Online Appendix to

"The Intensive Margin in Trade: How Big and How Important?"

Ana M. Fernandes  Peter J. Klenow  Sergii Meleshchuk
World Bank  Stanford and NBER  UC Berkeley

Martha Denisse Pierola  Andrés Rodríguez-Clare *
Inter-American Development Bank  UC Berkeley and NBER

October 2, 2019

*We are grateful to Arnaud Costinot, Caroline Freund, Cecile Gaubert, Keith Head, Sam Kortum, Thierry Mayer, Eduardo Morales and Jesse Perla for useful discussions, to seminar participants at various institutions, and to Matthias Hoelzel and Nick Sander for outstanding research assistance. 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, as well as the Stanford Institute for Economic Policy Research. 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. The views expressed herein are those of the authors and should not be attributed to the IMF, its Executive Board, or its management. The findings expressed in this paper are those of the authors and do not necessarily represent the views of the Inter-American Development Bank or its member countries.
A Estimation of $\tilde{\theta}$

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

$$x_{il|j} = \left( \frac{N_{ij}}{N_{il}} \right)^{-1/\tilde{\theta}}$$  \hspace{1cm} (1)

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 $l^* = \arg \max_k N_{ik}$, that is, the largest destination market for each origin country $i$ (e.g., the United States for Mexico). Letting

$$z_{ij} = \frac{x_{il^*(i)|j}}{x_{il^*(i)|l^*(i)}}$$  \hspace{1cm} (2)

and

$$m_{ij} = \frac{N_{ij}}{N_{il^*(i)}}$$  \hspace{1cm} (3)

then we have

$$\ln z_{ij} = -\frac{1}{\tilde{\theta}} \ln m_{ij}.$$  \hspace{1cm} (4)

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

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 $\tilde{\theta} = 1.75$. In their case, they keep in their estimating sample only firms with positive sales in France, so the variables $x_{FF|j}$ and $N_{Fj}$ are calculated based on the same set of firms. To implement an approach comparable to theirs, we drop all firms from country $i$ that do not sell to $l^*(i)$, so the sample includes only $N_{il^*(i)}$ firms for country $i$. This implies that all firms that make up $N_{ij}$ are also selling to $l^*(i)$. Figure 5 in the paper reproduces Figure 3 from Eaton et al. (2011) by plotting the variables in equation (4). The slope in the graph is equal to $1/\tilde{\theta}$, and the corresponding estimated values are reported in Table I8. Based on all observations in the core sample of countries and using no weighting, the estimated $\tilde{\theta}$ is over 19. But in Figure 5 for small values of $m_{ij}$, which correspond to small values of $N_{ij}$, there is a lot of dispersion in $z_{ij}$. To minimize the effect of that noise we weight observations by $\sqrt{N_{ij}}$ and this lowers the estimate of $\tilde{\theta}$ to 4.8. Finally, when we drop all observations with $N_{ij} < 100$ (remember that here $N_{ij}$ measures the number of firms from country $i$ that sell to country $j$ and also to $l^*(i)$) we obtain $\tilde{\theta} = 2.4$, which is still higher than in Eaton et al. (2011). We will use this estimate in our simulations of the intensive margin elasticity.
B Granularity

Theory

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

\[ \ln N_{ij} = \mu_i^{N,o} + \mu_j^{N,d} - \theta \ln \bar{\tau}_{ij} - \bar{\theta} \ln \bar{F}_{ij} + \xi_{ij} \] (5)

and

\[ \ln x_{ij} = \mu_i^{x,o} + \mu_j^{x,d} + \ln \bar{F}_{ij} + \varepsilon_{ij}, \] (6)

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

\[ \text{IME} = \frac{- (\bar{\theta} - 1) \var{\ln \bar{F}_{ij}} - \theta \cov{\ln \bar{\tau}_{ij}, \ln \bar{F}_{ij}} + \var{\varepsilon_{ij}} + \cov{\varepsilon_{ij}, \xi_{ij}}}{\var{-\theta \ln \bar{\tau}_{ij} - (\bar{\theta} - 1) \ln \bar{F}_{ij} + \varepsilon_{ij} + \xi_{ij}}}, \] (7)

where \( \cov{\equiv \cov{\ln \bar{F}_{ij} + \varepsilon_{ij}, \xi_{ij}} + \cov{\ln \bar{F}_{ij} + \ln \bar{\tau}_{ij}, \varepsilon_{ij}}. \) If \( \var{\varepsilon_{ij}} \) is large enough then we could have \( \text{IME} > 0 \) even with \( \cov{\ln \bar{F}_{ij}, \ln \bar{\tau}_{ij}} > 0 \). 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 \( \kappa \equiv \frac{\sigma \theta}{\bar{\theta} - 1} \). 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 \( \varepsilon_{ij} \) is an error term that arises from random realizations of productivity draws, the first moment of which is independent of any variables that determine bilateral fixed trade costs. If we further assume that
\[ F_{ij} = F_i^o F_j^d e^{\zeta \ln \text{dist}_{ij}} + v_{ij}/\kappa, \]
where \( v_{ij} \) satisfies \( \mathbb{E}(v_{ij}|\text{dist}_{ij}) = 0 \), we can then write
\[ x_{ij} = \kappa F_i^o F_j^d e^{\zeta \ln \text{dist}_{ij}} + u_{ij}, \] (8)

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

To take advantage of the time dimension of our data, we extend (8) to allow for an origin-time and destination-time specific components in the expression of fixed trade costs,
\[ x_{ijt} = \kappa F_i^{o_t} F_j^{d_t} e^{\zeta_t \ln \text{dist}_{ij}} + u_{ijt}, \] (9)

where again \( \mathbb{E}(u_{ijt}|\text{dist}_{ij}) = 0 \). We estimate (9) using Poisson pseudo maximum likelihood method as in Silva and Tenreyro (2011).

**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 \( \varphi \) and destination-specific demand shock \( \alpha_j \) can be calculated as
\[ x_{ij}(\varphi, \alpha_j) = \sigma F_{ij} \left( \frac{\alpha_j \varphi}{\alpha^*_i \varphi^*_j} \right)^{\sigma-1}, \] (10)

where \( \alpha^*_i \varphi^*_j \) 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 \( \varphi \) and \( \alpha_j \) 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 $\text{var} \left( F_{ij} \right) = 0$, 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 (10) to simulate the exports for each firm and to calculate average exports per firm (in total and in each percentile) for each origin-destination pair.

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

**Data**

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

We use equation (9) to estimate firm-level as well as firm-product-level fixed trade cost elasticities with respect to distance ($\zeta$). Table 8 shows that both of these elasticities are negative and statistically significant, so both firm-level and firm-product-level model-implied fixed trade costs are decreasing with distance, although with a much smaller elasticity than when not accounting for granularity (compare results of Tables 7 and 8).

Table 9 reports the estimated IME using simulated data for alternative values of $\bar{\theta}$ and for either zero or perfect correlation between the product of demand and productivity shocks across destinations. We consider three values of $\bar{\theta}$: our estimate $\bar{\theta} = 2.4$, the value $\bar{\theta} = 1.25$ that can be inferred from standard estimates of $\theta$ and $\sigma$ 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), and $\bar{\theta} = 1$ (as in Zipf’s Law).

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

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

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

C QQ-Estimation of $\sigma_\varphi$

Exports from country $i$ to country $j$ of a firm with productivity $\varphi$ 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 $\ln x_{ij}(\varphi) \sim N_{\text{trunc}}(\mu_{\varphi,ij}, \sigma_\varphi; \ln (\sigma F_{ij}))$, where $\sigma_\varphi = \sigma_\varphi (\sigma - 1), \mu_{\varphi,ij} = \mu_{\varphi,i} (\sigma - 1) + \ln (\sigma F_{ij}) + (1 - \sigma) \ln \left(\varphi^*_{ij}\right)$, 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 $\ln x_{ij,n}$, 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 $\bar{\mu}_{\varphi,ij} + \sigma_\varphi \Phi^{-1}\left(\hat{\Phi}_{ij,n}\right)$, 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 $\beta$ obtained from the regression

$$\ln x_{ij,n} = \alpha_{ij} + \beta \Phi^{-1}\left(\hat{\Phi}_{ij,n}\right) + \varepsilon_{ij,n}. \quad \text{(11)}$$

Table I9 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 I9. This comes from a not very surprising empirical failure of the simple Melitz-lognormal model outlined in the first part of Section 3:
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 $\sigma F_{ij}$), 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.

### D Quasi-Bayesian Estimation for the full Melitz-lognormal model

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

$$f_{X_{i1}, \ldots, X_{ij}}(x_{i1}, \ldots, x_{ij}) = \int_\omega f_{X_{i1}, \ldots, X_{ij}|\tilde{\varphi}_i}(x_{i1}, \ldots, x_{ij}|\omega) f_{\tilde{\varphi}_i}(\omega) d\omega$$  \hspace{1cm} \text{(12)}$$

$$= \int_\omega \prod_j f_{X_{ij}|\tilde{\varphi}_i}(x_{ij}|\omega) f_{\tilde{\varphi}_i}(\omega) d\omega,$$  \hspace{1cm} \text{(13)}$$

where the second equality comes from the fact that, conditional on productivity, sales are independent across markets (as well as the probability of selling to those markets). We now need to characterize $f_{X_{ij}|\tilde{\varphi}_i}(f_{ij}|\omega)$ to calculate the likelihood function. In general we have
\[ f_{X_{ij} | \bar{\phi}_i}(x_{ij} | \omega) = \left[ f_{Z_{ij} | \bar{\phi}_i}(x_{ij} | \omega) \Pr\{ Z_{ij} \geq \ln \sigma + \ln f_{ij} | \bar{\phi}_i = \omega, Z_{ij} = x_{ij} \} \right]^{|(x_{ij} \neq \emptyset)} \times \\
\times \left[ \Pr\{ \ln \sigma + \ln f_{ij} \geq Z_{ij} | \bar{\phi}_i = \omega \} \right]^{|(x_{ij} = \emptyset)}. \] (14)

The term on the first line of 14 corresponds to the density function for the cases when we observe exports, while the second line corresponds to the mass at the point \( x_{ij} = \emptyset \).

For the case when sales are not zero \( X_{ij} = Z_{ij} \) and

\[ Z_{ij} | [\bar{\phi}_i = \omega] = \omega + d_{ij} + \ln \alpha - \mu_\alpha, \] (15)

\[ Z_{ij} | [\bar{\phi}_i = \omega] \sim N(d_{ij} + \omega, \sigma_{\alpha,i}^2). \] (16)

In addition

\[ \Pr[Z_{ij} \geq \ln \sigma + \ln f_{ij} | \bar{\phi}_i = \omega, Z_{ij} = x_{ij}] = \Pr[\ln \sigma + \ln f_{ij} \leq x_{ij} | \ln \alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega], \] (17)

\[ \ln \sigma + \ln f_{ij} | [\ln \alpha - \mu_\alpha = x_{ij} - d_{ij} - \omega] \sim N(\mu_1, \sigma_{1,i}^2), \]

where

\[ \mu_1 \equiv \bar{\mu}_{f,ij} + \frac{\sigma_{\alpha,f,i}^2}{\sigma_{\alpha,i}^2} (x_{ij} - d_{ij} - \omega), \]

\[ \sigma_{1,i}^2 \equiv \sigma_{f,i}^2 (1 - \rho_i^2). \]

Finally we have

\[ \Pr[Z_{ij} \leq \ln \sigma + \ln f_{ij} | \bar{\phi}_i = \omega] = \Pr[-\ln \sigma - \ln f_{ij} + (\ln \alpha - \mu_\alpha) + d_{ij} \leq -\omega], \] (18)

\[ -\ln \sigma - \ln f_{ij} + (\ln \alpha - \mu_\alpha) + d_{ij} \sim N(-\bar{\mu}_{f,ij} + d_{ij}, \sigma_{2,i}^2), \] (19)

where

\[ \sigma_{2,i}^2 \equiv \sigma_{f,i}^2 + \sigma_{\alpha,i}^2 - 2\sigma_{\alpha f,i}. \]

Let \( \phi \) and \( \Phi \) denote the PDF and CDF of the standard normal, respectively. Plugging functional forms into 14 we can get the object of interest.
\[
X_{i1, \ldots, iJ}(x_{i1}, \ldots, x_{iJ}) = \int_{\omega} \prod_{j} \left\{ \frac{1}{\sigma_{\alpha_{ij}}} \phi \left( \frac{x_{ij} - d_{ij} - \omega}{\sigma_{\alpha_{ij}}} \right) \Phi \left( \frac{x_{ij} - \left[ \mu_{f,ij} + \frac{\sigma_{\alpha_{ij}}}{\sigma_{\alpha_{ij}}} (x_{ij} - d_{ij} - \omega) \right]}{\sqrt{\sigma_{f,i}^2 (1 - \rho_i^2)}} \right) \right\}^{I(x_{ij} \neq \emptyset)} \times \times \left\{ \Phi \left( \frac{-\omega + \mu_{f,ij} - d_{ij}}{\sqrt{\sigma_{f,i}^2 + 2 \sigma_{\alpha_{ij}} \sigma_{f,i}}} \right) \right\}^{I(x_{ij} = \emptyset)} \frac{1}{\sigma_{\bar{\varphi}_{ij}}} \phi \left( \frac{\omega}{\sigma_{\bar{\varphi}_{ij}}} \right) d\omega \quad (20)
\]

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

\[
g_{X_{i1, \ldots, iJ}}(x_{i1}, \ldots, x_{iJ}) = f_{X_{i1, \ldots, iJ} \cap \text{Is an exporter}}(x_{i1}, \ldots, x_{iJ} \cap \text{Is an exporter}), \quad (21)
\]

and hence

\[
g_{X_{i1, \ldots, iJ}}(x_{i1}, \ldots, x_{iJ}) = \frac{f_{X_{i1, \ldots, iJ}}(x_{i1}, \ldots, x_{iJ})}{\text{Pr}_i[\text{observe sales to at least 1 destination}]} = \frac{f_{X_{i1, \ldots, iJ}}(x_{i1}, \ldots, x_{iJ})}{1 - \text{Pr}_i[\text{observe sales to no destinations}]} = \frac{f_{X_{i1, \ldots, iJ}}(x_{i1}, \ldots, x_{iJ})}{1 - \int_{\omega} \prod_{j} \Phi \left( \frac{-\omega + \mu_{f,ij} - d_{ij}}{\sqrt{\sigma_{f,i}^2 + 2 \sigma_{\alpha_{ij}} \sigma_{f,i}}} \right) \frac{1}{\sigma_{\bar{\varphi}_{ij}}} \phi \left( \frac{\omega}{\sigma_{\bar{\varphi}_{ij}}} \right) d\omega}. \quad (22)
\]

Let \( C_i \) denote the probability that a firm from origin \( i \) sells to at least one of the destinations we consider.\(^1\) We know the number of firms, \( N_i^c \), that sell to those destinations, and we can thus infer the number of draws \( N_i = N_i^c / C_i \).

The likelihood function is a product of density functions as in \( 22 \). Parameters to estimate are

\[
\Theta_i = \left\{ \{d_{ij}, \mu_{f,ij}\}_{i,j}, \sigma_{\varphi_{ij}}, \sigma_{\alpha_{ij}}, \sigma_{f,i}, \rho_i \right\}.
\]

We next compute the density in the numerator of Equation \( 22 \). We can write this density in

\(^1\)\( C_i \) is given by the denominator in \( 22 \).
the following general form:

\[
f_{X_{i1},\ldots,X_{ij}}(x_{i1}, \ldots, x_{ij}) = \int_{\omega} G(\omega) \phi \left( \frac{\omega_{\phi,i}}{\sigma_{\phi,i}} \right) \exp \left( -\left( \frac{\omega_{\phi,i}}{\sqrt{2} \sigma_{\phi,i}} \right)^2 \right) d\omega,
\]

where \( G(\omega) \) is a known function of \( \omega \). Using change of variables \( \tilde{\omega} = \frac{\omega}{\sqrt{2} \sigma_{\phi,i}} \) and \( d\omega = \sqrt{2} \sigma_{\phi,i} d\tilde{\omega} \) we can write:

\[
f_{X_{i1},\ldots,X_{ij}}(x_{i1}, \ldots, x_{ij}) = \int_{\tilde{\omega}} G(\sqrt{2} \sigma_{\phi,i} \tilde{\omega}) \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2} \left( \frac{\tilde{\omega}}{\sigma_{\phi,i}} \right)^2 \right) d\tilde{\omega}.
\]

We can speed up the process to calculate object in 24 by applying a Gauss-Hermite quadrature. In general:

\[
\int g(x) \exp(-x^2) dx \approx \sum_k g(r_k) w_k,
\]

where \( r_k \) are the roots of the Hermite polynomial and \( \omega_k \) are associated weights. We calculate 33 values of \( r_k \) and corresponding weights \( w_k \) using the Gauss-Hermite method. The number of points was chosen to ensure that the quadrature approximation is accurate and that calculations take a reasonable amount of time.

### E MCMC algorithm

Since the likelihood function in the equation 33 in the paper is highly nonlinear there may exist multiple local maxima. We thus estimate a vector of parameters \( \Theta \) (defined in equation 34 in the paper) using the methodology developed in Chernozhukov and Hong (2003). This procedure not only gives us point estimates, but also yields confidence intervals for the estimated parameters, intensive marginal elasticity implied by the model, and elasticity of trade costs with respect to distance. We implement Chernozhukov and Hong (2003) procedure using the Metropolis-Hastings Monte-Carlo Markov chain algorithm. This algorithm yields a chain of parameter draws \( \{\Theta_i^{(n)}\}_{n=1}^N \) for each origin \( i \) such that \( \tilde{\Theta}_i \equiv \frac{1}{N} \sum_n \Theta_i^{(n)} \) is a consistent estimate of \( \Theta_i \). Moreover, using the values of the parameters in the chain we can construct confidence intervals for some functions \( f(\Theta_i) \). The chain of parameters \( \{\Theta_i^{(n)}\}_{n=1}^N \) for each origin is constructed in the following way:
Step 1. Randomly choose a starting guess $\Theta_i^{(0)}$.

Step 2. Draw a candidate vector of parameters for the chain’s $n+1$ value as $\tilde{\Theta}_i^{n+1} = \Theta_i^{(n)} + \psi^{(n)}$, where $\psi^{(n)}$ is a vector of iid shocks taken from the mean-zero normal distribution. The variance-covariance matrix of this distribution is diagonal. The initial values of the diagonal elements are set at $0.2\Theta_i^{(0)}$. At each step all but 1 elements of $\psi^{(n)}$ are zero. In other words, we only add an iid shock to one parameter at each step of the chain. Since the vector $\Theta_i$ has 34 elements for each $i$, we try a new value for each parameter every 34 steps.

Step 3. Calculate $\Theta_i^{(n+1)}$ in the following way:

$$\Theta_i^{(n+1)} = \begin{cases} 
\tilde{\Theta}_i^{(n+1)} & \text{with probability } \min\left[1; L\left(\tilde{\Theta}_i^{(n+1)}\right) - L\left(\Theta_i^{(n)}\right)\right] \\
\Theta_i^{(n)} & \text{otherwise,}
\end{cases}$$

where $L(\theta)$ is defined in equation 33 in the paper. Every 3,400 iterations (100 iterations per parameter) during the first 100,000 iterations we update diagonal elements of the variance-covariance matrix of the shocks so that the acceptance rate for each parameter is in the interval 0.25–0.35, as recommended by Chernozhukov and Hong (2003). We calculate the acceptance rate as a share of draws for which $\Theta_i^{(n+1)} = \tilde{\Theta}_i^{(n+1)}$.

We repeat the procedure until we have at least 34 million draws (1 million draws per parameter) in the chain after we discard the first 100,000 draws (known as “burn-in period”). For each origin we construct 5 different chains with different starting guesses to check that our estimates are robust with respect to the starting values (we discuss convergence of the chains in the Online Appendix F).

Having estimated the chains, we take 1,000 random draws from the chains for each origin with replacement. We use those draws to calculate point estimates (averages) as well as 95% confidence intervals. Using those draws we simulate the model 1,000 times and run IME regressions the way we run them in the data. Finally, we run 1,000 regressions of the estimates $d_{ij}$ and $\mu_{f,ij}$ on log distance with origin and destination fixed effects for the 4 destinations (USA, France, Germany, and Japan) and interpret the results as elasticity of variable and fixed trade costs with respect to distance.
F Convergence of the Monte Carlo Markov chains

As mentioned before, we ran our estimation algorithm 5 times per origin country which gave us 5 different chains that started at different random initial guesses. This lets us compare underlying distributions of parameter estimates across different chains. We checked the convergence of the chains using the following criteria:

1) comparing the means of parameter estimates in the first and second half of the chain;

2) comparing parameter means across different chains for the same origin.

If the means calculated for the two parts of a chain are statistically indistinguishable, we conclude that the chain converged. If the means across different chains are statistically indistinguishable, we conclude that multiple chains converged to the same region. It turns out that in some cases the chains didn’t converge. In those cases we disregarded the chains that didn’t converge. Out of remaining chains, we randomly choose one chain per origin to conduct our numerical analysis. Figures J1-J34 plot the means for the first and second half of each chain, as well as an overall mean with the range between 2.5th and 97.5th percentiles of the draws in those chains that converged. As can be seen, means and their confidence intervals are virtually indistinguishable across all the chains and subchains that converged for all countries and all parameters.

G Local trade elasticity in the full Melitz-lognormal model

Let $\bar{\varphi}$ be a combination of demand and productivity shocks (not demeaned and not in logs). Total exports from $i$ to $j$ is given by

$$X_{ij} = N_i (\bar{\sigma} w_i \tau_{ij})^{1-\sigma} P_j^{\sigma-1} X_j \int_0^\infty \int_0^\infty \varphi^{\sigma-1} g(\varphi | f) d\varphi g(f) df,$$

where the cutoff productivity for any $f$ is given by

$$\varphi_{0,ij}(\tau_{ij} | f) = \bar{\sigma} w_i \tau_{ij} \left[ \frac{\sigma w_i f}{P_j^{\sigma-1} X_j} \right]^{\frac{1}{\sigma-1}}.$$

\(^2\)It happens, for example, when some of the values for $\sigma_{\varphi,i}$ and $\sigma_{\alpha,i}$ exploded.
Assume that changes in $\tau_{ij}$ have no effect on $P_j$. Then trade elasticity is equal to

$$-\theta_{ij} = (1 - \sigma) + \frac{\partial \ln \int_0^\infty \int_0^\infty \varphi_0,ij(\tau_{ij}|f) \varphi^{\sigma-1} g(\varphi|f) df g(f) df}{\partial \ln \tau_{ij}}.$$  \hfill (28)

The first term is the intensive margin and the second term is the extensive margin (EM) of trade elasticity. Letting

$$\Gamma_{ij} = \int_0^\infty \int_0^\infty \varphi_0,ij(\tau_{ij}|f) \varphi^{\sigma-1} g(\varphi|f) df g(f) df,$$  \hfill (29)

then

$$EM = \int_0^\infty \frac{\partial \int_0^\infty \varphi_0,ij(\tau_{ij}|f) \varphi^{\sigma-1} g(\varphi|f) df}{\partial \tau_{ij}} \frac{g(f) df}{\Gamma_{ij}} \frac{\tau_{ij}}{\Gamma_{ij}},$$  \hfill (30)

$$= -\int_0^\infty \varphi_0,ij(\tau_{ij}|f) \varphi^{\sigma} g(\varphi_0,ij(\tau_{ij}|f)|f) g(f) df \frac{1}{\Gamma_{ij}},$$  \hfill (31)

where we used

$$\frac{\partial \varphi_0,ij(\tau_{ij}|f)}{\partial \tau_{ij}} = \frac{\varphi_0,ij(\tau_{ij}|f)}{\tau_{ij}}.$$  \hfill (32)

Now we can use the expression for the extensive margin of trade elasticity to calculate local trade elasticity implied by the full Melitz-lognormal model.

## H Relation to Melitz and Redding (2015)

In this section we explore how the implications of our model relate to those in Melitz and Redding (2015). We start with a symmetric, two-country economy, with heterogeneous firms as in our Melitz-lognormal model. Following Melitz and Redding (2015), we set the domestic trade share to be 0.9, the share of exporters to be 0.18, and international trade costs to be $\tau = 1.83$. We set the parameters of the productivity and fixed cost shocks at their median values (across origin countries) from our estimation. We then conduct a trade liberalization experiment and compare welfare responses in the true model and the Melitz-Pareto models in the following way:

1. Increase trade costs to $\tau_{big}$ and calculate the local trade elasticity $\theta^*$ implied by the true lognormal model as outlined in Online Appendix Section G. This is similar to the “starting” trade elasticity in Melitz and Redding (2015).

2. Reduce trade costs from $\tau_{big}$ to 1.25 and calculate the following welfare responses:
(a) Welfare gains in the true Melitz-lognormal model.

(b) Welfare gains in the Melitz-Pareto model using the ACR formula with $\theta^*$ and changes in trade shares implied by the true model. Following ACR, we refer to this as an ex-post welfare analysis for the Melitz-Pareto model.

(c) Welfare gains in the Melitz-Pareto model using the ACR formula with $\theta^*$ and changes in trade shares implied by the exact hat algebra in the Melitz-Pareto model. Following ACR, we refer to this as an ex-ante welfare analysis for the Melitz-Pareto model.

The results are reported on Figure J35. In contrast to Melitz and Redding (2015), we do not find large differences between the true model and the Pareto model. As the figure shows, the trade elasticity is below 4.3 even for high starting trade costs. Since the trade elasticity cannot fall below $\sigma - 1 = 4$, the starting trade elasticity $\theta^*$ is close to the average trade elasticity along the path of liberalization, and thus the differences between the welfare responses in the two models are not quantitatively large, both for the ex-post and the ex-ante analysis.

To highlight the importance of the variation in the trade elasticity for these results, we redo the exercise above but reducing the standard deviation of the sum of productivity and demand shocks from 4.15 used above to 1.4, which is far below the minimum standard deviation we find in our estimation. As can be seen in Figure J36, the trade elasticity becomes much more variable, and the differences in the welfare response become more pronounced. This establishes that our results are driven not by the nature of our exercise or the use of a lognormal distribution but by the discipline imposed by the estimation using the EDD data.

We thus conclude that while the full lognormal Melitz model in this paper can theoretically yield similar results to those in Table 1 in Melitz and Redding (2015), that does not happen when the model is estimated on the EDD data.
References


## I Appendix Tables

Table I1: Additional EDD countries and years in the Extended Sample (only EDD statistics available)

<table>
<thead>
<tr>
<th>ISO3</th>
<th>Country name</th>
<th>1st year</th>
<th>Last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRA</td>
<td>Brazil</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>DNK</td>
<td>Denmark</td>
<td>2003</td>
<td>2012</td>
</tr>
<tr>
<td>ESP</td>
<td>Spain</td>
<td>2005</td>
<td>2013</td>
</tr>
<tr>
<td>EST</td>
<td>Estonia</td>
<td>2003</td>
<td>2011</td>
</tr>
<tr>
<td>KWT</td>
<td>Kuwait</td>
<td>2009</td>
<td>2010</td>
</tr>
<tr>
<td>LKA</td>
<td>Sri Lanka</td>
<td>2013</td>
<td>2013</td>
</tr>
<tr>
<td>NOR</td>
<td>Norway</td>
<td>2003</td>
<td>2013</td>
</tr>
<tr>
<td>PRT</td>
<td>Portugal</td>
<td>2003</td>
<td>2012</td>
</tr>
<tr>
<td>SWZ</td>
<td>Swaziland</td>
<td>2012</td>
<td>2012</td>
</tr>
<tr>
<td>TUR</td>
<td>Turkey</td>
<td>2003</td>
<td>2013</td>
</tr>
</tbody>
</table>
Table I2: IME regressions, extended sample

| Panel a: country pairs with $N_{ij} \geq 100$ |  |
| IM elasticity | 0.437*** | 0.450*** | 0.381*** |
| Standard error | [0.0037] | [0.0028] | [0.0034] |
| $R^2$ | 0.55 | 0.75 | 0.83 |
| Variation in $\ln X_{ij}$ explained by FE, % | 0.01 | 0.14 | 0.55 |
| Observations | 14,318 | 14,300 | 13,964 |

| Panel b: all country pairs |  |
| IM elasticity | 0.434*** | 0.477*** | 0.516*** |
| Standard error | [0.0016] | [0.0016] | [0.0016] |
| $R^2$ | 0.69 | 0.77 | 0.80 |
| Variation in $\ln X_{ij}$ explained by FE, % | 0.00 | 0.22 | 0.56 |
| Observations | 52,775 | 52,775 | 52,658 |

Year FE | Yes |
Origin $\times$ year FE | Yes | Yes |
Destination $\times$ year FE | Yes |

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year fixed effects (column 1), origin $\times$ year fixed effects (column 2), origin $\times$ year and destination $\times$ year fixed effects (column 3) using the extended sample of the EDD. The data are aggregated at the year-origin-destination level for a set of origin-years listed in Table 1 in the paper and Table I1 in the Online Appendix. Panel a) presents the regression on the sample of country-pairs with at least 100 exporters. Panel b) presents the regression on the full sample. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
Table I3: IME regression by income group

<table>
<thead>
<tr>
<th></th>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity (richer countries)</td>
<td>0.437*** 0.443*** 0.379***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0048] [0.0045] [0.0053]</td>
</tr>
<tr>
<td>IM elasticity (poorer countries)</td>
<td>0.438*** 0.531*** 0.480***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0051] [0.0097] [0.0094]</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.54 0.74 0.85</td>
</tr>
<tr>
<td>Observations</td>
<td>7,767 7,754 7,310</td>
</tr>
<tr>
<td>HS2 × year FE</td>
<td>Yes</td>
</tr>
<tr>
<td>Origin × HS2 × year FE</td>
<td>Yes Yes</td>
</tr>
<tr>
<td>Destination × HS2 × year FE</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year fixed effects (column 1), origin × year fixed effects (column 2), origin × year and destination × year fixed effects (column 3) interacted with a dummy for a rich/poor country. Richer countries are defined as those with average GDP per capita over 2003-2013 greater than the median from the EDD sample. The data are aggregated at the year-origin-destination level for a set of origin-years listed in Table 1 in the paper. The sample is restricted to country pairs with $N_{ij} \geq 100$. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
### Table I4: IME regression, small firms excluded, extended sample

<table>
<thead>
<tr>
<th>Panel</th>
<th>Country pairs with $N_{ij} \geq 100$</th>
<th>All country pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</td>
<td>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</td>
</tr>
<tr>
<td></td>
<td>IM elasticity</td>
<td>$0.437^{***}$</td>
</tr>
<tr>
<td></td>
<td>Standard error</td>
<td>[0.0037]</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Variation in $\ln X_{ij}$ explained by FE, %</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Observations</td>
<td>14,216</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Year FE</th>
<th>Yes</th>
<th>Yes</th>
<th>Year FE</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Origin × year FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Destination × year FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year fixed effects (column 1), origin × year fixed effects (column 2), origin × year and destination × year fixed effects (column 3). The data are aggregated at the year-origin-destination level for a set of origin-years listed in Table 1 in the paper and Table I1 in the Online Appendix. Average and total exports per destination are calculated using the sales of firms with at least $1000 to that destination. Panel a) presents the regression on the sample of country-pairs with at least 100 exporters. Panel b) presents the regression on the full sample. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
Table I5: IME regressions, IV

<table>
<thead>
<tr>
<th></th>
<th>IV lag</th>
<th>IV lead</th>
<th>IV lag and lead</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a: country pairs with</strong> $N_{ij} \geq 100$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IM elasticity</td>
<td>0.392***</td>
<td>0.392***</td>
<td>0.399***</td>
</tr>
<tr>
<td>Standard error</td>
<td>[0.0056]</td>
<td>[0.0057]</td>
<td>[0.0061]</td>
</tr>
<tr>
<td>Observations</td>
<td>6,372</td>
<td>6,181</td>
<td>5,224</td>
</tr>
</tbody>
</table>

| **Panel b: all country pairs** |        |         |                 |
| IM elasticity           | 0.476*** | 0.479*** | 0.468***        |
| Standard error          | [0.0018] | [0.0022] | [0.0024]        |
| Observations            | 36,065  | 36,065  | 28,672          |

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin $\times$ year FE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Destination $\times$ year FE</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year fixed effects (column 1), origin $\times$ year fixed effects (column 2), origin $\times$ year and destination $\times$ year fixed effects (column 3) using lags and/or leads of the independent variable as instruments. The data are aggregated at the year-origin-destination level for a set of origin-years listed in Table 1 in the paper. Panel a) presents the regression on the sample of country-pairs with at least 100 exporters. Panel b) presents the regression on the full sample. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
Table I6: IME regression, disaggregated by HS2 product, extended sample

<table>
<thead>
<tr>
<th>IM elasticity</th>
<th>Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$</th>
<th>Standard error</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.646***</td>
<td>[0.0020]</td>
<td>58,609</td>
</tr>
<tr>
<td></td>
<td>0.598***</td>
<td>[0.0016]</td>
<td>56,560</td>
</tr>
<tr>
<td></td>
<td>0.518***</td>
<td>[0.0029]</td>
<td>29,906</td>
</tr>
<tr>
<td>Year × HS FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Origin × Year × HS FE</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Destination × Year × HS FE</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year × HS2 product fixed effects (column 1), origin × year × HS2 product fixed effects (column 2), origin × year × HS2 product and destination × year × HS2 product fixed effects (column 3). The data are aggregated at the year-origin-destination-HS2 industry level for a set of origin-years listed in Table 1 in the paper and Table I1 in the Online Appendix. The sample is restricted to country pairs with $N_{ij} \geq 100$. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
Table I7: IME regression, dropping industries with large share of intermediaries

<table>
<thead>
<tr>
<th>Panel a: country pairs with $N_{ij} \geq 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
</tr>
<tr>
<td>Standard error</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Variation in $\ln X_{ij}$ explained by FE, %</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel b: all country pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>IM elasticity</td>
</tr>
<tr>
<td>Standard error</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Variation in $\ln X_{ij}$ explained by FE, %</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year FE</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Origin × year FE</td>
<td>Yes</td>
</tr>
<tr>
<td>Destination × year FE</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Note: the table presents the estimated coefficients of the regression of log average exports on log total exports with year fixed effects (column 1), origin × year fixed effects (column 2), origin × year and destination × year fixed effects (column 3) dropping industries with a large share of intermediaries as defined in Chan (2017). The data are aggregated at the year-origin-destination level for a set of origin-years listed in Table 1 in the paper. Panel a) presents the regression on the sample of country-pairs with at least 100 exporters. Panel b) presents the regression on the full sample. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
Table I8: Estimates of $\bar{\theta}$

<table>
<thead>
<tr>
<th></th>
<th>$\bar{\theta}$</th>
<th>s. e.</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>All observations, no weights</td>
<td>18.61***</td>
<td>[0.787]</td>
<td>39,712</td>
</tr>
<tr>
<td>Weights $\sqrt{N_{ij}}$</td>
<td>4.481***</td>
<td>[0.0360]</td>
<td>39,712</td>
</tr>
<tr>
<td>Dropping $N_{ij} &lt; 100$</td>
<td>2.657***</td>
<td>[0.0175]</td>
<td>7,781</td>
</tr>
<tr>
<td>Dropping $M_{ij} &lt; 100$</td>
<td>2.360***</td>
<td>[0.0147]</td>
<td>5,267</td>
</tr>
</tbody>
</table>

Note: the table presents estimates of $\bar{\theta}$ as discussed in Section A of this Online Appendix. $N_{ij}$ denotes the number of exporters from $i$ to $j$ and $M_{ij}$ denotes the number of exporters from $i$ to $j$ that also export to $i$’s largest destination. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.

Table I9: QQ estimates of $\bar{\sigma}_\varphi$

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

Note: the table presents estimates of $\bar{\sigma}_\varphi$ as discussed in Section C of this Online Appendix. Robust standard errors are reported in brackets. *, **, and *** represent the 5%, 1%, and 0.1% significance levels respectively.
J  Appendix Figures

Note: Figures J1-J34 present the moments from the parameter draws that we obtained using the procedure to estimate the full Melitz-lognormal model parameters (see sections E and F of this Online Appendix for more details). Each graph corresponds to one parameter and consists of multiple panels for different origin countries in the EDD. Each panel consists of five horizontal areas corresponding to 5 different MCMC chains that we ran for each origin. The black, red, and blue stars denote the mean of the parameter draws in the full chain, the first half of the chain, and the second half of the chain, respectively. The black lines correspond to the 2.5 – 97.5 percentiles range for the draws.
Figure J1: Summary plots for QBE chains: $d_{i1}$
Figure J2: Summary plots for QBE chains: $d_{2}$

- mean: full chain
- mean: 1st half
- mean: 2nd half
- 95% CI
Figure J3: Summary plots for QBE chains: $d_{33}$
Figure J4: Summary plots for QBE chains: $d_{44}$
Figure J5: Summary plots for QBE chains: $d_{i5}$
Figure J6: Summary plots for QBE chains: $d_{ab}$
Figure J7: Summary plots for QBE chains: $d_{ij}$
Figure J8: Summary plots for QBE chains: $d_{38}$
Figure J9: Summary plots for QBE chains: $d_{ij}$
Figure J10: Summary plots for QBE chains: $d_{i10}$
Figure J11: Summary plots for QBE chains: $d_{i11}$
Figure J12: Summary plots for QBE chains: $d_{112}$
Figure J13: Summary plots for QBE chains: $d_{i13}$
Figure J14: Summary plots for QBE chains: $d_{i14}$
Figure J15: Summary plots for QBE chains: $d_{i15}$
Figure J16: Summary plots for QBE chains: $\mu_{f,i1}$
Figure J17: Summary plots for QBE chains: $\mu_{f,ij}$
Figure J18: Summary plots for QBE chains: $\mu_{f,i3}$
Figure J19: Summary plots for QBE chains: $\mu_{f,i,t}$
Figure J20: Summary plots for QBE chains: $\mu_{f,i5}$
Figure J21: Summary plots for QBE chains: $\mu_{f,i6}$
Figure J22: Summary plots for QBE chains: $\mu_{f,i7}$
Figure J23: Summary plots for QBE chains: $\mu_{f,i8}$
Figure J24: Summary plots for QBE chains: $\mu_{f,i9}$
Figure J25: Summary plots for QBE chains: $\mu_{f,i10}$
Figure J26: Summary plots for QBE chains: $\mu_{f,11}$
Figure J27: Summary plots for QBE chains: $\mu_{f,i2}$
Figure J28: Summary plots for QBE chains: $\mu_{f,i13}$
Figure J29: Summary plots for QBE chains: $\mu_{f_i,14}$
Figure J30: Summary plots for QBE chains: $\mu_{f,15}$
Figure J31: Summary plots for QBE chains: $\bar{\sigma}_{\phi,i}$
Figure J32: Summary plots for QBE chains: $\sigma_{\alpha,i}$
Figure J33: Summary plots for QBE chains: $\sigma_{f,i}$
Figure J34: Summary plots for QBE chains: $\rho_i$
Figure J35: 2-country world: changes in welfare after trade liberalization I

Note: the figure presents the change in welfare in response to a reduction in trade costs in a symmetrical two-country world in spirit of Melitz and Redding (2015) (see Online Appendix section H for the definitions and calibration of parameters). The x-axis represents the level of trade costs at which trade liberalization starts. The y-axis (left) represents the changes in welfare in different models. The y-axis (right panel) represents the local trade elasticity in the true model at the corresponding level of trade costs.
Figure J36: 2-country world: changes in welfare after trade liberalization II

Note: the figure represents the change in welfare in response to a reduction in trade costs in a symmetrical two-country world in spirit of Melitz and Redding (2015) (see Online Appendix section H for the definitions and calibration of parameters). The x-axis represents the level of trade costs at which trade liberalization starts. The y-axis (left) represents the changes in welfare in different models. The y-axis (right panel) represents the local trade elasticity in the true model at the corresponding level of trade costs.