The Intensive Margin in Trade: Moving Beyond Pareto

Ana M. Fernandes
World Bank

Peter J. Klenow
Stanford and NBER

Sergii Meleshchuk
UC Berkeley

Martha Denisse Pierola
World Bank

Andrés Rodríguez-Clare *
UC Berkeley and NBER

December 4, 2015
Preliminary

Abstract

The Melitz model with Pareto-distributed firm productivity has become a tractable benchmark in international trade. It predicts that, conditional on the fixed costs of exporting, all variation in exports across trading partners will occur on the extensive margin (the number of firms exporting). In the World Bank's *Exporter Dynamics Database* on firm-level exports from 50 countries, however, we find that half of the variation in exports occurs on the intensive margin (exports per exporting firm). We explore several ways to explain this discrepancy. Most importantly, firm productivity does not follow a Pareto distribution.

*We are grateful to Caroline Freund, Keith Head, Thierry Mayer for useful discussions and to seminar participants at Berkeley, the Paris School of Economics, Stanford, and the Minneapolis Fed 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. The findings expressed 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

The Melitz (2003) model with Pareto-distributed firm productivity has become a useful and tractable benchmark in international trade. It is consistent with many firm-level facts – see, for example, Eaton et al. (2011). It generates a gravity equation (Chaney, 2008). And it yields a simple summary statistic for the welfare gains from trade (Arkolakis et al., 2012).

The Melitz-Pareto model makes 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 exporter, but draw in marginal exporters to the point that average exports per exporter – the intensive margin – is unchanged. This exact offset is a special property of the Pareto distribution. But it is not so dependent on other aspects of the Melitz model. The dominance of the extensive margin extends to some environments with firm-market idiosyncratic shocks to demand and fixed costs (e.g., Eaton et al. (2011)), non-CES preferences (e.g., Arkolakis et al. (2015)), non-monopolistic competition (e.g., Jensen et al. (2003)), multi-national production (e.g., Arkolakis et al. (2014)), and convex marketing costs (e.g., Arkolakis (2010)).

It has been difficult to test the sharp prediction of the Melitz-Pareto model because most firm-level empirical studies have only a few exporting origins or importing destinations. Eaton et al. (2008) analyze firm-level exports for Colombia, Eaton et al. (2011) France, Eaton et al. (2012a) (henceforth EKS) Denmark and France, Manova and Zhang (2012) China, and Arkolakis and Muendler (2013) Brazil, Chile, Denmark and Norway. With just one destination or origin, it is possible that exports vary on the intensive margin because of differences in fixed costs of exporting to a given market, rather than because of any deviation from Pareto-distributed firm productivity.[1]

[1] One could look at whether firm exports from a given origin follow a Pareto distribution. But this begs the question of whether the deviation is widespread across origins and, more to our point, generates a lot of variation on the intensive margin.

We use the World Bank's *Exporter Dynamics Database* (hereafter EDD) to examine the importance of the extensive and intensive margins empirically. The EDD covers firm-level exports from 50 developing countries to all destination countries in most years from 2003–2013.[2] Because the EDD contains so many different origin-destination pairs, we can control for both origin-specific and destination-specific fixed trade costs. Controlling for such fixed trade costs, the elasticity of the intensive margin with respect to overall exports is around 50 percent in the EDD. Considering mean exports in each percentile of a country's exporter size distribution, the intensive margin elasticity with respect to overall exports rises steadily from around 20 percent for the smallest exporters to over 50 percent for the largest exporters.

We next explore potential explanations for the positive intensive margin elasticity (i.e., the tendency of big exports per exporter to go along with big overall exports). First, we consider the possibility that fixed trade costs vary by origin-destination pairs. The problem with this story is that higher fixed trade costs raise average exports per exporter, but also *lower* overall exports. To explain why the intensive margin is increasing in overall exports, one would need 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, one would need the fixed trade costs to fall with distance between trade partners. Moreover, this explanation would require that the intensive margin elasticity be equally important for the smallest and the largest exporters, rather than rising steadily with exporter size percentile.

Second, we explore the role of multi-product firms. If the typical firm exports more products where overall export flows are larger, this could account for the importance of the intensive margin for exports. We show that the number of HS 6-digit products per exporting firm do indeed account for about 12 percent of the variation in overall exports, or about one-fourth of the 50 percent contri-

[2] Within exporting firms and their destination countries, the exports can be broken into products at least down to the HS 6-digit level. Fernandes et al. (2015) describe the dataset in detail.

bution of the intensive margin to overall exports. However, in the context of the multi-product Melitz-Pareto model developed by Bernard, Redding and Schott (2011), 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 continuum of firms, the extensive margin should explain all variation in overall exports (conditional on fixed costs) in the Melitz-Pareto model. With a finite number of firms, however, the intensive margin (and overall exports) can be high because of good productivity draws from the Pareto distribution within a country. We develop a general method of moments (GMM) estimator for the elasticity of fixed trade costs to distance that is valid under granularity as in Eaton et al. (2012a). This estimator yields a negative elasticity, implying that even under granularity fixed trade costs are required to 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.

Lastly, we depart from the Pareto distribution to 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. With this motivation, we show that lognormally-distributed firm productivity (with realistic dispersion) 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). Shifting to lognormal productivity also changes our inference about fixed trade costs, rendering them positively correlated with variable trade costs and *rising* with distance. It also implies that, as in the data, the intensive margin elasticity rises steadily with the size percentile of exporters.

The rest of the paper is organized as follows. In Section 2 we formalize the predictions of the Melitz-Pareto model. Section 3 describes the EDD data. In Section 4 we document the empirical importance of the intensive margin in accounting for cross-country variation in exports. Sections 5 and 6 explore how allowing for multi-product firms and granularity, respectively, affects the implications of the Melitz-Pareto model for the intensive margin and contrast those implications against the data. Finally, in Section 7 we study how the implications of the Melitz model change when we drop the Pareto assumption and assume instead that the productivity distribution is lognormal. Section 8 concludes.

2. Properties of the Melitz-Pareto Model

We start with the Melitz model with a continuum of single-product firms with a Pareto distribution for productivity as in Chaney (2008) and Arkolakis et al. (2008). 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 denoted by σ and common across countries. Each firm produces one variety under monopolistic competition. Firm-level productivity φ is distributed Pareto with shape parameter $\theta > \sigma - 1$, $\Pr(\varphi \leq \varphi_0) = G_i(\varphi_0) = 1 - (\varphi_0/b_i)^{-\theta}$. Firms from country i also incur fixed trade costs F_{ij} as well as iceberg trade costs τ_{ij} to sell in country j. In each country i there is a large pool of prospective entrants which can pay an entry cost F_i^e to draw a productivity φ

from $G_i(\varphi)$.[3]

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

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

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, N_i is the total number of firms in i, $\bar{\sigma} \equiv \sigma / (\sigma - 1)$ is the markup, and φ_{ij}^* is the productivity cutoff for exports from i to j, which is defined implicitly by

$$x_{ij}(\varphi_{ij}^*) = \sigma F_{ij}. \tag{2}$$

The value of overall exports from i to j is then

$$X_{ij} = N_i \int_{\varphi \geq \varphi_{ij}^*} x_{ij}(\varphi) dG_i(\varphi). \tag{3}$$

The free-entry condition entails $F_i^e = \sum_j \left[(X_{ij}/\sigma) - F_{ij}\left(1 - G_i(\varphi_{ij}^*)\right)\right]$. Combined with Equations (3) and (2) and using the fact that $G_i(\varphi)$ is Pareto, this implies $N_i = \frac{\sigma-1}{\sigma\theta}\frac{L_i}{F_i^e}$. The equilibrium is a set of wages w_i such that $w_i L_i = \sum_j X_{ij}$.

The number of firms from i that export to j is $N_{ij} = N_i \int_{\varphi \geq \varphi_j^*} dG_i(\varphi)$. Using again the fact that $G_i(\varphi)$ is Pareto we get from (3) that

$$X_{ij} = \left(\frac{\theta}{\theta - (\sigma - 1)}\right) A_j \left(w_i \tau_{ij}\right)^{1-\sigma} b_i^\theta N_i \left(\varphi_{ij}^*\right)^{\sigma-\theta-1} \tag{4}$$

and

$$N_{ij} = b_i^\theta N_i \left(\varphi_{ij}^*\right)^{-\theta}. \tag{5}$$

Combining (2), (4) and (5), the extensive margin is

$$N_{ij} = N_i \left(\frac{w_i}{b_i}\right)^{-\theta} \left(\frac{\sigma}{A_j}\right)^{-\theta/(\sigma-1)} \tau_{ij}^{-\theta} F_{ij}^{-\theta/(\sigma-1)}, \tag{6}$$

[3]Both F_{ij} and F_i^e are 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.

while the intensive margin is

$$x_{ij} \equiv \frac{X_{ij}}{N_{ij}} = \left(\frac{\theta\sigma}{\theta-(\sigma-1)}\right) F_{ij}. \tag{7}$$

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 (6) and (7), and defining variables appropriately, we have

$$\ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \widetilde{\tau}_{ij} - \bar{\theta} \ln \widetilde{F}_{ij} \tag{8}$$

and

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \widetilde{F}_{ij}, \tag{9}$$

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.

Consider the OLS regression of $\ln x_{ij}$ on $\ln X_{ij}$ with origin and destination fixed effects,

$$\ln x_{ij} = FE_i^o + FE_j^d + \alpha \ln X_{ij} + \varepsilon_{ij}. \tag{10}$$

The estimated regression coefficient is given by

$$\hat{\alpha} = \frac{cov(\ln \tilde{x}_{ij}, \ln \tilde{X}_{ij})}{var\left(\ln \tilde{X}_{ij}\right)}. \tag{11}$$

where $\ln \tilde{z}_{ij}$ denotes variable $\ln z_{ij}$ demeaned by origin and destination fixed effects, and likewise for other variables. The Intensive Margin Elasticity (IME) is IME $\equiv \hat{\alpha}$. The extensive margin elasticity is EME $= \frac{cov(\ln \tilde{N}_{ij}, \ln \tilde{X}_{ij})}{var\left(\ln \tilde{X}_{ij}\right)} = 1 -$ IME.

According to the model (i.e., using equations 8 and 9), the coefficient $\hat{\alpha}$ is given by

$$\text{IME} = -\frac{\left(\bar{\theta}-1\right) var(\ln \widetilde{F}_{ij}) + \theta cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})}{var\left(-\theta \ln \widetilde{\tau}_{ij} - \left(\bar{\theta}-1\right) \ln \widetilde{F}_{ij}\right)}. \tag{12}$$

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, 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 $= 0$.

Combined with the assumption that $\bar{\theta} > 1$, the result in (12) 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 9 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: If $\frac{cov(\ln \tilde{x}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})} < 0$ then $\frac{cov(\ln \widetilde{F}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})} < 0$.

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. Denote average exports per firm in percentile pct as x_{ij}^{pct} and consider the following regression:

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

and define the IME for each percentile as IME$^{pct} \equiv \hat{\alpha}^{pct}$. The previous argument implies that the intensive magin elasticity calculated separately for each percentile is the same as the overall intensive margin elasticity.[4]

[4]Observation 4 no longer holds when fixed trade costs vary across firms. However, simula-

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

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 8 and 9 to solve for $\ln \widetilde{F}_{ij}$ and $\ln \widetilde{\tau}_{ij}$ in terms of $\ln x_{ij}$ and $\ln N_{ij}$ yields

$$\ln \widetilde{F}_{ij} = \delta_i^{F,o} + \delta_j^{F,d} + \ln x_{ij} \tag{13}$$

and

$$\theta \ln \widetilde{\tau}_{ij} = \delta_i^{\tau,o} + \delta_j^{\tau,d} - \bar{\theta} \ln x_{ij} - \ln N_{ij}. \tag{14}$$

Model-implied values for $\ln \widetilde{F}_{ij}$ are (ignoring origin and destination fixed effects) directly given by $\ln x_{ij}$, but for $\ln \widetilde{\tau}_{ij}$ we need a value for $\bar{\theta}$ to go from $\ln x_{ij}$ and $\ln N_{ij}$ in the data to model-implied values for $\theta \ln \tilde{\tau}_{ij}$.[5]

3. World Bank *Exporter Dynamics Database*

We use the Exporter Dynamics Database (EDD) described in Fernandes et al. (2015) to study the intensive and extensive margins of trade using firm-level data. The EDD is based on customs data covering the universe of export transactions provided by customs agencies from 70 countries (56 developing and 14 developed countries). For each country, the raw customs data contains annual export flows disaggregated by firm, destination and Harmonized System (HS) 6-digit product, based on a time-consistent product classification. Oil exports (HS chapter 27) 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 COM-

tions in that case suggest that IME^{pct} may be downward sloping, which is contrary to what we observe in the data. This is something we plan to explore further.

[5]To show results for τ_{ij}, we need an estimate of $\bar{\theta}$. The Appendix outlines the procedure that we use, which follows that in Eaton et al. (2011). Alternatively, we appeal to estimates of θ and σ from Head and Mayer (2014) and Bas et al. (2015)

TRADE/WITS. The set of more than 100 statistics included in the EDD are publicly available at the exporting country-year, exporting country-product-year, exporting country-destination country-year, or exporting country-product-destination country-year levels for download, and include in particular, measures of average exports per firm as well as the number of exporters, which are comparable across countries.

In this paper we exploit the underlying exporter-level customs data for 49 countries included in the second release of the EDD as well as China.[6] Hence we use an unbalanced panel of 50 developing countries whose sample periods cover some or all of the years between 2003 and 2013, as shown in Table 1.

Using these data we calculate variants of average exports per firm, number of exporters, and total exports at the exporting country-destination country-year level or at the exporting country-product or industry-destination country-year level. We focus on products or industries belonging to the broad manufacturing sector. Specifically, using a concordance across the ISIC rev. 3 classification and the HS 6-digit classification we consider only HS 6-digit codes that correspond to ISIC manufacturing sub-sectors 15-37. The product disaggregations that we use are HS 2-digit, HS 4-digit, or HS 6-digit. The industries that we use are groups of HS 2-digit codes as defined in Table 2. For exporter size percentiles, we calculate average exports per firm based on firms within each size percentile at the exporting country-destination country-year level.

4. The Melitz-Pareto Model vs. the Data

We start by plotting the intensive margin (x_{ij}), the extensive margin (N_{ij}), and total exports (X_{ij}) from all countries in the EDD to four large destinations: the United States (US), Japan, Germany and France for each year – see Figure 1.[7]

[6]China is not included in the EDD due to confidentiality concerns, but statistics based on exporter-level customs data for China can be used in this paper.

[7]We restrict destination countries to be the US, Japan, Germany and France to reduce noise associated with country pairs with few exporters. We further restrict the sample to the 676 coun-

Each dot corresponds to (x_{ij}, X_{ij}) (panel a) or (N_{ij}, X_{ij}) (panel b). The plotted line represents the prediction of the Melitz-Pareto model with a continuum of firms and fixed trade costs that vary by origin and destination but not across country pairs, i.e., $var(\widetilde{F}_{ij}) = 0$. Observation 1 in Section 2 indicates this is a horizontal line for the intensive margin (panel a) and has a unit slope for the extensive margin (panel b). The IME is the slope of the regression line running through the data in panel a – it corresponds to $\hat{\alpha}$ from regression 10. The intensive margin elasticity is clearly positive while the extensive margin elasticity is clearly less than one.

We supplement Figure 1 with Table 3, which shows results for the IME from the benchmark regression 10 with several alternative specifications. In the benchmark specification it ranges between 0.45 and 0.52, with the latter being our preferred estimate due to the inclusion of origin and destination fixed effects. The results are robust in wider samples. When all destinations are included, the IME estimate only falls to 0.43 in the third column of Table 4. In a sample with country pairs with less than 100 exporters, IME estimates remain higher than 0.48, as reported in Table 5, and reach 0.67 when origin and destination fixed effects are included.

As a further robustness check, we estimated the IME based on data with an industry disaggregation for each country pair each year. It could be the case that average exports per firm are constant across different countries within the same industry, but differ across industries. Positive IME estimates could thus arise from a different sectoral composition of exports across country pairs, but be consistent with the basic Melitz-Pareto model at the sector level. We show, however, that this is not the case. Consider Figure 2 where we plot the intensive and extensive margins against total exports at the exporting country-industry-destination country-year level. The slope of the regression line going through the data is still positive for the intensive margin. Table 6 shows that the IME

try pairs with more than 100 exporters (i.e., ij pairs for which $N_{ij} > 100$ when j is US, Japan, Germany or France).

actually increases when moving from aggregate to disaggregated data, implying that the sectoral composition of trade is not an explanation for the puzzle. At the lowest level of aggregation available, which is HS 6-digit products, the IME is as high as 0.71.

Finally, we check whether the results are driven by the prevalence of small firms in the data. To do this, we calculate average exports per firm, number of firms and total exports using only data on exporters whose annual exports were at least $1,000. The results reported in Table 7 change only slightly, with the IME never falling below 0.42.

We further explore the implications of a positive IME for fixed and variable trade costs. Using Equations (13) and (14) we uncover model-implied fixed and variable trade costs components that vary by origin-destination pairs. As per Observation 2, the Melitz-Pareto model can have a positive IME, but this would imply a negative correlation between fixed and variable trade costs. Indeed, Table 8 reports that this correlation is negative and strong.

Moreover, Observation 3 implies that fixed trade costs must decline with distance if $cov(\ln \widetilde{x}_{ij}, \ln \widetilde{dist}_{ij}) < 0$. We plot fixed trade costs against distance in Figure 3. As expected from the positive sign of the IME and the fact that total exports fall with distance, the figure shows that fixed trade costs are decreasing with distance, while variable trade costs are increasing with distance. This finding is puzzling, since one would expect trade barriers to increase rather than fall with distance. Table 9 confirms that the fixed trade cost elasticity with respect to distance, reported in the first column, is negative and statistically significant.

We also check the implications of Observation 4 by plotting the IME^{pct} for each exporter size percentile in Figure 4. The horizontal line corresponds to the theoretical prediction of a common elasticity across percentiles while the dots represent the estimated elasticities for each percentile. Contrary to Observation 4, the IME is not constant across percentiles, it is significantly lower than the overall IME of 0.52 for the first 90 percentiles, and is increasing gradually. The largest IME is observed in the top percentile.

To conclude, in the data 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, disaggregation and exclusion of small firms, and is at odds with the simple version of the Melitz-Pareto model with fixed trade costs varying only because of origin and destination effects. 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. The data is also at odds with the implication from the Melitz-Pareto model of a constant IME across percentiles.

5. Multi-Product Extension of Melitz-Pareto

In Section 4 we show that around half of the variation in bilateral exports is explained by the firm-level intensive margin. This contradicts the basic Melitz-Pareto model with a continuum of single-product firms and no variation in bilateral fixed trade costs, $var(\widetilde{F}_{ij}) = 0$. One can always assume that bilateral fixed trade costs are such that the model matches the data, but this leads to the awkward implications that the covariance between fixed trade costs and either variable trade costs or distance is negative. In this section we explore whether the contradictions between the model and the data are maintained for a popular extension of the Melitz-Pareto model to multi-product firms. The idea is that what the basic Melitz-Pareto model considers as the intensive margin is really an extensive margin operating at the level of the firm. Thus, perhaps the reason why average exports per firm fall with distance in the data is that firms export fewer products (because of higher product-level fixed trade costs) even though they export more per product to more distant destinations.

5.1. 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 i with productivity φ that wants to export to j, $\lambda^*_{ij}(\varphi)$, is given implicitly by

$$A_j \left(\frac{w_i \tau_{ij}}{\varphi \lambda^*_{ij}(\varphi)} \right)^{1-\sigma} = \sigma f_{ij}. \tag{15}$$

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

$$\pi_{ij}(\varphi) \equiv \int_{\lambda^*_{ij}(\varphi)}^{\infty} \left[\left(\frac{\lambda}{\lambda^*_{ij}(\varphi)} \right)^{\sigma-1} - 1 \right] f_{ij} dG^p(\lambda). \tag{16}$$

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 $f_{ij} = f^o_i f^d_j \tilde{f}_{ij}$,

$F_{ij} = F_i^o F_j^d \tilde{F}_{ij}$, and $\tau_{ij} = \tau_i^o \tau_j^d \tilde{\tau}_{ij}$, and defining variables appropriately we get

$$\ln X_{ij} = \mu_i^{X,o} + \mu_j^{X,d} - \theta^f \ln \widetilde{\tau}_{ij} - \left(\frac{\theta^f}{\sigma - 1} - \frac{\theta^f}{\theta^p}\right) \ln \widetilde{f}_{ij} - \left(\frac{\theta^f}{\theta^p} - 1\right) \ln \widetilde{F}_{ij}, \quad (17)$$

$$\ln x_{ij}^p \equiv \ln X_{ij} - \ln M_{ij} = \mu_i^{x^p,o} + \mu_j^{x^p,d} + \ln \widetilde{f}_{ij}, \quad (18)$$

and

$$\ln x_{ij}^f \equiv \ln X_{ij} - \ln N_{ij} = \mu_i^{x^f,d} + \mu_j^{x^f,d} + \ln \widetilde{F}_{ij}. \quad (19)$$

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.

From regressions similar to regression 10 we now have two intensive margin elasticities, one for the intensive margin defined over products and one for the intensive margin defined over firms:

$$\text{IME}^s \equiv \frac{cov(\ln \widetilde{x}_{ij}^s, \ln \widetilde{X}_{ij})}{var(\ln \widetilde{X}_{ij})}, \quad (20)$$

with $s = f, p$. Letting $\bar{\theta} \equiv \theta^f / (\sigma - 1)$ and $\chi \equiv \theta^f / \theta^p$, then from Equations (17) to (19) we have

$$\text{IME}^p = -\frac{(\bar{\theta} - \chi)\, var\left(\ln \widetilde{f}_{ij}\right) + (\chi - 1)\, cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) + \theta cov(\ln \widetilde{f}_{ij}, \ln \widetilde{\tau}_{ij})}{var(\ln \widetilde{X}_{ij})} \quad (21)$$

and

$$\text{IME}^f = -\frac{(\chi - 1)\, var\left(\ln \widetilde{F}_{ij}\right) + (\bar{\theta} - \chi)\, cov(\ln \widetilde{f}_{ij}, \ln \widetilde{F}_{ij}) + \theta cov(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij})}{var(\ln \widetilde{X}_{ij})} \quad (22)$$

Letting $m_{ij} \equiv M_{ij}/N_{ij}$ be the average number of products sold by firms from i in j, we have that $x_{ij}^f = x_{ij}^p m_{ij}$ and $X_{ij} = x_{ij}^p m_{ij} N_{ij}$. This implies that IME^f is equal to IME^p plus the extensive product margin elasticity, $cov(\ln \tilde{m}_{ij}, \ln \tilde{X}_{ij})/var(\ln \tilde{X}_{ij})$.

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 $\text{IME}^f > 0$ 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: If $cov(\ln \widetilde{x}^p_{ij}, \ln dist_{ij}) < 0$ then $cov(\ln \widetilde{f}_{ij}, \ln dist_{ij}) < 0$.

As in the single-product model, we can use the model to back out the implied trade costs. Equation (19) 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 (18) can be used to obtain a model-implied $\widetilde{f}_{ij}$.

5.2. Data

Table 10 reports the estimated product-level IME, IME^p. In our basic sample with the 4 large destinations the IME^p never falls below 0.37 (see panel a). When all destinations are included, the IME^p remains virtually unchanged (see panel b). Of course, the IME^f is the same as that reported in Section 4. While we could have the model match the data by allowing both firm-level and product-level fixed trade costs to vary appropriately across country pairs, from observations 6 and 7 this would imply that those fixed trade costs would be negatively correlated, or that the covariances between those fixed trade costs and variable trade costs would be negative.

Observation 8 indicates that if product-level fixed trade costs increase with distance, then a positive covariance between average exports per product-firm and distance should be observed. In fact, the opposite is verified in the data: as shown in Figure 5, $\tilde{x}_{ij}^{p}$ declines with distance,[8] implying that product-level fixed trade costs are also falling with distance, as was the case for firm-level fixed trade costs in Section 4 (see the third column of Table 9).

6. 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.

6.1. Theory

Eaton et al. (2012b) extend the Melitz-Pareto model above to allow for granularity. Equations 8 and 9 then become

$$\ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \widetilde{\tau}_{ij} - \bar{\theta} \ln \widetilde{F}_{ij} + \xi_{ij} \tag{23}$$

and

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \widetilde{F}_{ij} + \varepsilon_{ij}, \tag{24}$$

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, and assuming for the sake of argument that all covariances between $\ln \widetilde{\tau}_{ij}$ or $\ln \widetilde{F}_{ij}$ with ξ_{ij} or ε_{ij} are equal to zero, the previous

[8] Bernard, Redding and Schott (2011) report a positive but statistically insignificant coefficient for the regression of $\ln x_{ij}^{p}$ on distance (see their Table II). Their regression uses only data for U.S. exports data, so it doesn't include origin and destination fixed effects (they include a control for destination-country size).

equations imply that

$$IME = -\frac{\left(\bar{\theta}-1\right) var(\ln \widetilde{F}_{ij}) + \theta cov\left(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij}\right) - var(\varepsilon_{ij}) - cov(\varepsilon_{ij}, \xi_{ij})}{var\left(-\theta \ln \widetilde{\tau}_{ij} - \left(\bar{\theta}-1\right) \ln \widetilde{F}_{ij} + \varepsilon_{ij} + \xi_{ij}\right)}. \quad (25)$$

Even if $cov(\varepsilon_{ij}, \xi_{ij}) = 0$, since $var(\varepsilon_{ij}) > 0$, 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 $\kappa \equiv \frac{\sigma\bar{\theta}}{\bar{\theta}-1}$. 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

$$x_{ij} = \kappa F_i^o F_j^d e^{\zeta \ln dist_{ij}} + u_{ij}, \quad (26)$$

where $u_{ij} \equiv \upsilon_{ij} + \varepsilon_{ij}$ 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 $\mathbb{E}(\upsilon_{ij}|\ln dist_{ij})$ 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 proceed, we follow Charbonneau (2012) to derive a GMM estimate of ζ. To take advantage of the time dimension of our data, we extend 26 to allow for a time-specific component in the expression of fixed trade costs,

$$x_{ijt} = \kappa F_i^o F_j^d F_t e^{\zeta \ln dist_{ij}} + u_{ijt}, \tag{27}$$

where again $\mathbb{E}(u_{ijt}|dist_{ij}) = 0$. As is shown in the Appendix, manipulating equation 27 yields the moment condition

$$\begin{aligned} \mathbb{E}\Big[\Big(x_{ikt}x_{ljt} - x_{lkt}x_{ijt}e^{\zeta\left(-\ln dist_{ij}+\ln dist_{ik}+\ln dist_{lj}-\ln dist_{lk}\right)}\Big) \times \\ (-\ln dist_{ij} + \ln dist_{ik} + \ln dist_{lj} - \ln dist_{lk})\Big] = 0 \end{aligned} \tag{28}$$

In the next subsection we will use this moment condition to estimate ζ.

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

$$x_{ij}(\varphi, \alpha_j) = \sigma F_{ij} \left(\frac{\alpha_j \varphi}{\alpha_{ij}^* \varphi_{ij}^*}\right)^{\sigma-1}, \tag{29}$$

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 $F_{ij} = F_i^o F_j^d$. This will allow us to study the IME generated by granularity by itself.

3. Use Equation (29) 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 10 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 .

6.2. 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 (28) to estimate firm-level as well as product-level fixed trade cost elasticities with respect to distance (ζ). Table 11 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 even when granularity is taken into account. Hence granularity does not help to eliminate one of the puzzles emerging from the comparison between the Melitz-Pareto model and the data.

Table 12 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 $\bar{\theta} = 2.3$, 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 $\bar{\theta}$ ($\bar{\theta} = 2.3$) and with no demand shocks (so there is perfect correlation in firm-level productivity across destinations), the simulated IME of 0.002 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 0.37, not too far from our preferred estimate based on the data of 0.52. 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 6 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.

7. Lognormal

In this section we depart from the assumption of a common Pareto distribution of firm-level productivity and instead assume a lognormal distribution. 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.

7.1. 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. Following Bas et al. (2015) (henceforth BMT), let

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

Since $x_{ij}(\varphi) \propto \varphi^{\sigma-1}$ then the ratio of average to minimum exports for each country pair satisfies

$$x_{ij}/x_{ij}(\varphi_{ij}^{*}) = H(\varphi_{ij}^{*}). \tag{30}$$

If firm productivity is distributed Pareto with parameter $\theta > \sigma - 1$ (as in Section 2) then it is easy to show that $H(\varphi) = \frac{\theta}{\theta-(\sigma-1)}$ (see Equation 7), 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 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

$$G_i(\varphi) = \Phi\left(\frac{\ln\varphi - \mu_{\varphi,i}}{\sigma_\varphi}\right). \tag{31}$$

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

$$H\left(\varphi^*_{ij}\right) = \frac{h\left[-(\ln\varphi^*_{ij} - \mu_{\varphi,i})/\sigma_\varphi\right]}{h\left[-(\ln\varphi^*_{ij} - \mu_{\varphi,i})/\sigma_\varphi + \bar{\sigma}_\varphi\right]}, \tag{32}$$

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

$$\frac{x_{ij}}{x_{ij}(\varphi^*_{ij})} = \Omega\left(\frac{N_{ij}}{N_i}\right) \equiv \frac{h\left(\Phi^{-1}\left(\frac{N_{ij}}{N_i}\right)\right)}{h\left(\Phi^{-1}\left(\frac{N_{ij}}{N_i}\right) + \bar{\sigma}_\varphi\right)}. \tag{33}$$

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 (33) to get the model-implied average exports per firm (in logs),

$$\ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln\Omega\left(N_{ij}/N_i\right). \tag{34}$$

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

$$\ln\widetilde{F}_{ij} = \delta_i^{F,o} + \delta_j^{F,d} + \ln x_{ij} - \ln\Omega\left(\frac{N_{ij}}{N_i}\right). \tag{35}$$

Second, to compute $\widetilde{\tau}_{ij}$, we combine Equations (1), (2), (31) and (33) to get (with appropriately defined fixed effects)

$$(\sigma-1)\ln\widetilde{\tau}_{ij} = \delta_i^{\tau,o} + \delta_j^{\tau,d} - \ln x_{ij} + \ln\Omega_i\left(\frac{N_{ij}}{M_i}\right) + \bar{\sigma}_\varphi\Phi^{-1}\left(1-\frac{N_{ij}}{N_i}\right). \tag{36}$$

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.[9] We acknowledge slippage

[9]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 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 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 productivity shocks that vary by destination. This is similar to the setup in Eaton et al. (2011), except that we further assume that there is zero correlation between demand shocks and fixed costs shocks within destinations. More importantly, 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

between log number of firms in the BR dataset and log population with an elasticity of 0.945, as reported in Table 13 and in Figure 7.

each origin i:

$$\begin{bmatrix} \ln\varphi \\ \ln\alpha_1 \\ \vdots \\ \ln\alpha_J \\ \ln f_1 \\ \vdots \\ \ln f_J \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{\varphi,i} \\ \mu_\alpha \\ \vdots \\ \mu_\alpha \\ \mu_{f,i1} \\ \vdots \\ \mu_{f,iJ} \end{bmatrix}, \begin{bmatrix} \sigma_\varphi^2 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_\alpha^2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_\alpha^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \sigma_f^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & \sigma_f^2 \end{bmatrix}\right). \tag{37}$$

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 are assumed to be uncorrelated 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. The assumption of joint lognormality as well as independence between demand shocks and fixed trade costs makes the likelihood function tractable and the estimation procedure much faster. In future work we plan to explore relaxing the assumption of independence between demand shocks and fixed trade costs.

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 $D_{ij} \equiv \ln\left[A_j\left(w_i\tau_{ij}\right)^{1-\sigma}\right]$ and let $Z_{ij} \equiv D_{ij} + \ln\alpha_j + (\sigma-1)\ln\varphi$ be a latent variable which we observe only if a firm actually exports. We then have

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

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

$$\begin{bmatrix} Z_{i1} \\ \vdots \\ Z_{iJ} \end{bmatrix} \sim N\left(\begin{bmatrix} d_{i1} \\ \\ d_{iJ} \end{bmatrix}, \begin{bmatrix} \bar{\sigma}_\varphi^2 + \sigma_\alpha^2 & \cdots & \bar{\sigma}_\varphi^2 \\ \vdots & \ddots & \vdots \\ \bar{\sigma}_\varphi^2 & \cdots & \bar{\sigma}_\varphi^2 + \sigma_\alpha^2 \end{bmatrix}\right), \tag{38}$$

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 (38) as well as mean log fixed trade costs (up to a constant) and their dispersion using maximum likelihood methods. The density $g_{X_i}(x_i)$ is easy to write down for the case in which $X_{ij} > 0$ for all j except one, but otherwise it is computationally expensive to compute because of the correlation across Z_{ij} that arises from the common productivity term $\ln\varphi$. For now, we simplify the analysis by considering only data for two destinations, which we label $j = 1, 2$. To be consistent with our previous empirical exercises, we also restrict the sample to include only origins with at least 100 exporters to the two chosen destinations. The Appendix shows how to derive the density $g_{(X_{i1},X_{i2})}\left(x_{i1}, x_{i2}\right)$ for this case. We compute $g_{(X_{i1},X_{i2})}\left(x_{i1}, x_{i2}\right)$ for each observation in our dataset (which is a realization of $\{X_{i1}, X_{i2}\}$ that we observe). Since all random variables are independent across firms, we can compute the log-likelihood function as a sum of log-densities,

$$\ln L\left(\theta|\left\{x_{i1}\left(k_i\right),x_{i2}\left(k_i\right)\right\}_{i,k_i}\right)=\sum_i\sum_{k_i=1}^{N_i}\ln\left[g_{\left(X_{i1},X_{i2}\right)}\left(x_{i1}\left(k_i\right),x_{i2}\left(k_i\right)\right)\right], \tag{39}$$

where N_i is the number of firms from i that sell to either of the two destinations we consider, and where k_i is an index for a particular observation in our dataset (for origin i it takes values in $1,...,N_{i1}$) and θ is a vector of parameters that we want to estimate,

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

where $\bar{\mu}_{f,ij}=\ln\sigma+\mu_{f,ij}$. Because the likelihood is potentially not concave in θ and because there are 75 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 $\theta^{(n)}$[10], dropping the first 750k runs ("burn in" period) and then continuing until $n=3MM$. 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 σ_{α}: the more correlated firm sales are across destinations, the larger is σ_{φ} relative to σ_{α}. Finally, to understand how σ_f is identified, imagine for simplicity that there is only one

[10]We use Gibbs sampling technique. At each iteration $\theta^{(n)}$ and $\theta^{(n-1)}$ can differ by at most one element.

destination. We then have

$$g_{X_{i1}}(x_{i1}) = \frac{g_{Z_{i1}}(x_{i1}) \times \Pr\{F_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\}}{C}$$

where $C \equiv \Pr\{\mathrm{Z}_{i1} \leq F_{i1}\}$ 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 $\frac{\Pr\{F_{i1} \leq x_{i1} | Z_{i1} = x_{i1}\}}{C}$ 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.

7.2. Data

Simple Melitz model with lognormal distribution

Table 14 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, $\bar{\sigma}_{\varphi} = 4.55$, which corresponds to the subsample with the largest 25% of firms. [11]

We present the results for the IME (computed as explained above) for $\bar{\sigma}_{\varphi} = \{2.4; 4.55\}$ in Table 15. First, a lognormal distribution allows the intensive mar-

[11] 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.

gin elasticity to be positive even under the assumption of a continuum of firms. Second, for our estimate of the shape parameter $\bar{\sigma}_\varphi = 4.55$, the size of IME is very close to that from the data. Third, most of the action comes from the right tail of the exporter size distribution, as seen in Figure 8.

We use Equations (35) and (36) to compute the model-implied fixed and variable trade costs. The correlations between those costs and distance are reported in Table 16 and plotted in Figure 9. In contrast to our results under Pareto, now under lognormal both the model-implied variable and fixed trade costs are increasing with distance.

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 -0.9). 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

We estimate the parameters of the full Melitz-lognormal model using firm-level data from the EDD across different origins and two destinations, the US and Germany. An origin is included in the subsample only if it has more than 100 firms exporting to both destinations; 18 countries in our whole sample satisfy this condition.

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 10 shows a plot of the CDF for firm-level sales from one origin (name undisclosed for reasons of confidentiality) to the United States.[12]

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

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 the least popular destination would also export to the most popular destination. The share of firms that only sell in the least popular destination is then a measure of the extent to which this strict hierarchy predicted by the simplest model is violated. According to Figure 11, the share predicted by the estimated model is quite close to the one in the data.

Finally, Figure 12 shows the correlation in sales across the two destinations for firms from a given origin that sell in both destinations. The estimated model implies that this correlation is 0.39 (with tiny deviations across origins) while we see that this correlation exhibits some dispersion across our 18 origin countries. However, note that the country with the most firms lines up right on the 45 degree line, indicating that its implied correlation is the same as in the estimated model, while the origins with correlations that deviate mostly from the estimated model are those with few firms.

The results of the estimation for the dispersion parameters $(\bar{\sigma}_\varphi, \sigma_\alpha, \sigma_f)$ are shown in Table 17. The estimated values for $\bar{\sigma}_\varphi$ and σ_α are close to 2.6, while the estimate for σ_f is close to 3.6, all with very tight 95% confidence intervals. The estimate of $\bar{\sigma}_\varphi$ is close to the estimate of 2.4 in BMT, and much lower than the one we estimated in the simple lognormal model of 4.5. 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 $(\bar{\sigma}_\varphi^2+\sigma_\alpha^2)^{1/2} = 3.7$, although selection due to fixed trade costs brings the implied standard deviation for actual sales down to around 3, which is what we we observe in the data.

Table (18) and Figure (13) 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 1MM 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 1MM 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 (i.e., 0.58) is actually higher than the one in the data, although the difference is not statistically significant. [13] We plot the associated IME for each percentile in Figure (13) – the pattern of the IME across percentiles is remarkably close to what we see in the data.

In Table (18) we also report the IME implied by the estimated model when we suppress the systematic variation in fixed trade costs that does not come from origin and destination fixed effects (i.e., we set $\bar{\mu}_{f,ij}$ equal to $\delta_i^o + \delta_j^d$ coming from running the regression $\bar{\mu}_{f,ij} = \delta_i^o + \delta_j^d + \varepsilon_{ij}$). Interestingly, the IME actually increases in this case, revealing that the variation in $\bar{\mu}_{f,ij}$ not coming from origin and destination fixed effects actually is lowering the IME, which is exactly what we should expect if variable and fixed trade costs are positively correlated. We check this directly by computing the correlation between t_{ij} and $\bar{\mu}_{f,ij}$ (after removing origin and destination fixed effects from both) for a random sample of 50,000 values of the $\theta^{(n)}$ in the chain of estimates of θ. As shown in Table (19), the correlation is now positive and highly significant. This Table also shows the elasticity of variable and fixed trade costs with respect to distance (ignoring origin and destination fixed effects), computed in the same way. We see that now both types of trade costs are strongly increasing in distance.

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

[13] The confidence interval in Table Table (18) 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.

in the EDD data 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.

8. 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* to test this prediction. The EDD covers 50 countries for varying subsets of 2003–2013. Compared to existing studies, the EDD 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 about 50 percent of 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 one-half of this. This finding is robust to looking at all destinations or only the largest destinations, is robust to including all firms or ignoring very small firms, is robust to including all country pairs or only ones for which more than 100 firms export, and is robust to dissagregating across industries.

Although variation in fixed trade costs can make the Melitz-Pareto model fit this fact, this requires fixed trade costs that are negatively correlated with variable trade costs and with distance, and the resulting model does not reproduce the pattern for the IME across exporter percentiles. Allowing firms to export multiple products or taking into account granularity does not reverse these implications.

In contrast, moving away from a Pareto distribution and assuming that the productivity distribution is lognormal resolves the puzzles. A Melitz model with lognormally distributed firm productivity, demand shocks and fixed costs, estimated using maximum likelihood methods on the EDD firm-level data is consistent with the positive IME and its pattern across exporter percentiles while exhibiting fixed trade costs that are positively correlated with variable trade costs and distance.

References

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.

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.

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

Charbonneau, Karyne, "Multiple fixed effects in nonlinear panel data models," 2012.

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

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.

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*, 2015.

Head, Keith and Thierry Mayer, "Gravity Equations: Workhorse, Toolkit, and Cookbook, chapter to appear in the Handbook of International Economics Vol. 4, eds. Gopinath, Helpman, and Rogoff," 2014.

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

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

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.

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

$$\frac{x_{il|j}}{x_{il|l}} = \left(\frac{N_{ij}}{N_{il}}\right)^{-1/\bar{\theta}} \tag{40}$$

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

$$z_{ij} \equiv \frac{x_{il^*(i)|j}}{x_{il^*(i)|l^*(i)}} \tag{41}$$

and

$$m_{ij} \equiv \frac{N_{ij}}{N_{il^*(i)}} \tag{42}$$

then we have

$$\ln z_{ij} = -\frac{1}{\bar{\theta}} \ln N_{ij}. \tag{43}$$

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 A1 reproduces Figure 3 from Eaton et al. (2011) by plotting the variables in Equation (43). The slope in the graph is equal to $1/\bar{\theta}$, and the corresponding estimated values are reported in Table A1. Based on the full sample and using no weighting, the estimated $\bar{\theta}$ is over 19. But in Figure A1 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} is a measure of firms that sell from country i 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.

GMM Estimation of the Distance Elasticity of Fixed Trade Costs under Granularity

Let $f_i^o \equiv \ln F_i^o$, $f_j^d \equiv \ln F_j^d$, and $f_t \equiv \ln F_t$. From equation (27) it follows that

$$x_{ijt} e^{-\zeta \ln dist_{ij}} = \kappa e^{f_i^o + f_j^d + f_t} + u_{ijt} e^{-\zeta \ln dist_{ij}} \tag{44}$$

$$e^{\zeta \ln dist_{ik}} = x_{ikt} \kappa^{-1} e^{-f_i^o - f_k^d - f_t} - u_{ikt} \kappa^{-1} e^{-f_i^o - f_k^d - f_t} \tag{45}$$

Multiply equations (44) and (45) to get

$$x_{ijt} e^{-\zeta \ln dist_{ij} + \zeta \ln dist_{ik}} = x_{ikt} e^{f_j^d - f_k^d} - u_{ik} e^{f_j^d - f_k^d} + u_{ijt} \kappa^{-1} e^{-f_i^o - f_k^d - f_t - \zeta \ln dist_{ij}} \left[x_{ikt} - u_{ikt} \right] \tag{46}$$

and

$$x_{lkt} e^{f_j^d - f_k^d} - u_{lkt} e^{f_j^d - f_k^d} + u_{ljt} \kappa^{-1} e^{-f_l^o - f_k^d - \zeta \ln dist_{lj}} \left[x_{lkt} - u_{lkt} \right] = x_{ljt} e^{-\zeta \ln dist_{lj} + \zeta \ln dist_{lk}} \tag{47}$$

Multiply equations (46) and (47) to get

$$x_{ijt}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}} \left\{x_{lkt}e^{f_j^d-f_k^d} - u_{lkt}e^{f_j^d-f_k^d} + u_{ljt}\kappa^{-1}e^{-f_l^o-f_k^d-\zeta \ln dist_{lj}} \left[x_{lkt} - u_{lkt}\right]\right\} =$$
$$x_{ljt}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}} \left\{x_{ikt}e^{f_j^d-f_k^d} - u_{ikt}e^{f_j^d-f_k^d} + u_{ijt}\kappa^{-1}e^{-f_i^o-f_k^d-\zeta \ln dist_{ij}} \left[x_{ikt} - u_{ikt}\right]\right\} \tag{48}$$

Dividing both sides of (48) by $e^{f_j^d-f_k^d}$

$$x_{ijt}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}} \left\{x_{lkt} - u_{lkt} + u_{ljt}\kappa^{-1}e^{-f_l^o-f_j^d-\zeta \ln dist_{lj}} \left[x_{lkt} - u_{lkt}\right]\right\} =$$
$$x_{ljt}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}} \left\{x_{ikt} - u_{ikt} + u_{ijt}\kappa^{-1}e^{-f_i^o-f_j^d-\zeta \ln dist_{ij}} \left[x_{ikt} - u_{ikt}\right]\right\} \tag{49}$$

Rearranging (49) yields

$$\begin{aligned} &x_{lkt}x_{ijt}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}} - u_{lkt}x_{ijt}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}} + \\ &x_{ijt}\kappa^{-1}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}-\zeta \ln dist_{lj}-f_l^o-f_j^d}u_{ljt}\left[x_{lkt} - u_{lkt}\right] = \\ &x_{ikt}x_{ljt}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}} - u_{ikt}x_{ljt}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}} + \\ &x_{ljt}\kappa^{-1}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}-\zeta \ln dist_{ij}-f_i^o-f_j^d}u_{ijt}\left[x_{ikt} - u_{ikt}\right] \end{aligned} \tag{50}$$

Taking conditional expectation of (50) yields

$$\mathbb{E}\left[x_{lkt}x_{ijt}e^{-\zeta \ln dist_{ij}+\zeta \ln dist_{ik}}|\ln \mathbf{dist}\right] = \mathbb{E}\left[x_{ikt}x_{ljt}e^{-\zeta \ln dist_{lj}+\zeta \ln dist_{lk}}|\ln \mathbf{dist}\right]$$

This gives us the moment condition

$$\mathbb{E}\Big[\Big(x_{ikt}x_{ljt} - x_{lkt}x_{ijt}e^{\zeta\left(-\ln dist_{ij}+\ln dist_{ik}+\ln dist_{lj}-\ln dist_{lk}\right)}\Big)$$
$$\left(-\ln dist_{ij} + \ln dist_{ik} + \ln dist_{lj} - \ln dist_{lk}\right)|t\Big] = 0$$

Taking expectation with respect to time then gives us the moment condition in Equation 28.

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:

$$Q^E_{ij,n} = \ln x_{ij,n} \tag{51}$$

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

$$Q^T_{ij,n} = \bar{\mu}_{\varphi,ij} + \bar{\sigma}_\varphi \Phi^{-1}\left(\hat{\Phi}_{ij,n}\right), \tag{52}$$

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

$$\ln x_{ij,n} = \alpha_{ij} + \beta\Phi^{-1}\left(\hat{\Phi}_{ij,n}\right) + \varepsilon_{ij,n}. \tag{53}$$

Table 14 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 14. 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

Without risk of confusion, let's change notation and use $F_{ij} \equiv \ln \sigma + \ln f_i$. The density function for the case in which we consider two destinations is

$$g_{(X_{i1},X_{i2})}(x_{i1},x_{i2}) = \frac{1}{C} \times \left[g_{(Z_{i1},Z_{i2})}(x_{i1},x_{i2}) \Pr \left\{ \begin{array}{ccc} F_{i1} \leq x_{i1} & & Z_{i1} = x_{i1} \\ & | & \\ F_{i2} \leq x_{i2} & & Z_{i2} = x_{i2} \end{array} \right\} \right]^{I(x_{i1},x_{i2} \neq \emptyset)} \tag{54}$$

$$\times \left[g_{Z_{i1}}(x_{i1}) \Pr \left\{ \begin{array}{c} F_{i1} \leq x_{i1} \\ \\ F_{i2} > Z_{i2} \end{array} \mid Z_{i1} = x_{i1} \right\} \right]^{I(x_{i1} \neq \emptyset; x_{i2} \emptyset)}$$

$$\times \left[g_{Z_{i2}}(x_{i2}) \Pr \left\{ \begin{array}{c} F_{i1} > Z_{i1} \\ \\ F_{i2} \leq x_{i2} \end{array} \mid Z_{i2} = x_{i2} \right\} \right]^{I(x_{i2} \neq \emptyset; x_{i1} = \emptyset)},$$

where $g_{(Z_{i1},Z_{i2})}(x_{i1},x_{i2})$ is the joint density of latent variables Z_{ij}, which can be computed using distributional assumptions stated in (38), and $g_{Z_{ij}}(x_{ij})$ are the associated marginal densities. Our assumptions guarantee that $g_{(Z_{i1},Z_{i2})}(x_{i1},x_{i2})$ will be joint-normal. Note that the probabilities $\Pr\{\cdot|\cdot\}$ in (54) reflect the probability that a firm is an exporter to some market and a non-exporter to the other market conditional on observable sales. Finally, C is the probability that a firm exports to at least 1 destination; it can also be expressed as a CDF of joint normal distribution. We can calculate (54) for each observation in out dataset (which is a realization of (X_{i1}, X_{i2}) that we observe). We compute the term

$$\Pr \left\{ \begin{array}{ccc} F_{i1} \leq x_{i1} & & Z_{i1} = x_{i1} \\ & | & \\ F_{i2} \leq x_{i2} & & Z_{i2} = x_{i2} \end{array} \right\},$$

by noting that this is equal to $G_{(F_{i1},F_{i2})}(x_{i1},x_{i2})$ (by independence between F and Z) and then using

$$\begin{bmatrix} F_{i1} \\ F_{i2} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{f,i1} \\ \mu_{f,i2} \end{bmatrix}, \begin{bmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_f^2 \end{bmatrix}\right).$$

The other two conditional probabilities require a bit more work. First, note that

$$\Pr\left\{\begin{array}{c} F_{i1} \leq x_{i1} \\ F_{i2} > Z_{i2} \end{array} \mid Z_{i1} = x_{i1}\right\} = G_{\begin{bmatrix} F_{i1} \\ Z_{i2} - F_{i2} \end{bmatrix}_{|Z_{i1}=x_{i1}}}\begin{pmatrix} x_{i1} \\ 0 \end{pmatrix}$$

To compute the object on the RHS, we use the fact that

$$\begin{bmatrix} F_{i1} \\ Z_{i2} - F_{i2} \\ Z_{i1} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_{f,i1} \\ d_{i2} - \mu_{f,i2} \\ d_{i1} \end{bmatrix}, \begin{bmatrix} \sigma_f^2 & 0 & 0 \\ 0 & \widetilde{\sigma}_\varphi^2 + \sigma_\alpha^2 + \sigma_f^2 & \widetilde{\sigma}_\varphi^2 \\ 0 & \widetilde{\sigma}_\varphi^2 & \widetilde{\sigma}_\varphi^2 + \sigma_\alpha^2 \end{bmatrix}\right).$$

Using the properties of conditional normal distribution we can get

$$G_{\begin{bmatrix} F_{i1} \\ Z_{i2} - F_{i2} \end{bmatrix}_{|Z_{i1}=x_{i1}}}\begin{pmatrix} x_{i1} \\ 0 \end{pmatrix} = G_{\begin{bmatrix} Y_{i1} \\ Y_{i2} \end{bmatrix}_{|Z_{i1}=x_{i1}}}\left(\begin{bmatrix} x_{i1} \\ 0 \end{bmatrix} - \left(\begin{bmatrix} \mu_{f,i1} \\ d_{i2} - \mu_{f,i2} \end{bmatrix} + \beta\left[x_{i1} - d_{i1}\right]\right)\right),$$

where

$$\beta \equiv \begin{bmatrix} 0 \\ \\ \bar{\sigma}_{\varphi}^2 \end{bmatrix} \left[\bar{\sigma}_{\varphi}^2 + \sigma_{\alpha}^2\right]^{-1}.$$

The other conditional probability is computed analogously.

Table 1: EDD sample countries and years

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	2005	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: Industry groupings

	HS 2-digit codes that are included in each industry
Food and beverages	02-05; 07-12; 14-24
Mineral	25-26
Chemicals	28-38
Plastic and rubber	39-40
Apparel	41-43; 60-67
Wood	44-46
Paper	47-49
Textiles	50-59
Glass	68-70
Precious metals	71
Metals	72-83
Machinery	84
Electrical machinery	85
Transport equipment	86-89
Miscellaneous	90-96

Table 3: Benchmark IME regression

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
IM elasticity	0.459***	0.452***	0.522***
Standard error	[0.0135]	[0.0146]	[0.0127]
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes

Note: 4 main destinations, $N_{ij} > 100$, 676 obs.

Robust standard errors in brackets

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

Table 4: IME regression, all destinations

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
IM elasticity	0.501***	0.426***	0.429***
Standard error	[0.00553]	[0.00621]	[0.00521]
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes

Note: all destinations, $N_{ij} > 100$, 7211 obs.

Robust standard errors in brackets

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

Table 5: IME regression, countries with low N_{ij} included

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
IM elasticity	0.484***	0.494***	0.670***
Standard error	[0.00730]	[0.00775]	[0.0120]
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes

Note: 4 main destinations, 1485 obs.

Robust standard errors in brackets

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

Table 6: IME regression, disaggregated

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
Panel a: manufacturing industries			
IM elasticity	0.589***	0.590***	0.678***
Standard error	[0.00813]	[0.00814]	[0.00747]
Observations	2620	2620	2620
Panel b: HS2 (within manufacturing)			
IM elasticity	0.614***	0.623***	0.719***
Standard error	[0.00570]	[0.00575]	[0.00693]
Observations	3752	3752	3752
Panel c: HS4 (within manufacturing)			
IM elasticity	0.661***	0.671***	0.707***
Standard error	[0.00484]	[0.00515]	[0.00555]
Observations	7298	7298	7298
Panel d: HS6 (within manufacturing)			
IM elasticity	0.671***	0.684***	0.712***
Standard error	[0.00469]	[0.00496]	[0.00550]
Observations	9514	9514	9514
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes
Industry/HS FE	Yes	Yes	Yes

Note: 4 main destinations, $N_ij > 100$

Robust standard errors in brackets

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

Table 7: IME regression, small firms excluded

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
Panel a: 4 main destinations, $N_{ij} > 100$			
IM elasticity	0.455***	0.446***	0.524***
Standard error	[0.0136]	[0.0148]	[0.0124]
Observations	673	673	673
Panel b: all destinations, $N_{ij} > 100$			
IM elasticity	0.501***	0.423***	0.427***
Standard error	[0.00559]	[0.00632]	[0.00523]
Observations	7128	7128	7128
Panel c: 4 main destinations			
IM elasticity	0.480***	0.489***	0.671***
Standard error	[0.00727]	[0.00772]	[0.0119]
Observations	1485	1485	1485
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes

Firms with sales less than $1000 excluded

Robust standard errors in brackets

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

Table 8: Fixed and variable trade costs

	$corr(\ln N_{ij}, \ln x_{ij})$	$corr(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})$
Raw data	0.366	
	[0.035]	
Purged of:		
Origin FE	0.500	
	[0.033]	
Destination FE	0.352	
	[0.036]	
Origin and Destination FE	0.418	-0.891
	[0.034]	[0.017]

Note: $\theta = 5$, $\sigma = 5$, $N_{ij} > 100$, 676 obs.

Standard errors in brackets

Table 9: Trade costs and distance

	$\ln \widetilde{F}_{ij}$	$\ln \widetilde{\tau}_{ij}$	$\ln \widetilde{f}_{ij}$
$\ln dist_{ij}$	-0.440***	0.317***	-0.211***
Standard error	[0.0539]	[0.0168]	[0.0504]

Note: 4 main destinations, $N_{ij} > 100$, 676 obs.

Robust standard errors in brackets

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

Table 10: Product-level IME

	Coefficient from $\ln x_{ij}^p$ on $\ln X_{ij}$		
Panel a: 4 main destinations			
IM elasticity	0.369***	0.369***	0.391***
Standard error	[0.0156]	[0.0171]	[0.0157]
Observations	676	676	676
Panel b: all destinations			
IM elasticity	0.446***	0.341***	0.325***
Standard error	[0.00733]	[0.00768]	[0.00682]
Observations	7211	7211	7211
Year FE	Yes	Yes	Yes
Destination FE		Yes	Yes
Origin FE			Yes

Note: $N_{ij} > 100$

Robust standard errors in brackets

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

Table 11: Fixed trade costs distance elasticity and granularity

	Fixed trade costs elasticity	
	Firm level	Product level
ζ	-0.587***	-0.165***
Standard error	[0.0318]	[0.0333]
Observations	3912	3912

Note: $N_{ij} > 100$

Robust standard errors in brackets

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

Table 12: IME under granularity

	$corr(\alpha_j \varphi, \alpha_k \varphi)$	
	0	1
$\widetilde{\theta} = 2.3$	0.004	0.002
$\widetilde{\theta} = 1.75$	0.020	0.005
$\widetilde{\theta} = 1.25$	0.146	0.036
$\widetilde{\theta} = 1$	0.368	0.103

Note: $N_{ij} > 100$

N_{ij} data as of 2009, 74 obs.

Table 13: Number of firms and population

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

Robust standard errors in brackets, 468 obs.

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

Table 14: QQ estimates of $\bar{\sigma}_\varphi$

	All firms	Top 50%	Top 25%
$\bar{\sigma}_\varphi$	7.884***	5.456***	4.545***
	[0.00332]	[0.00212]	[0.00260]
Observations	1,166,885	582,361	290,806
R^2	0.846	0.937	0.941
Bilateral FE	Yes	Yes	Yes

Robust standard errors in brackets

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

Table 15: IME under lognormal distribution

	IME
$\bar{\sigma}_\varphi = 4.55$	0.431
$\bar{\sigma}_\varphi = 2.4$	0.163

Table 16: Trade costs and distance, lognormal

	log fixed costs	log variable costs
$\ln dist$	0.379***	0.366***
Standard error	[0.0514]	[0.0179]

Note: 4 main destinations, $N_{ij} > 100$, 676 obs.

Robust standard errors in brackets

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

Table 17: Estimates of dispersion, full lognormal model

	Estimate	95% CI
$\bar{\sigma}_\varphi$	2.60	[2.45, 2.78]
σ_α	2.57	[2.46, 2.67]
σ_f	3.59	[3.36, 3.78]

Table 18: Estimates of IME, full lognormal model

	IME	95% CI
Unrestricted model	0.58	[0.49...0.65]
Setting $\bar{\mu}_{f,ij} = \delta_i^o + \delta_j^d$	0.63	[0.58...0.68]

Table 19: Implies trade costs in full lognormal model

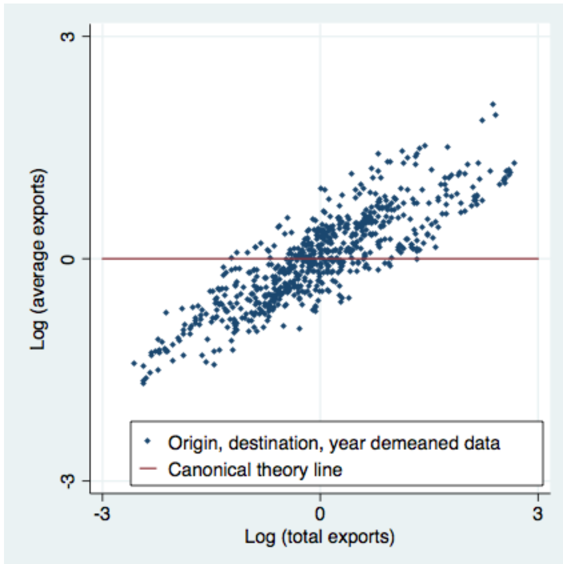
	Estimate	95% CI
Correlation	0.69	[0.53. . .0.80]
	Distance elasticity	
Fixed costs	1.93	[1.45. . .2.33]
Variable costs	0.53	[0.47. . .0.57]

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

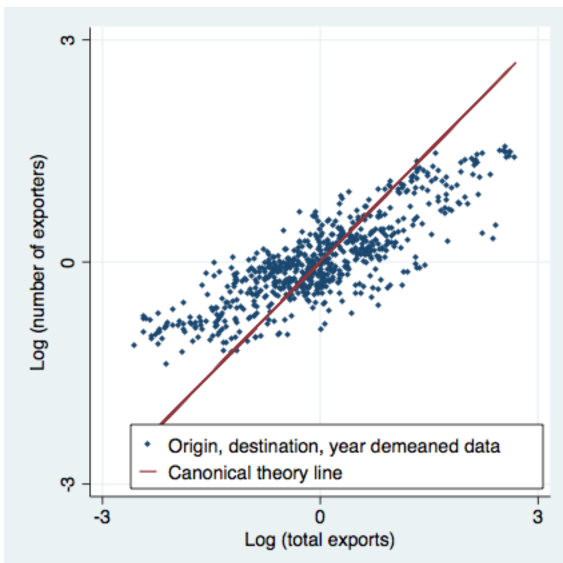
	$\bar{\theta}$	s. e.	Observations
Full sample	19.82***	[0.902]	39,054
Weights $\sqrt{N_{ij}}$	4.803***	[0.0431]	39,054
Dropping $N_{ij} < 100$	2.622***	[0.0185]	7,211
Dropping $M_{ij} < 100$	2.267***	[0.0140]	4,713

Figure 1: Intensive and Extensive margins of exporting

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



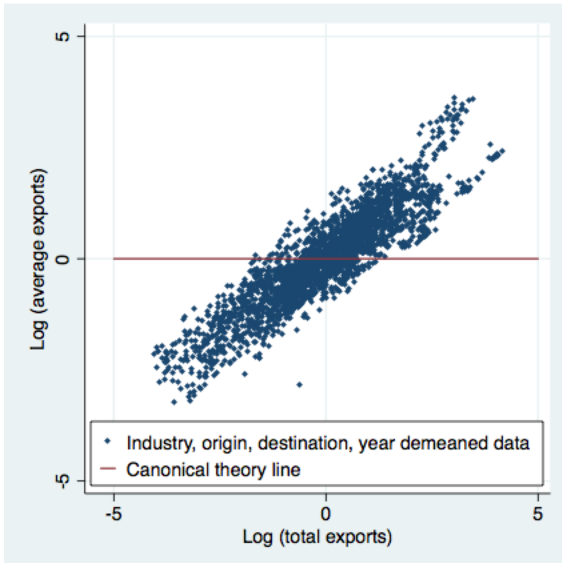
Panel b: Number of exporters (extensive margin) and total exports



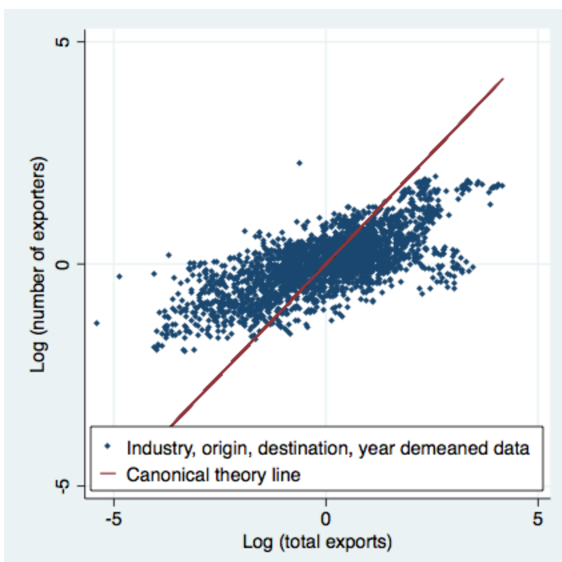
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log total exports at the exporting country-destination country-year level demeaned by country, destination, and year fixed effects. Only four destination countries are considered: France, Germany, Japan, and the US. 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



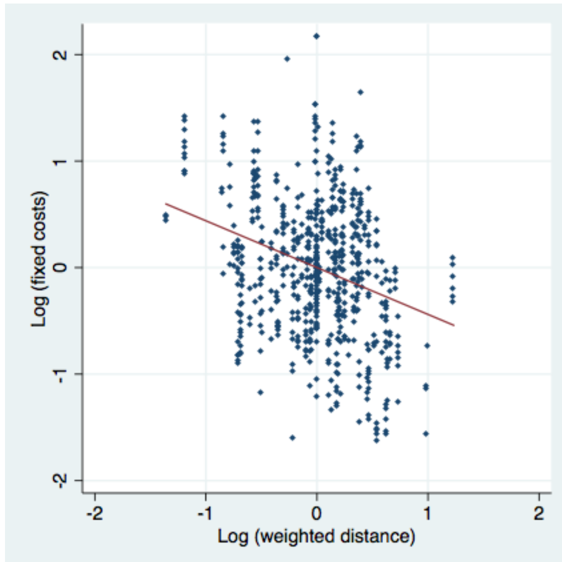
Panel b: Number of exporters (extensive margin) and total exports



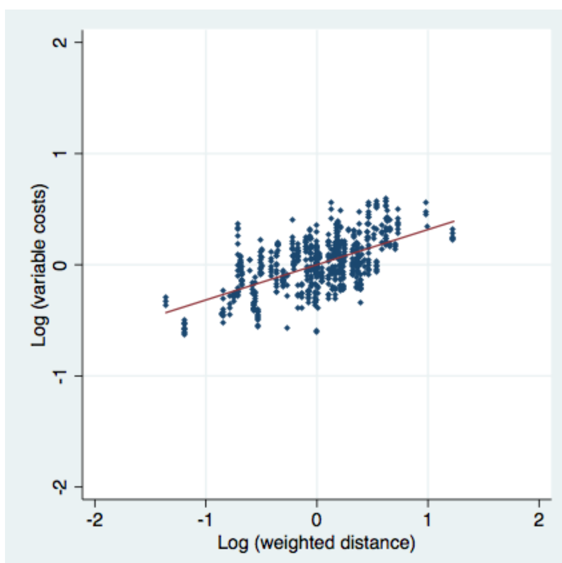
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log total exports at the exporting country-industry-destination country-year level demeaned by country, destination, industry, and year fixed effects. Only four destination countries are considered: France, Germany, Japan, and the US. The dots represent the raw measures. The line is the slope predicted by the Melitz-Pareto model.

Figure 3: Fixed and variable trade costs and distance

Panel a: fixed trade costs and distance

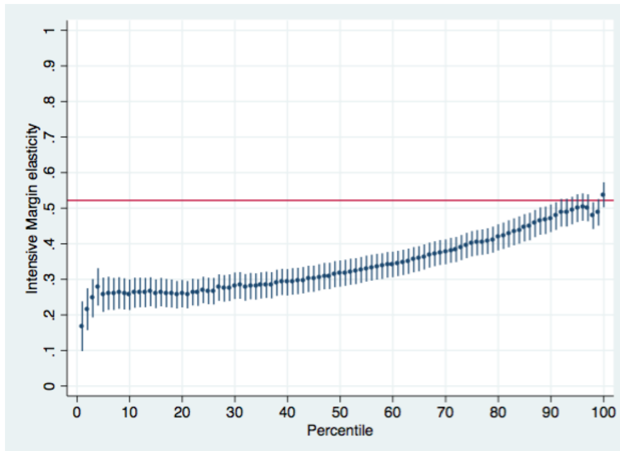


Panel b: variable trade costs and distance



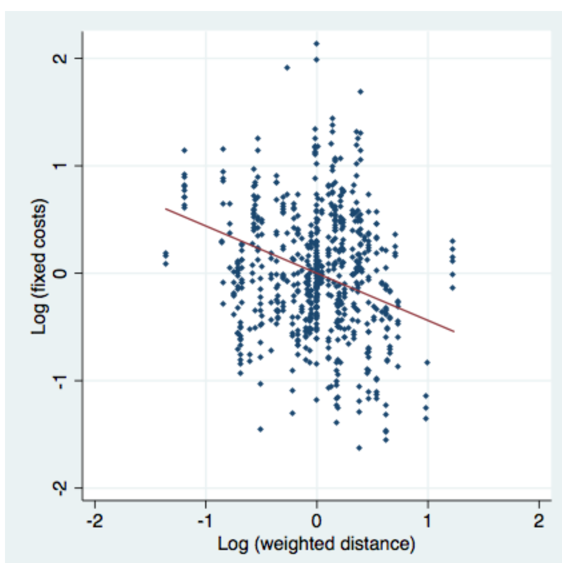
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance (demeaned by origin and destination) taken from Mayer and Zignago (2011). The y-axis represents fixed or variables trade costs demeaned by origin and destination. Only four destination countries are considered: France, Germany, Japan, and the US. To calculate the model-implied fixed and variable trade costs we use $\bar{\theta} = 1.25$ from Head and Mayer (2014) and $\sigma = 5$ from Bas et al. (2015)

Figure 4: IME for each percentile, data



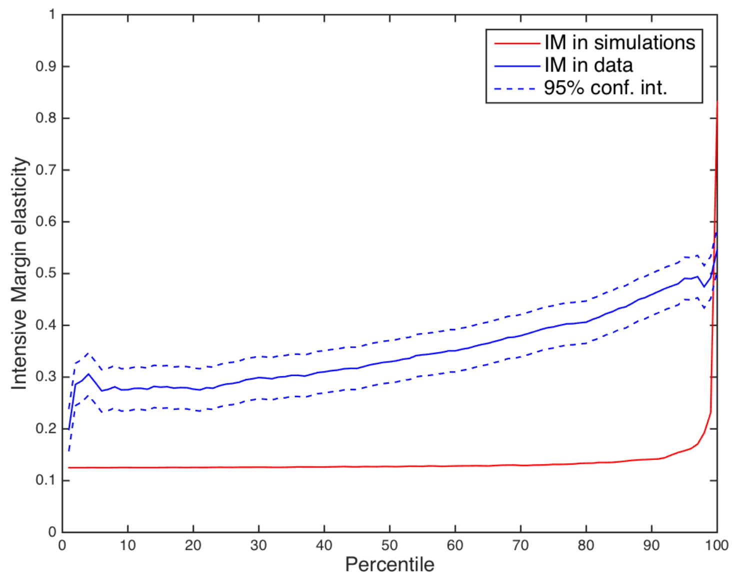
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represent percentiles. Each dot represents coefficient from the regression of log average exports in each percentile on log total exports. Data is demeaned by origin, destination, and year. Only four destination countries are considered: France, Germany, Japan, and the US.

Figure 5: Fixed product-level trade costs and distance



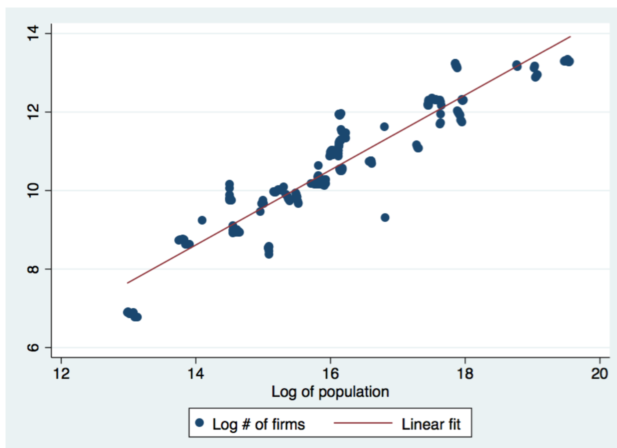
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011). Only four destination countries are considered: France, Germany, Japan, and the US.

Figure 6: IME for each percentile, Pareto and granularity



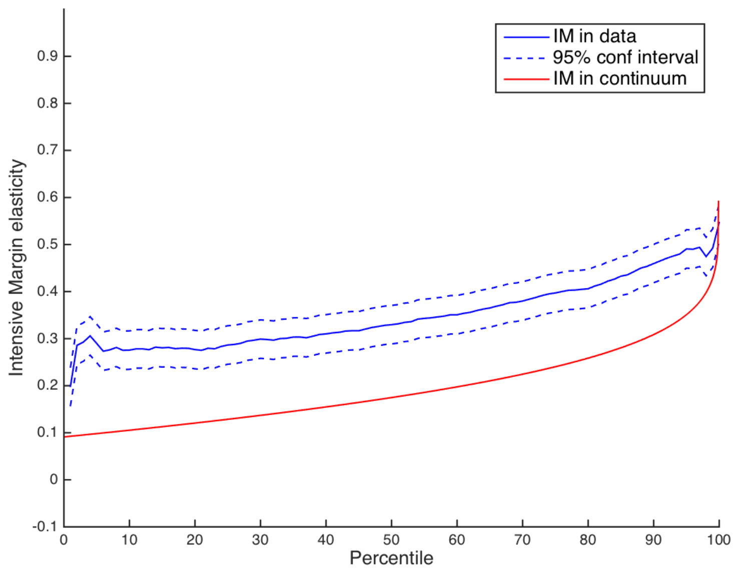
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and USA. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME 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 7: 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).

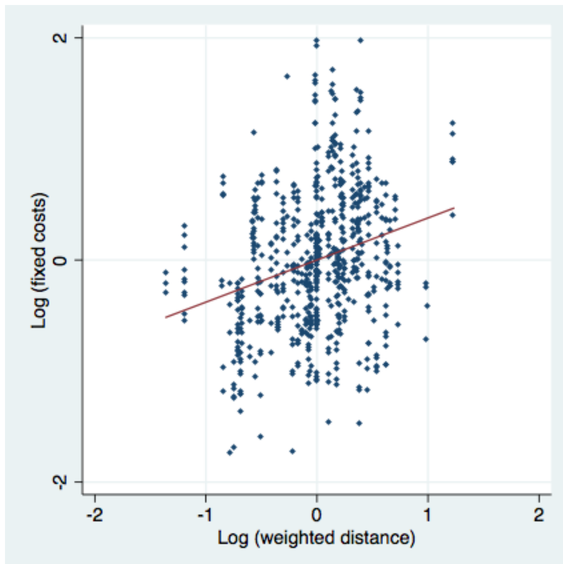
Figure 8: IME for each percentile, lognormal



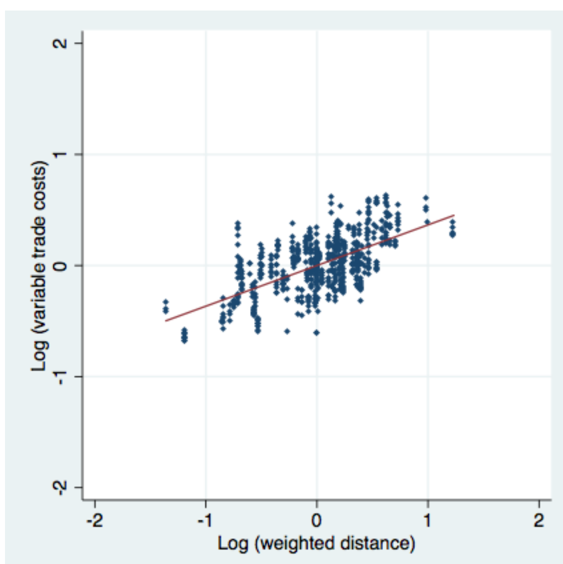
Note: the source is the exporter-level data used for the Exporter Dynamics Database. The darker solid line corresponds to IME for each percentile estimated using EDD and four main destinations: France, Germany, Japan and USA. Dashed lines indicate 95% confidence intervals. The lighter solid line is IME for each percentile implied by the model with lognormal distribution of productivity, $\bar{\sigma}_\varphi = 4.55$ (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 9: Fixed and variable trade costs and distance, lognormal

Panel a: fixed trade costs and distance

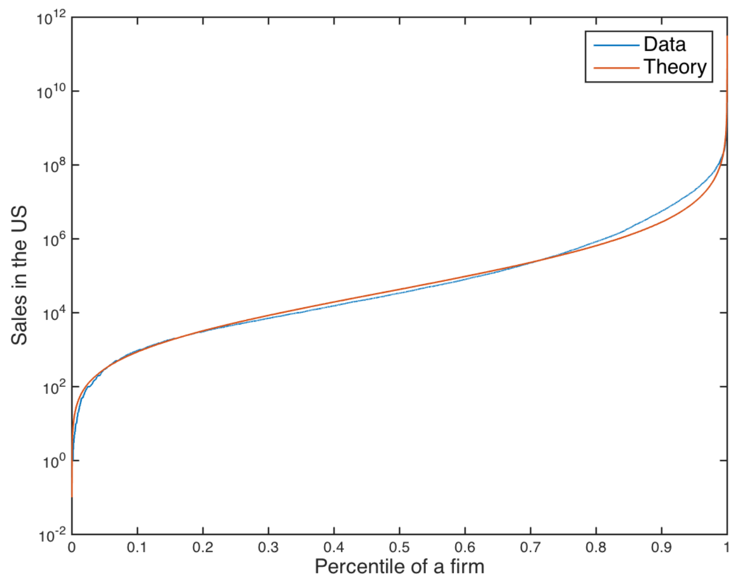


Panel b: variable trade costs and distance



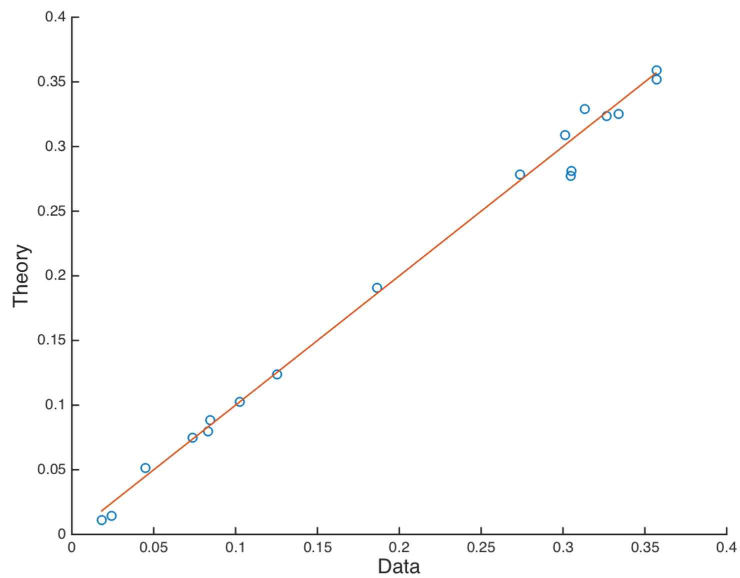
Note: source is the exporter-level data used for the Exporter Dynamics Database. The x-axis represents log distance taken from Mayer and Zignago (2011). Only four destination countries are considered: France, Germany, Japan, and the US. To calculate the model-implied fixed and variable trade costs we use our estimate of $\sigma_\varphi = 4.55$ and $\sigma = 5$ from Bas et al. (2015), and implied number of firm from Bento and Restuccia (2015)

Figure 10: Full lognormal model, inverse CDF of log sales



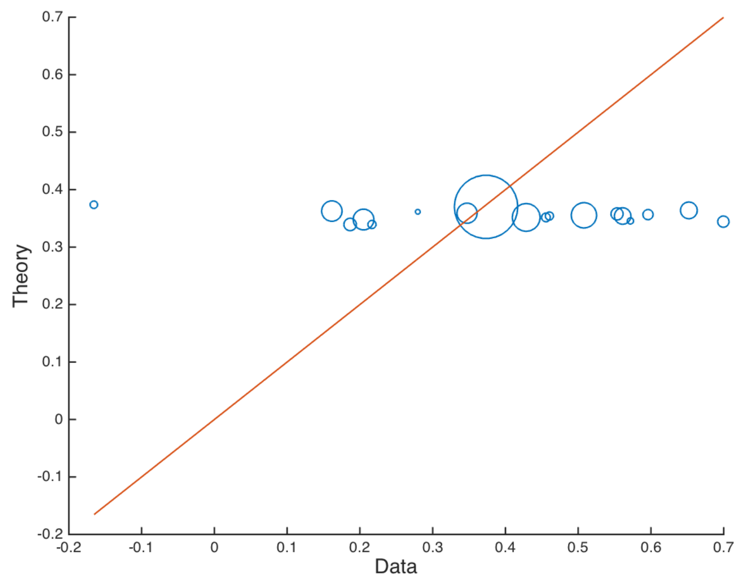
Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. The darker solid line corresponds to empirical CDF of log sales from some origin to the US. The lighter solid line corresponds to CDF implied by estimated full lognormal model

Figure 11: Share of firms selling only to the less popular market



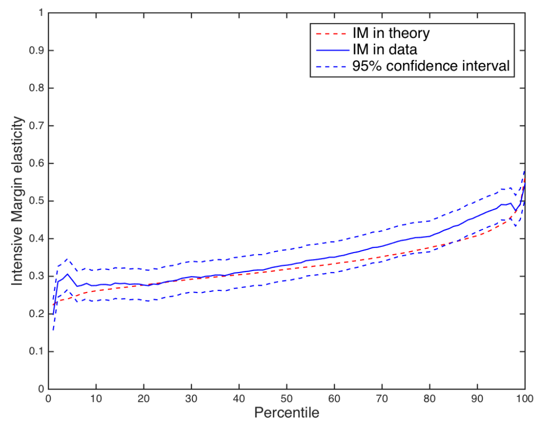
Note: the source is the exporter-level data used for the Exporter Dynamics Database and authors' calculations. The horizontal axis corresponds to the share of firms exporting to less popular market in the data, same shares implied by estimated model are plotted on the vertical axis. The line is a 45^o line

Figure 12: 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. The horizontal axis corresponds to correlation between log exports to the US and Germany (for those firms who sell in both markets) in the data, same shares implied by estimated model are plotted on the vertical axis. The size of the circles is proportional to the number of exporters

Figure 13: 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. Dark dashed lines represent 95% confidence intervals. The light dashed line represents the regression of log average exports in each percentile on log total exports in the simulated full lognormal model. We use all years and 4 main destinations to calculate IME for each percentile in the data, and estimated full lognormal model using 2009 data for 18 origins and 2 destinations: USA and Germany.

Figure A1: Exports to largest destination and market entry