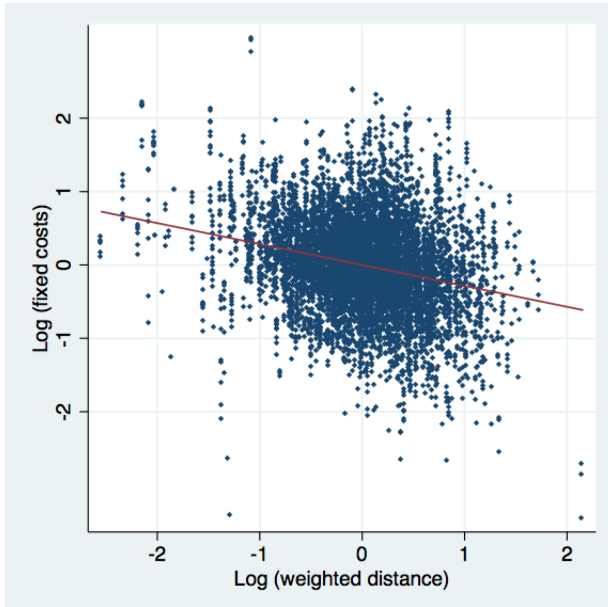
Figure 5: Exports to largest destination and market entry



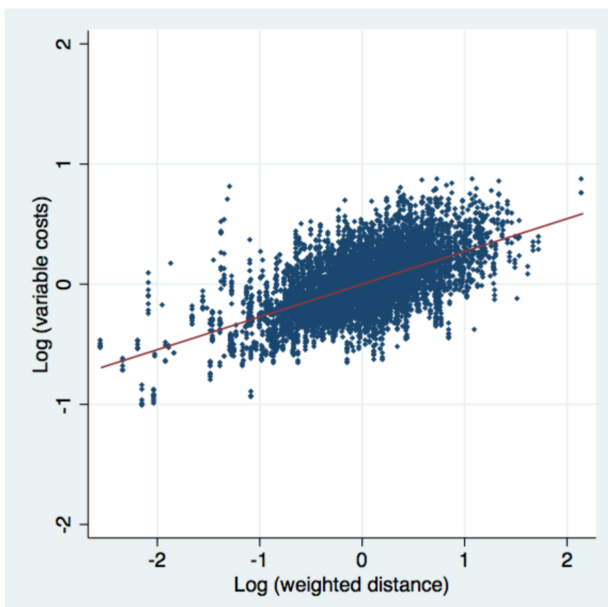
Note: the source is the exporter-level data used for the Exporter Dynamics Database for the core sample. The x-axis represents for each country i the log of the ratio of average exports per exporter to destination j to average exports per exporter to i s most popular destination market. The y-axis represents for each country i the log of the ratio of the number of exporters to destination j to the number of exporters to i s most popular destination market. For the calculation of both average exports per exporter and number of exporters we focus only on firms from i that sell both in j and in the most popular destination.

Figure 6: Model-implied fixed and variable trade costs and distance

Panel a: fixed trade costs and distance



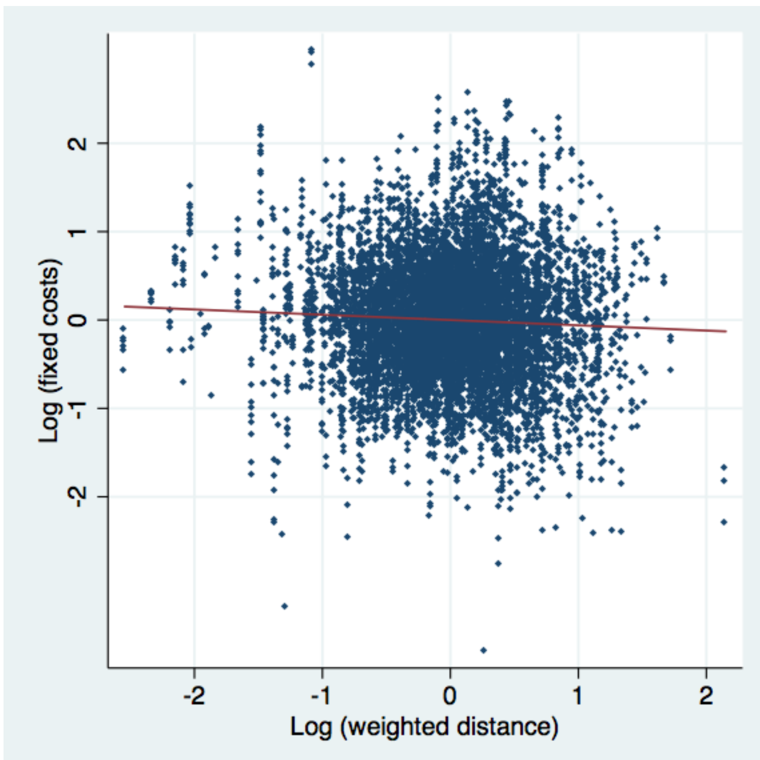
Panel b: variable trade costs and distance



Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance demeaned by origin and destination fixed effects taken from Mayer and Zignago (2011). The y-axis represents model-implied fixed or variable trade costs demeaned by origin and destination fixed effects. To calculate the model-implied fixed and variable trade costs we use $\theta = 5$ from Head (2014) and $\sigma = 5$ from Bas et al. (2015)

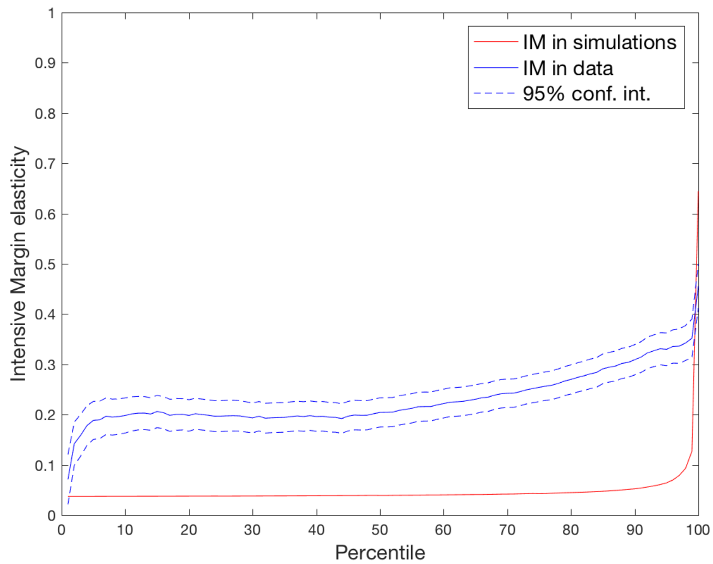
.

Figure 7: Fixed product-level trade costs and distance



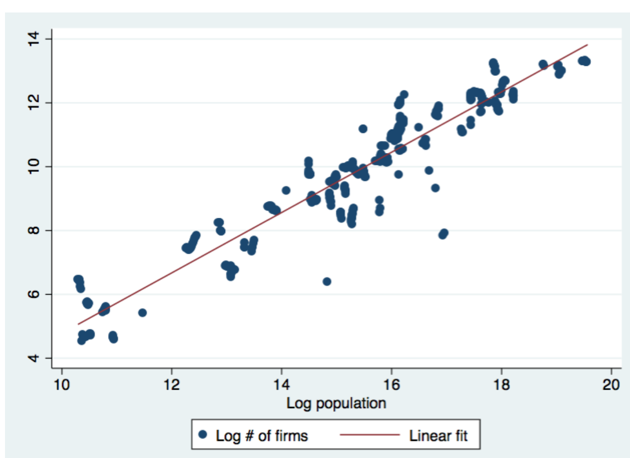
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011).

Figure 8: IME for each percentile, Pareto and granularity



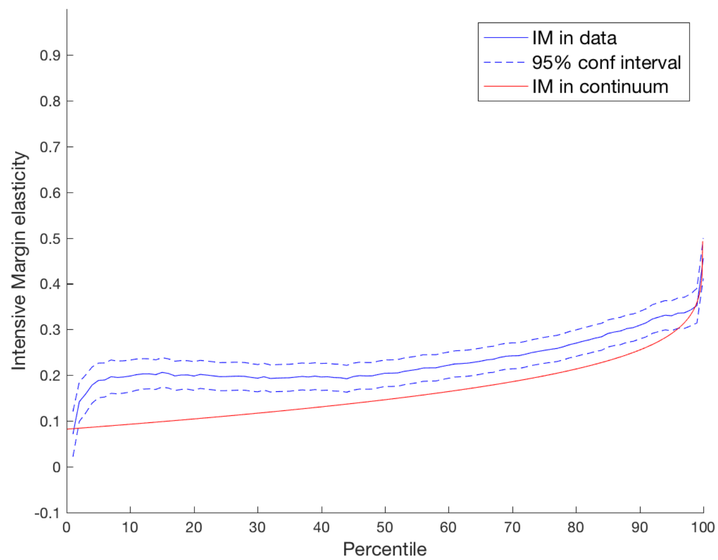
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and USA. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME for each percentile implied by the model with Pareto distribution of productivity and granularity, $\widetilde{\theta} = 1$. The level of bilateral fixed trade costs was chosen to match overall IME in the data. The number of draws for each origin-destination pair is equal to the number of exporters from origin to destination in EDD as of 2009.

Figure 9: Number of firms and population



Note: the x-axis represents log of population taken from the World Development Indicators. The y-axis represents the number of firms as computed by Bento and Restuccia (2015).

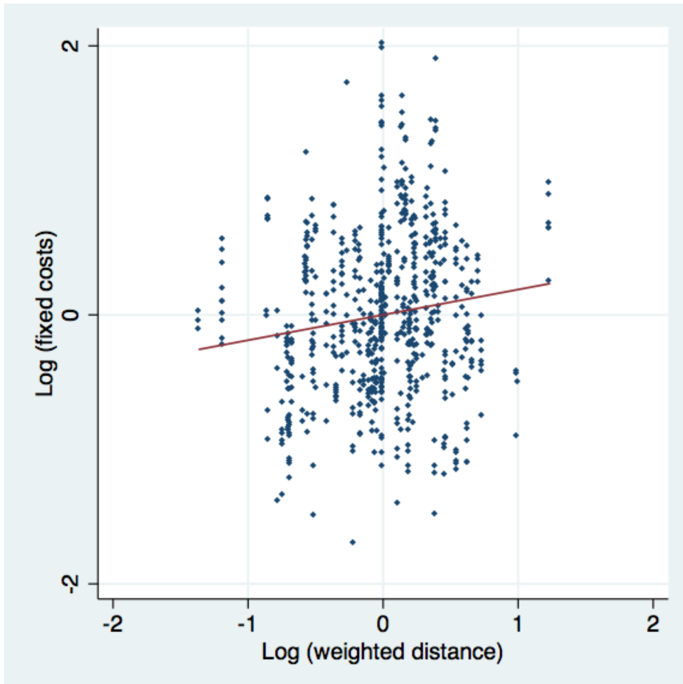
Figure 10: IME for each percentile, lognormal



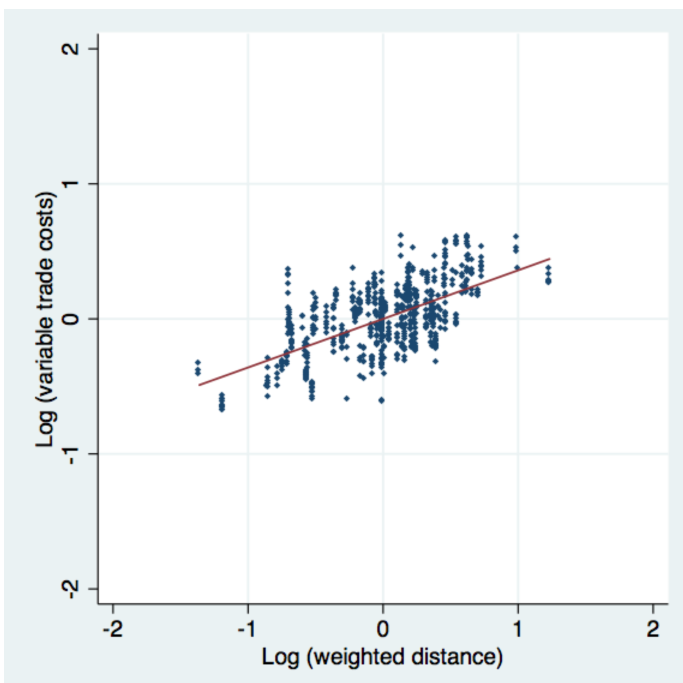
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and USA. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME for each percentile implied by the model with lognormal distribution of productivity, $\bar{\sigma}_{\varphi} = 4.02$ (our estimate) and $\sigma = 5$ from Bas et al. (2015). The level of bilateral fixed trade costs was chosen to match overall IME in the data. The total number of firms was imputed from Bento and Restuccia (2015).

Figure 11: Fixed and variable trade costs and distance, lognormal

Panel a: fixed trade costs and distance

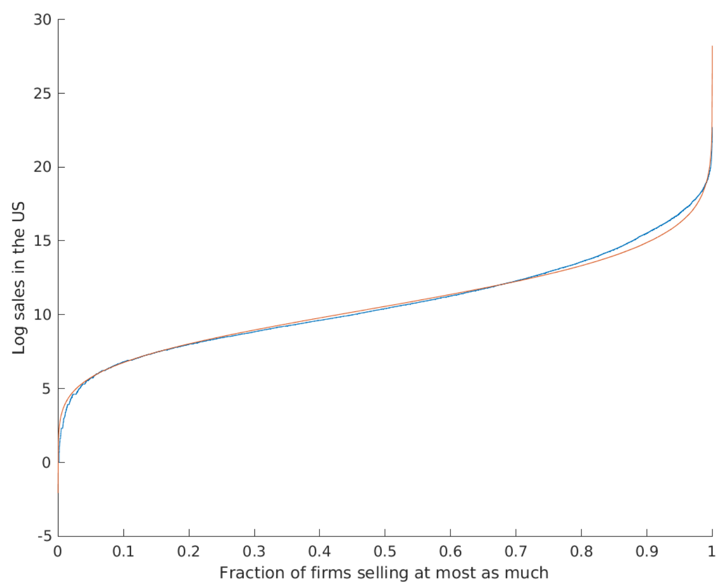


Panel b: variable trade costs and distance



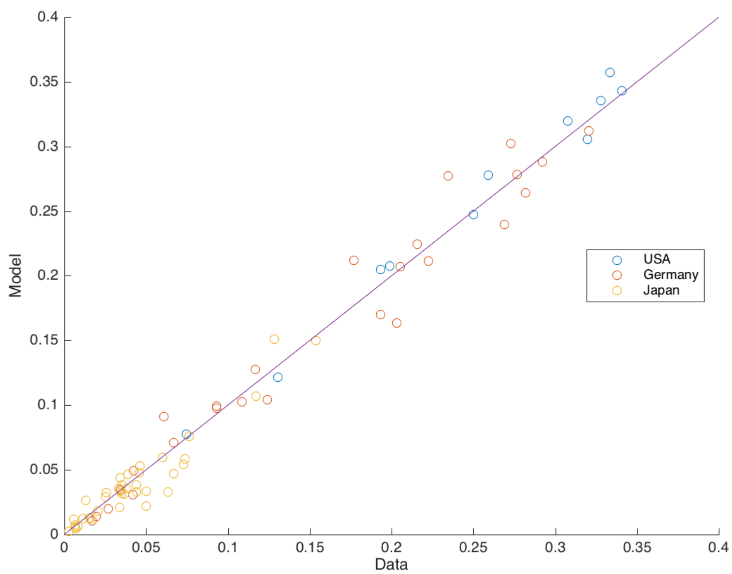
Note: source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011). Only four destination countries are considered: France, Germany, Japan, and the US. To calculate the model-implied fixed and variable trade costs we use our estimate of $\sigma_{\varphi} = 4.02$ and $\sigma = 5$ from Bas et al. (2015), and implied number of firm from Bento and Restuccia (2015).

Figure 12: Full lognormal model, inverse CDF of log sales



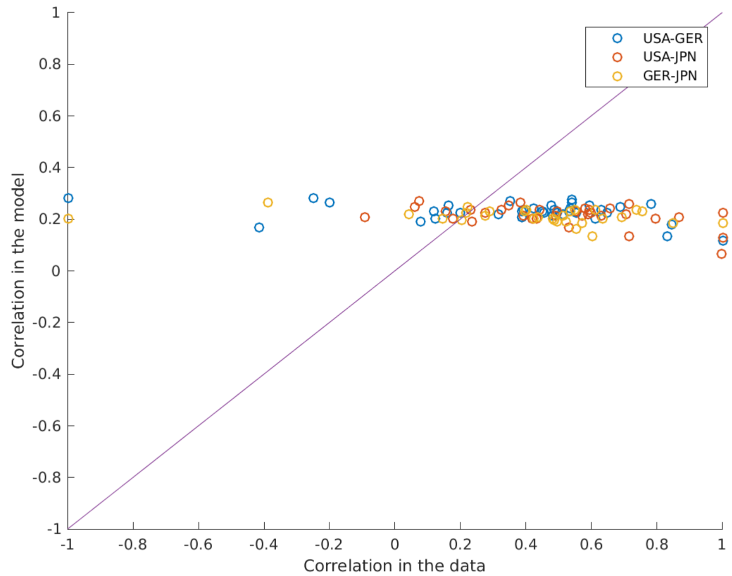
Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. The darker solid line corresponds to empirical CDF of log sales from some origin to the US. The lighter solid line corresponds to CDF implied by estimated full lognormal model.

Figure 13: Share of firms selling only to the less popular market



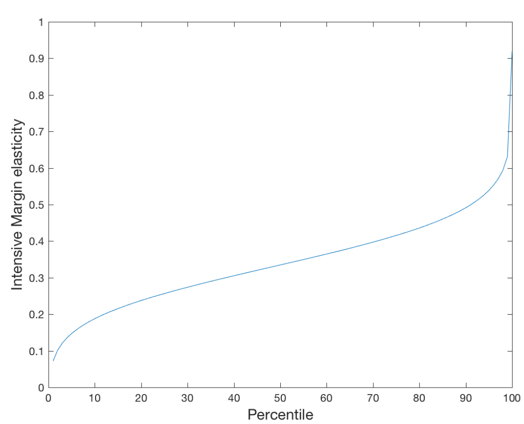
Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. Each point corresponds to the share of firms exporting only to less popular markets in the data (horizontal axis) and according to the estimated model (vertical axis). The figure also includes a 45^o line.

Figure 14: Correlation between log exports to the US and Germany



Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. Each point corresponds to the correlation between log exports to the US and Germany (for those firms who sell in both markets) in the data (horizontal axis) and according to the estimated model (vertical axis). The size of the circles is proportional to the number of exporters.

Figure 15: IME for each percentile, data and full lognormal model



Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. The x-axis represent percentiles. The dark solid line represents coefficient from the regression of log average exports in each percentile on log total exports in the data.