The Intensive Margin in Trade: Moving Beyond Pareto

Ana M. Fernandes, World Bank

Pete Klenow, Stanford University

Sergii Meleshchuk, UC Berkeley

Martha Denisse Pierola, World Bank

Andres Rodríguez-Clare, UC Berkeley

September 22, 2016

Motivation

- **Canonical trade model: Melitz with Pareto**
  - Consistent with some firm-level facts (EKK)
  - Implies gravity equation (Chaney 2008) and simple summary statistics for welfare (ACR 2012)
- **Sharp prediction:**
  - Extensive margin should explain all variation in exports
    – conditional on fixed costs of exporting
- **Is this prediction supported by firm-level data?**
  - We use the World Bank's *Exporter Dynamics Database* to find out.

Theoretical Robustness of the Prediction

- Melitz-Pareto + demand shocks + fixed-cost shocks + Arkolakis (2010)
  - Eaton, Kortum and Kramarz (2011)
- Non-CES preferences:
  - Melitz and Ottaviano (2008), Arkolakis, Costinot, Donaldson and Rodríguez-Clare (2015)
- Non-monopolistic competition
  - Bernard, Eaton, Jensen and Kortum (2003)
- Multinational production
  - Arkolakis, Ramondo, Rodríguez-Clare and Yeaple (2014)

Relation to the Literature

- Eaton, Kortum and Kramarz (2011)
  - We exploit multiple destinations and origins in the EDD
  - MLE estimation with lognormal distribution of productivity
- Closely related papers using EDD
  - Freund and Pierola (2015)
  - Fernandes, Freund and Pierola (2015)
  - Spearot (2015)
- Papers arguing that lognormal $\succ$ Pareto
  - Head, Mayer and Thoening (2014)
  - Bas, Mayer and Thoening (2015)
- Trade elasticity and welfare under lognormal or truncated Pareto
  - Feenstra (2014)
  - Head, Mayer and Thoening (2014)
  - Melitz and Redding (2015)
  - Bas, Mayer and Thoening (2015)

Outline

1. Canonical trade model
2. Exporter Dynamics Database
3. Empirical results – Intensive Margin puzzle
4. Potential explanations:
    a. Lognormal distribution

Canonical model

Plain vanilla Melitz-Pareto

- Monopolistic competition with CES preferences, $EoS = \sigma$
- In each country firms draw productivity φ from a Pareto distribution G_i with shape parameter $\theta > \sigma - 1$
- A firm in i pays fixed costs F_{ij} and iceberg costs τ_{ij} to sell in market j
- Notation:
  - $N_{ij} \equiv$ number of firms exporting from i to j
  - $x_{ij} \equiv$ exports per firm exporting from i to j
  - $X_{ij} \equiv x_{ij} \cdot N_{ij} \equiv$ total exports from i to j

Canonical model

- Using $\bar{\theta} \equiv \frac{\theta}{\sigma-1}$, the model implies

$$N_{ij} = N_i \left(\frac{w_i}{b_i}\right)^{-\theta} \left(\frac{\sigma}{A_j}\right)^{-\bar{\theta}} \tau_{ij}^{-\theta} F_{ij}^{-\bar{\theta}}$$

$$x_{ij} = \frac{\bar{\theta}}{\bar{\theta}-1} F_{ij}$$

- Using $\tau_{ij} = \tau_i^o \tau_j^d \widetilde{\tau}_{ij}$ and $F_{ij} = F_i^o F_j^d \widetilde{F}_{ij}$ we get

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

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

Canonical model

- The Intensive Margin Elasticity (IME) is the coefficient α from

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

- We then have

$$\hat{\alpha} = \frac{cov(\ln \widetilde{x}_{ij}, \ln \widetilde{X}_{ij})}{var\left(\ln \widetilde{X}_{ij}\right)} = -\frac{\left(\bar{\theta} - 1\right) var(\ln \widetilde{F}_{ij}) + \theta\, cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})}{var\left(-\theta \ln \widetilde{\tau}_{ij} + \left(1 - \bar{\theta}\right) \ln \widetilde{F}_{ij}\right)}$$

Canonical model

- Intensive Margin Elasticity (IME):

$$\hat{\alpha} = -\frac{(\bar{\theta}-1)\, var(\ln \widetilde{F}_{ij}) + \theta\, cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})}{var\left(-\theta \ln \widetilde{\tau}_{ij} + (1-\bar{\theta}) \ln \widetilde{F}_{ij}\right)}$$

- This leads to

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

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

Canonical model

- Equation

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

implies that

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

- This leads to:

Observation 3: $\frac{cov(\ln \tilde{x}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})} < 0 \Rightarrow \frac{cov(\ln \widetilde{F}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})} < 0$

Canonical model

- Denote average exports per firm in percentile *pct* as x_{ij}^{pct}
- Consider the following regression:

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

- Defining $\text{IME}^{pct} \equiv \hat{\alpha}^{pct}$, we have:

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

Canonical model

- We can use

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

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

to infer $\widetilde{\tau}_{ij}$ and $\widetilde{F}_{ij}$ from data on x_{ij} and N_{ij}

- Use results to compute

$$corr(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij}), \quad \frac{cov(\ln \widetilde{F}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})}, \quad \frac{cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})}$$

Exporter Dynamics Database (EDD)

- 50 countries over subsets of 2003–2013
- Annual exports disaggregated by:
  - firms
  - destinations
  - HS 6-digit products
- EDD total non-oil exports $\approx$ non-oil exports in COMTRADE/WITS
- $\sim$ 5% of non-oil exports w/o China, $\sim$ 12% w/ China
- See Fernandes, Freund and Pierola (2015) for documentation

Sample country-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 Rep.	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

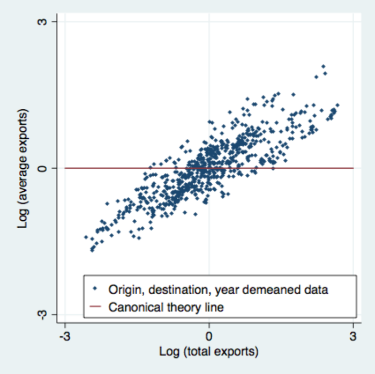
* indicates that Uganda does not have data for 2006

IM and EM in the data

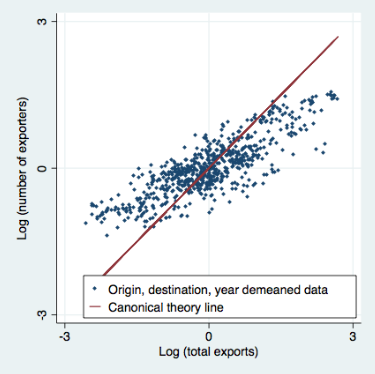
- 4 main destinations: USA, France, Germany, Japan
- Origin-destination pairs with over 100 exporters

IM and EM: data vs. theory

Intensive Margin

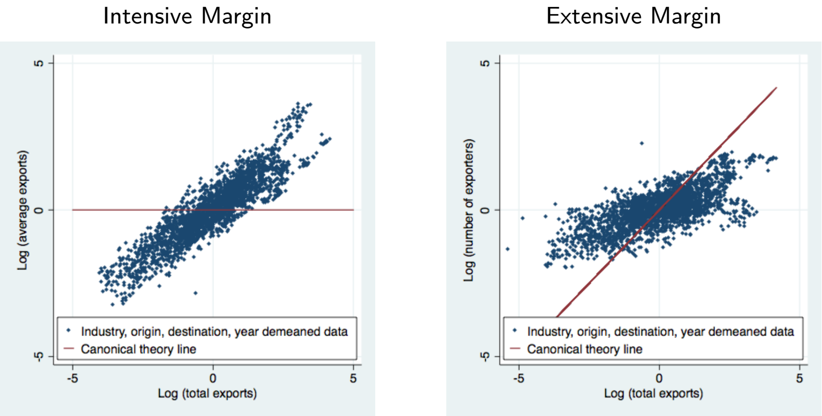


Extensive Margin



- 4 main destinations: USA, France, Germany, Japan
- Origin-destination pairs with over 100 exporters

IM and EM disaggregated by industry



- 4 main destinations; pairs with over 100 exporters
- 15 industries (combinations of HS 2-digit industries)

The Intensive Margin Elasticity (IME)

Regression evidence

	Coefficient from $\ln x_{ij}$ on $\ln X_{ij}$		
IME	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.

- Results are robust to:
  - including all destinations
  - instrumenting total exports with lead or lag of itself
  - disaggregating by industry
  - excluding small firms
  - including origin-destination pairs with fewer than 100 exporters

Empirical correlations

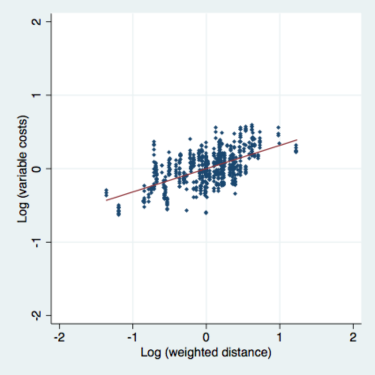
	$corr(\ln N_{ij}, \ln x_{ij})$	$corr(\ln \widetilde{\tau}_{ij}, \ln \widetilde{F}_{ij})$
Raw data	0.366	
	[0.035]	
Purged of:		
Origin FE	0.500	
	[0.033]	
Destination FE	0.352	
	[0.036]	
Origin and Destination FE	0.418	-0.891
	[0.034]	[0.019]

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

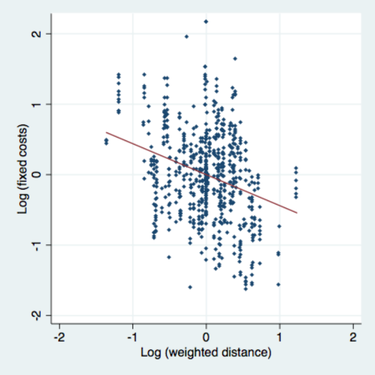
- Positive correlation between IM and EM in the data
- Implies highly negative correlation between fixed and variable trade costs

Trade costs and distance

Variable costs



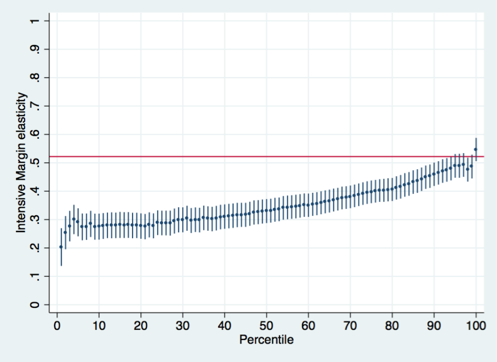
Fixed costs



$$\sigma = 5, \theta = 5$$

- Variable costs are increasing with distance: elasticity = 0.31 (s.e. 0.02)
- Fixed costs are decreasing with distance: elasticity = –0.44 (s.e. 0.05)

IME by percentiles



- Red line indicates theoretical IME

Potential explanations of 50% IM elasticity

- Multi-product firms [in the paper - fails]
- Granularity [in the paper - fails]
- **Lognormal distribution [today]**

Lognormal distribution

Simple Model: Theory

- Start with an identity, $x_{ij} = \sigma F_{ij} \frac{x_{ij}}{x_{ij}^{min}}$, then write

$$x_{ij} = \sigma F_{ij} \frac{\int_{\varphi_{ij}^*}^{\infty} \varphi^{\sigma-1} \frac{dG_i(\varphi)}{1-G_i(\varphi_{ij}^*)}}{\left(\varphi_{ij}^*\right)^{\sigma-1}} \equiv \sigma F_{ij} H_i(\varphi_{ij}^*)$$

- With Pareto, $H_i(\varphi_{ij}^*) = \frac{\overline{\theta}}{\overline{\theta}-1}$
- With lognormal (Bas, Mayer and Thoenig, 2015 – BMT),

$$H_i(\varphi_{ij}^*) = \Omega\left(\frac{N_{ij}}{N_i}; \bar{\sigma}_\varphi\right),$$

where $\bar{\sigma}_\varphi \equiv (\sigma - 1)\sigma_\varphi$

- $\Omega(N_{ij}/N_i; \bar{\sigma}_\varphi)$ is increasing in N_{ij}/N_i

$\Longrightarrow \quad \frac{x_{ij}}{x_{ij}^{\mathrm{min}}}$ increases in $\frac{N_{ij}}{N_i}$, so $IME > 0$ even with $var(\widetilde{F}_{ij}) = 0$

Lognormal distribution

Simple Model: Data

- Combine N_{ij}/N_i and estimate of $\bar{\sigma}_\varphi$ to compute $\Omega(N_{ij}/N_i)$, then proceed as above to generate IME and model-implied trade costs.
- Estimate $\bar{\sigma}_\varphi$ using QQ regression as in Head-MT but allowing for truncation and using N_i from Bento and Restuccia (2015)
- IME is 0.43, with positive slope across percentiles, as in the data
- Partial success regarding model-implied trade costs:

$$\frac{cov(\ln \widetilde{F}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})}, \frac{cov(\ln \widetilde{\tau}_{ij}, \ln \widetilde{dist}_{ij})}{var(\ln \widetilde{dist}_{ij})} > 0$$

$$corr(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij}) < 0$$

Lognormal distribution

Simple model: unrealistic predictions – as discussed by EKK

- Perfect hierarchy of destinations
- Perfect correlation of sales across markets
- Minimum exports pins down fixed costs

Full Melitz-lognormal model

Basic assumptions

- Allow for firm-destination demand α_j and fixed costs f_j shocks
- Assume:

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

Full Melitz-lognormal model

Latent sales

- Let $Z_{ij} \equiv \ln \left[A_j \left(w_i \tau_{ij}\right)^{1-\sigma}\right] + \ln \alpha_j + (\sigma - 1) \ln \varphi$

- Then

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

where

$$d_{ij} \equiv \ln \left[A_j \left(w_i \tau_{ij}\right)^{1-\sigma}\right] + \mu_\alpha + (\sigma - 1)\, \mu_{\varphi,i}$$

$$\bar{\sigma}_\varphi \equiv (\sigma - 1)\, \sigma_\varphi$$

Full Melitz-lognormal model

Likelihood function for firm sales

- We observe log sales

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

- With two destinations, the likelihood function is

$$L\left(\theta | \left\{x_{i1}\left(k_i\right), x_{i2}\left(k_i\right)\right\}_{i,k_i}\right) = \prod_i \prod_{k_i=1}^{N_i} \left[g_{(X_{i1},X_{i2})}\left(x_{i1}\left(k_i\right), x_{i2}\left(k_i\right)\right)\right]$$

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

Full Melitz-lognormal model

MCMC estimation using likelihood function

- Restrict to i with $N_{ij} > 100$ for $j =$ US, Germany
- Likelihood is potentially non-concave in θ, and θ has 75 elements
- Estimation approach by Chernozhukov and Hong (2003)
  - Metropolis-Hastings MCMC algorithm to construct chain $\theta^{(n)}$
  - Accept with probability that increases in likelihood
    - Drop first 750k runs and continue until $n = 3$ million
  - $\bar{\theta} \equiv \frac{1}{N} \sum_{n=1}^{N} \theta^{(n)}$ is a consistent estimator of θ
  - Variance of $\theta^{(n)}$ used to construct confidence intervals for $\bar{\theta}$

Full Melitz-lognormal model

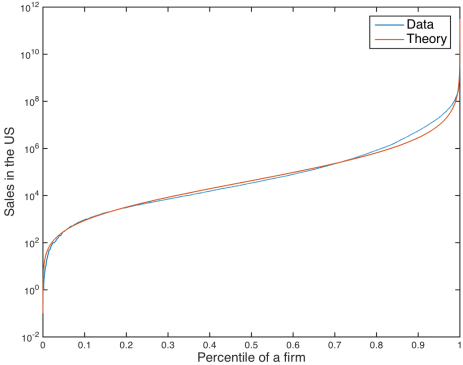
Identification

- Data on X_{ij} and N_{ij} "identifies" d_{ij} and $\bar{\mu}_{f,ij}$
- Variance of sales within each ij "identifies" $\bar{\sigma}_{\varphi}^2 + \sigma_{\alpha}^2$
- Cov. of sales across destinations "identifies" $\bar{\sigma}_{\varphi}^2$ from σ_{α}^2
- Skewness of sales distribution within each ij "identifies" σ_f^2

Full Melitz-lognormal model

Theoretical and empirical CDF

Log sales and percentile of a firm

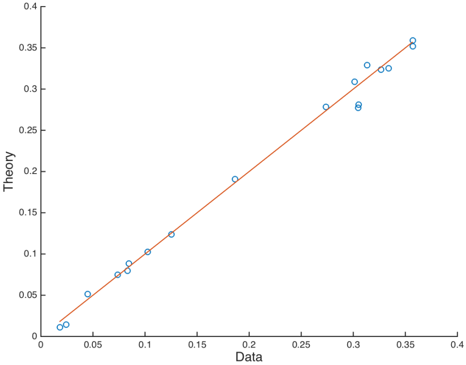


- Sales to the US from some origin

Full Melitz-lognormal model

Hierarchy

Fraction of firms selling to less attractive destination only



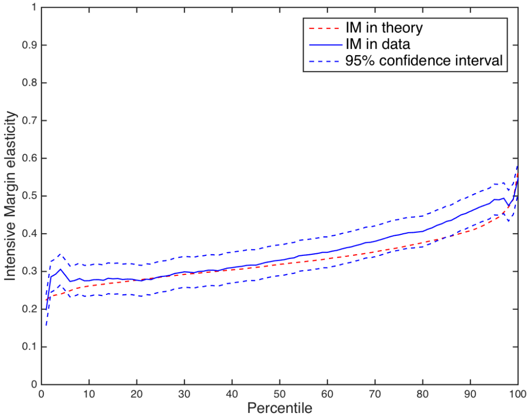
Full Melitz-lognormal model
IME

	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]

- $\sigma_f > 0$ helps explain IME > 0

Full Melitz-lognormal model

IME for each percentile



Full Melitz-lognormal model

Implied trade costs

	Estimate	95% CI
$corr\left(\widetilde{F}_{ij}, \widetilde{\tau}_{ij}\right)$	0.69	[0.53, 0.80]
	Distance elasticity	
Fixed costs	1.93	[1.45, 2.33]
Variable costs	0.53	[0.47, 0.57]

Full Melitz-lognormal model
Summary

	Canonical model	Lognormal
Overall IME with $var\left(\widetilde{F}_{ij}\right) = 0$	No	**Yes**
$corr\left(\ln \widetilde{F}_{ij}, \ln \widetilde{\tau}_{ij}\right) > 0$	No	**Yes**
$corr\left(\ln \widetilde{\tau}_{ij}, \ln distance\right) > 0$	**Yes**	**Yes**
$corr\left(\ln \widetilde{F}_{ij}, \ln distance\right) > 0$	No	**Yes**
IME for each percentile	No	**Yes**

Some implications

- Any model with trade and firm heterogeneity shall be able to reproduce significant intensive margin
- Possible solution to total export elasticity with respect to exchange rate puzzle:
  - In Berman et. al. (2012) model with Pareto distribution extensive margin elasticity is much larger than in the data
  - Model with lognormal distribution of productivities implies lower extensive margin elasticity

Conclusions

- **Exporter Dynamics Database (EDD)**
  - Wide coverage (50 exporting countries, subsets of 2003–2013)
  - Reveals that IM accounts for 50% of export variation
  - IME rises systematically with size percentile of exporting firms
- **EDD facts pose a puzzle for canonical Melitz-Pareto model**
- **Puzzle solved if productivity is lognormal and heterogeneous fixed trade costs**
  - Melitz-Lognormal model fits EDD with fixed trade costs that correlate positively with variable trade costs and with distance
  - Model also matches shape of IME across percentiles