A Tale of Two Tails: Productivity Distribution and the Gains from Trade

Sergey Nigai

ETH Zurich, KOF and CESifo

December 2016

Productivity Distribution and Selection Effects

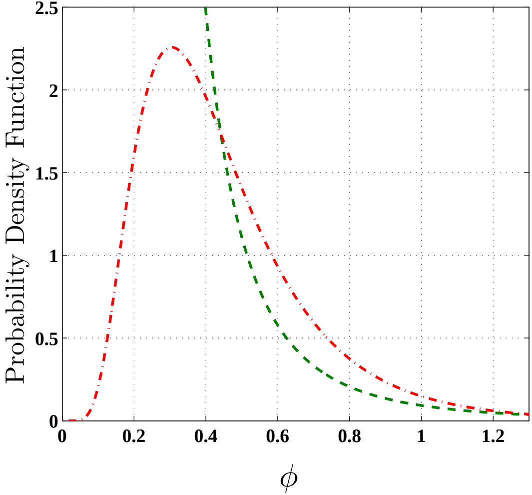

Figure: WHEN YOU OBSERVE EVERYTHING

Productivity Distribution and Selection Effects

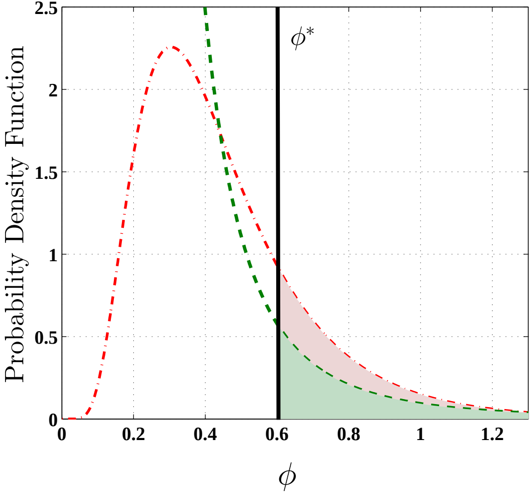

Figure: WHEN YOU OBSERVE EVERYTHING

Left Tail (bottom 95%) : Pareto or Log-normal?

Left Tail (bottom 95%) : Pareto or Log-normal?

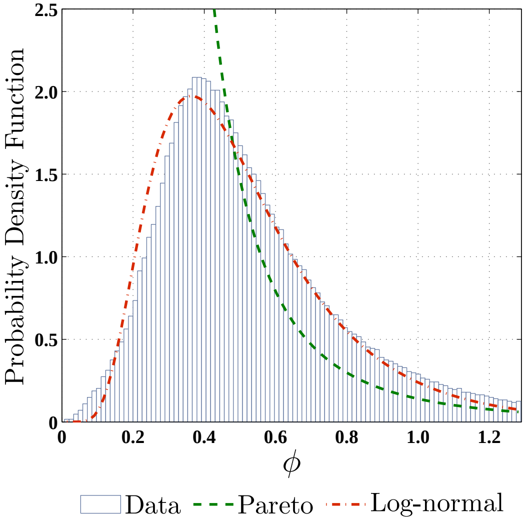

Figure: Empirical and Parametric p.d.f.'s (Left Tail)

Right Tail (top 5%) : Pareto or Log-normal?

Right Tail (top 5%) : Pareto or Log-normal?

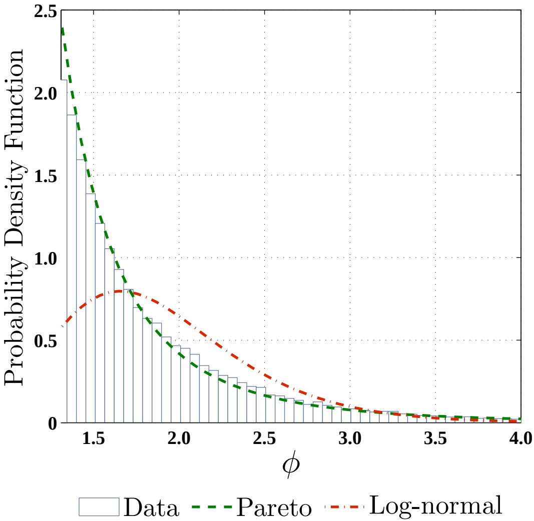

Figure: Empirical and Parametric p.d.f.'s (Right Tail)

Right Tail (top 5%) : Pareto or Log-normal?

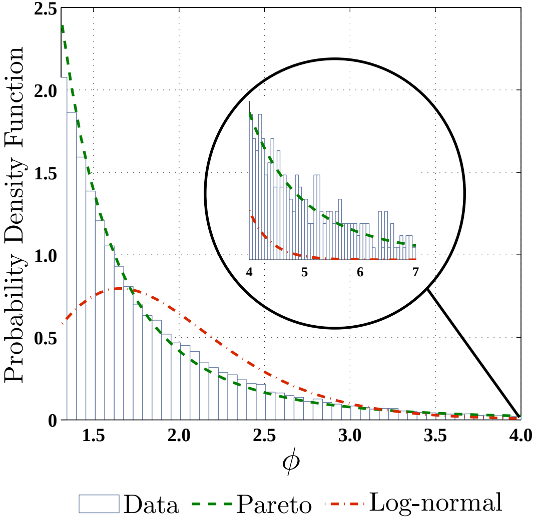

Figure: Empirical and Parametric p.d.f.'s (Right Tail)

Pareto or Log-normal?

Three main observations:

- Pareto does not capture the shape of the left tail
- Log-normal underpredicts the thickness of the right tail
- Neither captures the empirical distribution over the entire support

So what?

So what?

Are deviations from the empirical distribution harmful?

- Both (un-)bounded Pareto and Log-normal lead to significant errors in trade outcomes:
  - Welfare gains from trade
  - Extensive margin of trade
  - Intensive margin of trade

So what?

Are deviations from the empirical distribution harmful?

- Both (un-)bounded Pareto and Log-normal lead to significant errors in trade outcomes:
  - Welfare gains from trade
  - Extensive margin of trade
  - Intensive margin of trade

Why?

- Efficiency distribution determines magnitude of the selection effects:
  - Entry and exporting
  - Available varieties and their prices

Contribution of this paper

This paper proposes using a mixed distribution dubbed *Two-piece*. The distribution has several advantages:

- Models left tail as Log-normal (captures bell shape)
- Models right tail as Pareto (captures fat right tail)
- Fits the data considerably better than (un-)bounded Pareto and Log-normal almost everywhere
- Produces negligible errors in the estimates of the gains from trade and other trade outcomes
- Still parametric, tractable and well-behaved distribution

Large and rich literature on both Pareto and Log-normal in trade!

Lit.Review

Log-normal meets Pareto

Following Cooray and Ananda (2005) and Scollnik (2005) start with:

$$f_L(\phi) = \frac{1}{\sqrt{2\pi} s\phi} e^{-\frac{1}{2}\left(\frac{\ln \phi - \mu}{s}\right)^2} \text{ and } f_P(\phi) = \frac{\alpha\theta^\alpha}{\phi^{\alpha+1}}. \tag{1}$$

Derive Two-piece distribution by imposing the following conditions:

- Random variable ϕ follows Log-normal up to a threshold, θ, and Pareto after that
- Two-piece is a well-behaved distribution:
  - Continuous
  - Differentiable
  - p.d.f. and c.d.f. have necessary properties

Two-piece distribution

The resulting distribution has the following c.d.f. :

$$F(\phi) = \begin{cases} \dfrac{\bar{p}}{\Phi[\alpha s(\alpha, \bar{p})]} \Phi \left(\alpha s(\alpha, \bar{p}) + \dfrac{\ln \phi - \ln \theta}{s(\alpha, \bar{p})} \right) & \text{for } \phi \in (0, \theta] \\ 1 - (1 - \bar{p}) \dfrac{\theta^\alpha}{\phi^\alpha} & \text{for } \phi \in [\theta, \infty), \end{cases}$$

and p.d.f. :

$$f(\phi) = \begin{cases} \dfrac{\bar{p}}{\Phi[\alpha s(\alpha, \bar{p})]} \dfrac{1}{\sqrt{2\pi} s(\alpha, \bar{p}) \phi} e^{-\frac{1}{2} \left(\alpha s(\alpha, \bar{p}) - \frac{\ln \theta - \ln \phi}{s(\alpha, \bar{p})} \right)^2} & \text{for } \phi \in (0, \theta] \\ (1 - \bar{p}) \dfrac{\alpha \theta^\alpha}{\phi^{\alpha+1}} & \text{for } \phi \in [\theta, \infty), \end{cases}$$

Two-piece distribution

$\Phi(\cdot)$ is c.d.f. of standard normal and $s(\bar{p}, \alpha)$ is an implicit function which defines s given $\bar{p}$ and α according to:

$$\Phi\left[\alpha s(\alpha, \bar{p})\right] \sqrt{2\pi}[\alpha s(\alpha, \bar{p})] e^{\frac{1}{2}[\alpha s(\alpha, \bar{p})]^2} = \frac{\bar{p}}{1 - \bar{p}}.$$

The Two-piece distribution is characterized by the following parameters:

- First scale parameter, θ, identifies the cut-off point
- Second scale parameter, $\bar{p}$, identifies the share of population that follows Log-normal
- Shape parameter, α, comes from the original Pareto distribution.

Parameterized example

Set $\theta = 1$, $\bar{p} = 0.95$, $\alpha = 3$. Choose parameters of Log-normal and Pareto to match the first two moments.

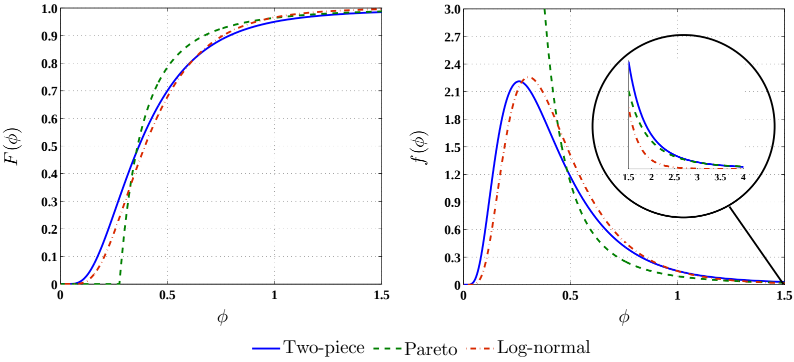

Figure: TWO-PIECE, LOG-NORMAL AND PARETO DISTRIBUTIONS WITH IDENTICAL FIRST TWO MOMENTS

Estimation

The data come from ORBIS and cover almost 1 mln. French entities in 2012.

1. Keep all firms in the production sectors.
2. Calculate domestic sales as total sales net of export revenues
3. Calculate productivity distribution
4. Generate 100,000 quantiles (for numerical purposes)
   - This grid covers c.d.f. on the support $[0.00001; 0.99999]$

QQ-estimator

The estimator solves the following:

$$\min_{\Theta_\ell} \left\{ \sum_q \left(\ln\left[Q_e(q)\right] - \ln\left[Q_\ell(q|\Theta_\ell)\right]\right)^2 \right\}, \tag{2}$$

where

- q is the grid of the c.d.f. 100,000 data points in $[0.00001; 0.99999]$
- $Q_e(q)$ the empirical quantile function evaluated at q.
- $Q_\ell(q)$ is the parametric quantile function of type ℓ evaluated at q.
- Θ_ℓ is the vector of parameters of the parametric function of type ℓ to estimate

Results I: Fit on different intervals of the support

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	3.033	1.185	0.938	0.058	0.465	0.221	0.026	0.033
	(0.006)	(0.005)	(0.001)					
Log-normal	0.569	-0.701		0.069	0.415	0.194	0.156	0.304
	(0.001)	(0.001)						
Pareto	1.914	0.294		0.236	1.405	0.840	0.344	0.648
	(0.005)	(0.001)						
B. Pareto	0.372	0.214	1.438	0.183	1.105	0.582	0.406	0.799
	(0.026)	(0.001)	(0.015)					

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in the case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: Estimation Results

Results II: p.d.f.

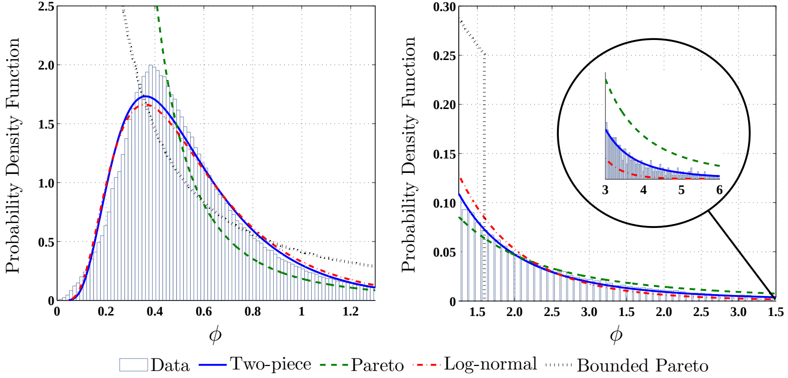

Figure: DENSITY OF TWO-PIECE, LOG-NORMAL AND PARETO VS. DATA

Workhorse Heterogeneous Trade Model

The model is standard and follows Arkolakis, Demidova, Klenow and Rodríguez-Clare (2008) and only slightly deviate from Arkolakis, Costinot and Rodríguez-Clare (2012), and Melitz and Redding (2014).

- CES preferences
- Labor is the only factor of production: L_i
- Variable and fixed cost of exporting: τ_{ij} and f_{ij}
- Fixed cost of entry: f_i^e
- Productivity is randomly drawn upon paying f_i^e: ϕ
- Free entry, markets clear

Workhorse Heterogeneous Trade Model

The model's solution for J countries is governed by $2 \times J$ equations. The first set of $1 \times J$ equation comes from the Free entry condition:

$$\sum_{j \in J} \left(\int_{\phi_{ij}^*}^{\bar{\phi}} w_j f_{ij} (\phi_{ij}^*)^{1-\sigma} \phi^{\sigma-1} f(\phi) d\phi - \int_{\phi_{ij}^*}^{\bar{\phi}} w_j f_{ij} f(\phi) d\phi \right) = w_i f_i^e,$$

The second set of $1 \times J$ equations comes from the Labor market clearing condition:

$$\frac{N_i}{1 - F(\phi_{ii}^*)} \sum_{j \in J} \left(\frac{(\sigma-1) w_j}{w_i} f_{ij} (\phi_{ij}^*)^{1-\sigma} \int_{\phi_{ij}^*}^{\bar{\phi}} \phi^{\sigma-1} f(\phi) d\phi + f_i^e \right) + \sum_{j \in J} \frac{N_j}{1 - F(\phi_{jj}^*)} f_{ji} \int_{\phi_{ji}^*}^{\bar{\phi}} f(\phi) d\phi = L_i.$$

Solution

The solution of the system depends on two selection statistics:

- $1 - F(\phi_{ij}^*)$ which measures the probability of firms from i being active in j for all i, j
- $\int_{\phi_{ij}^*}^{\bar{\phi}} \phi^{\sigma-1} f(\phi) d\phi$, which is required to calculate total revenues of firms from i in market j for all i, j
- Third statistics $\int_{\phi_{ij}^*}^{\bar{\phi}} f(\phi) d\phi$ is redundant due to the following identity:

$$\int_{\phi_{ij}^*}^{\bar{\phi}} f(\phi) d\phi = \int_0^{\bar{\phi}} f(\phi) d\phi - \int_0^{\phi_{ij}^*} f(\phi) d\phi = 1 - F(\phi_{ij}^*).$$

Benchmark

Results under all parametric distributions will be compared to those under numerical benchmark:

- $1 - F(\phi_{ij}^*)$ is calculated by using empirical c.d.f. in a non-parametric form
- $\int_{\phi_{ij}^*}^{\bar{\phi}} \phi^{\sigma-1} f(\phi) d\phi$ is calculated by numerical trapezoidal integration

Parameterization of the model

Without loss of generality, the model's primitives are chosen as follows:

Parameter	J	L_1	L_2	f_1^e	f_2^e	f_{11}	f_{22}	σ
	2	100	50	1	1	0.001	0.001	4

Table: Primitives of the model

Simplistic but general framework, amenable to increasing the number of countries/sectors.

Counterfactual experiments

Define changes in consumer welfare as τ gradually goes from 3 to 1:

$$\text{Welfare Gains} = 100\% \times \left(\frac{w_i(\tau)}{P_i(\tau)} \frac{P_i(\tau')}{w_i(\tau')} - 1 \right),$$

Define error in the estimates of the welfare gains from trade of parametric model ℓ as:

$$Error_\ell = \text{Welfare Gains}(\tau') - \text{Welfare Gains}_\ell(\tau'),$$

$\ell =$ Two-piece, Log-normal, (un-)bounded Pareto.

Experiment 1: Falling τ_{ij} at high f_{ij}

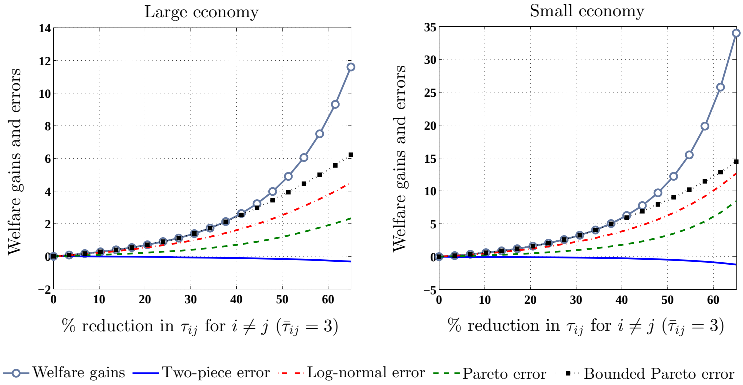

Figure: BENCHMARK WELFARE GAINS AND ERRORS: EXPERIMENT 1

Experiment 2: Falling τ_{ij} at low f_{ij}

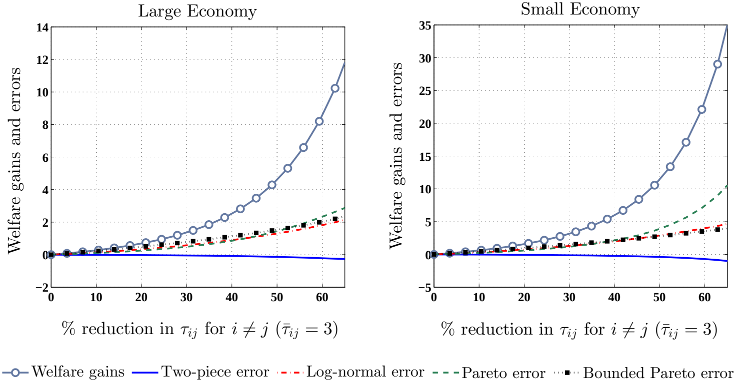

Figure: Benchmark welfare gains and errors: Experiment 2

Other trade outcomes

Researchers are often interested in two other trade statistics: intensive and extensive margins of trade. define mean squared errors as follows:.

$$MSE_{\ell}(\lambda) = \sqrt{\frac{1}{J}\sum_{j}\left(\lambda_{jj} - \lambda_{jj,\ell}\right)^2}; \quad MSE_{\ell}(\chi) = \sqrt{\frac{1}{J}\sum_{j}\left(\chi_{j} - \chi_{j,\ell}\right)^2},$$

where λ_{jj} is the share of intra-trade (governs intensive margin) and χ_{ij} is the share of exporters (extensive margin).

Other trade outcomes

Repeat Experiments 1 and 2 and calculate mean squared errors under four parametric distributions.

	Variable	Share of intratrade				Share of Exporters			
	τ_{ij} for $i \neq j$	3.0	2.4	1.8	1.2	3.0	2.4	1.8	1.2
Exp. 1	Two-piece	1.57	2.69	5.02	7.53	0.04	0.08	0.20	1.85
	Log-normal	31.38	51.16	78.91	81.65	2.29	7.80	27.77	93.02
	Bounded Pareto	12.72	22.06	40.39	63.01	0.51	1.02	2.25	7.20
	Pareto	37.30	71.05	121.29	75.83	0.75	1.51	26.44	155.79
Exp. 2	Two-piece	1.49	2.59	4.57	6.67	0.27	0.76	2.41	6.32
	Log-normal	21.61	29.11	33.42	22.13	23.66	54.58	126.29	265.54
	Bounded Pareto	14.79	25.99	48.96	80.46	1.92	3.88	9.25	23.73
	Pareto	35.13	34.51	29.28	16.23	8.67	91.78	197.69	326.43

Table notes: For expositional purposes, due to the fractional nature of the variables the results are reported in one thousandths.

Table: MEAN SQUARED ERRORS IN THE SHARE OF INTRATRADE AND EXPORTERS

If time permits

Several extensions and sensitivity checks:

1. Comparing to 3-parameter distributions 3-param. distr.
2. Alternative weighting scheme Weighting
3. Removing data points at the extremes Extreme data points
4. Testing for sensitivity to the choice of the country Other countries
5. Alternative measures of productivity Other measures of ϕ
6. Truncation of the Two-piece distribution Truncation
7. External Validity: City Size Distribution Out of sample test

Thank you!

Three-parameter distributions

Generalized Pareto and Three-parameter Log-normal c.d.f.s:

$$F_{GP}(\phi) = 1 - \left(1 + \frac{\eta(\phi - \psi)}{\xi}\right)^{-1/\eta} \quad \text{and} \quad F_{TLN}(\phi) = \Phi\left(\frac{\ln(\phi - \nu) - \mu}{\delta}\right),$$

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot.m 5%	Top 5%	Top 1%
Generalized Pareto	5.399	-0.081	0.424	0.144	1.011	0.512	0.196	0.409
	(0.354)	(0.006)	(0.003)					
3 P. Log-Normal	0.518	-0.608	-0.042	0.067	0.252	0.120	0.195	0.367
	(0.004)	(0.009)	(0.004)					

Table notes: (I), (II) and (III) refer to the shape, location and scale parameters. All parameters are estimated using 100,000 quantile data points.

Table: ESTIMATION RESULTS (ALTERNATIVE PARAMETRIC DISTRIBUTIONS) Back to main

Extreme data points

Remove 1000 largest data points from the right:

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	3.338	1.326	0.959	0.060	0.457	0.217	0.073	0.133
	(0.023)	(0.012)	(0.001)					
Log-normal	0.564	-0.704		0.064	0.425	0.199	0.113	0.192
	(0.001)	(0.001)						
Pareto	1.945	0.296		0.240	1.411	0.845	0.382	0.755
	(0.005)	(0.001)						
B. Pareto	0.274	0.209	1.373	0.177	1.086	0.567	0.378	0.704
	(0.025)	(0.001)	(0.012)					

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: TRUNCATED DISTRIBUTION (FROM THE RIGHT)

Extreme data points

Remove 1000 smallest data points from the left: Back to main

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	2.983	1.133	0.930	0.044	0.308	0.156	0.028	0.041
	(0.004)	(0.003)	(0.001)					
Log-normal	0.563	-0.698		0.060	0.249	0.122	0.165	0.317
	(0.001)	(0.001)						
Pareto	1.923	0.296		0.224	1.242	0.780	0.339	0.639
	(0.004)	(0.001)						
B. Pareto	0.460	0.220	1.479	0.171	0.956	0.530	0.388	0.773
	(0.022)	(0.001)	(0.014)					

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: TRUNCATED DISTRIBUTION (FROM THE LEFT)

Other countries

Country	Parameters			Root Mean Squared Error			
	(I)	(II)	(III)	Two-piece	Log-normal	Pareto	Bounded Pareto
France	2.880	1.057	0.923	0.053	0.070	0.228	0.181
	(0.006)	(0.004)	(0.001)				
Italy	3.820	2.463	0.993	0.068	0.068	0.308	0.191
	(0.018)	(0.025)	(0.001)				
Japan	2.723	1.143	0.921	0.043	0.060	0.215	0.143
	(0.014)	(0.009)	(0.002)				
Norway	3.842	3.088	0.997	0.064	0.064	0.345	0.182
	(0.025)	(0.042)	(0.000)				
Portugal	2.637	0.925	0.885	0.038	0.070	0.198	0.154
	(0.008)	(0.004)	(0.001)				
Spain	3.049	1.308	0.952	0.053	0.061	0.247	0.181
	(0.008)	(0.006)	(0.001)				
Sweden	3.471	2.035	0.988	0.071	0.073	0.317	0.195
	(0.014)	(0.015)	(0.001)				
Average	3.065	1.648	0.955	0.052	0.062	0.276	0.182

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in the case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: Estimation Results for Different Countries

Alternative measure of productivity

So far, I've looked at measures consistent with Melitz (2003). This measure would not be accurate under:

- Variable mark-ups
  - Bernard, Eaton, Jensen and Kortum (2003), Melitz and Ottaviano (2008), Simonovska (2015), Edmond, Midrigan and Xu (2015)
- Second dimension of heterogeneity across firms:
  - Egger and Kreickemeier (2009), Bustos (2011) and many others.

Unfortunately, the data don't allow deriving measures of ϕ under these conditions (no wage data, no mark-up data). Instead, I use a cost-based approach.

Alternative measure of productivity

I use data on total cost of goods sold and total cost of employees in Japan 2012 (219,454 observations) to calibrate the productivity parameter from the following relationship:

$$c(\phi)q = \frac{w(\phi)q}{\phi},$$

where $c(\phi)q$ and $w(\phi)q$ are observed total cost of goods sold and observed cost of employees, respectively.

Alternative measure of productivity

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	1.997	1.224	0.773	0.162	1.491	0.673	0.166	0.362
	(0.009)	(0.014)	(0.005)					
Log-normal	0.621	-0.177		0.210	1.319	0.621	0.558	1.154
	(0.002)	(0.001)						
Pareto	1.653	0.457		0.252	2.208	1.056	0.197	0.266
	(0.003)	(0.001)						
B. Pareto	1.653	0.457	(2.E+09)	0.252	2.208	1.056	0.197	0.266
	(0.003)	(0.001)	(4.E+09)					

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in the case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: Measure of productivity under variable markups

Truncation of the Two-piece distribution

Such an extension would be straightforward and would require mixing the following two p.d.f.s:

$$f_L(\phi) = \frac{1}{\sqrt{2\pi} s \phi} e^{-\frac{1}{2}\left(\frac{\ln \phi - \mu}{s}\right)^2} \quad \text{and} \quad f_{BP}(\phi) = \frac{\alpha \theta^\alpha \phi^{-\alpha-1}}{1 - \theta^\alpha \phi_h^{-\alpha}},$$

where ϕ_h would serve as an upper bound.

Back to main

External validity test

Debate about Pareto vs. Log-normal is not unique to International Trade. Long history in the literature on the city size distribution:

- Gabaix (1999, QJE) – Upper tail is Pareto
- Eeckhout (2004, AER) – But Log-normal fits the data better over the whole support
- Levy (2009, AER) – Yes but the tail is not Log-normal at all. Look at the top 0.6 percent of the largest cities in the QQ-plot!
- Eeckhout (2009, AER) – Yes but who cares about top 0.6 percent. Fitting the distribution on top 0.6 percent of observations only is useless.

Applying Two-Piece to Eeckhout (2009) data

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	1.205	6.178	0.978	0.089	0.209	0.222	0.155	0.256
	(0.034)	(0.774)	(0.003)					
Log-normal	1.751	-1.738		0.099	0.257	0.261	0.191	0.322
	(0.001)	(0.001)						
Pareto	0.619	0.035		0.679	2.909	1.900	1.412	2.728
	(0.004)	(0.001)						
B. Pareto	0.126	0.013	5.264	0.354	1.932	1.029	0.868	1.745
	(0.006)	(0.001)	(0.136)					

Table notes: In case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in case of the Pareto, (i) and (II) refer to the shape and scale parameters; in case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: Estimation Results for the City Size Distribution

Applying Two-Piece to Eeckhout (2009) data

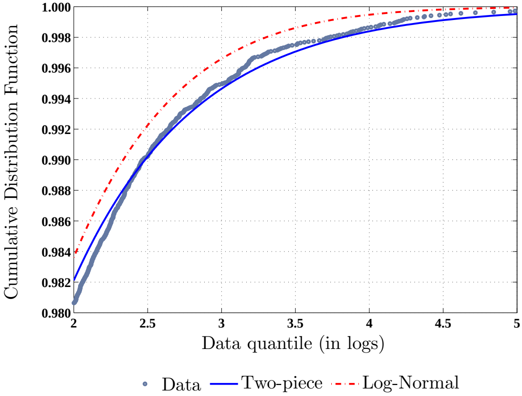

Figure: City size c.d.f: Two-piece and Log-normal vs. Data

Back to main

Pareto or Lognormal? Related literature

Why one may choose Pareto:

- Many firm-specific outcomes follow Pareto (at least in the upper tail):
  - Simon and Bonini (1958), Luttmer (2007), Axtell (2001), Gabaix (2008), Levchenko and di Giovanni (2012).
- By far the most popular choice in quantitative models:
  - Following Baldwin (2005) and Chaney (2008), hundreds of papers use Pareto, e.g., Melitz and Ottaviano (2008), Arkolakis, Costinot and Rodríguez-Clare (2012), Melitz and Redding (2014) and many others.
- Elegant and easy to handle analytically

Pareto or Log-normal? Related literature

Why one may choose Log-normal:

- Fits the data better on a larger interval of the support ($> 90\%$):
  - Head, Mayer and Thoenig (2014), Freund and Pierola (2015)
- Leads to non-linear trade elasticities which is supported by the data:
  - Yang (2014) and Bas, Mayer and Thoenig (2015), and Fernandes, Klenow, Meleshchuk, Pierola, and Rodríguez-Clare (2015)
- Not as elegant as Pareto but still tractable

Pareto vs. Log-normal? Theoretical literature

The paper is related to:

- Arkolakis (2015) shows how endogenous growth processes can lead to mixture distribution of productivities.
- Mrazova, Neary and Parenti (2015) show how different assumptions about the structure of demand and technology affect distribution of firms outcomes.
- Papers when the choice between Pareto and Log-normal is unclear, e.g., debate about the city size distribution Gabaix (1999), Eeckhout (2004, 2009) and Levy (2009).

Back to main

Alternative weighting

I employ an alternative weighting scheme which weights observations according to their size in the data using the following estimator:

$$\min_{\Theta_\ell} \left\{ \sum_q Q_e(q) \Big(\ln\left[Q_e(q)\right] - \ln\left[Q_\ell(q|\Theta_\ell)\right] \Big)^2 \right\}, \tag{3}$$

Giving larger weights to observations in the right tail does not alter the main results, i.e., the Two-piece distribution still dominates the alternatives according to the size of the RMSE along the entire support and especially in the right tail.

	Parameters			Root Mean Squared Error				
	(I)	(II)	(III)	All	Bot. 1%	Bot. 5%	Top 5%	Top 1%
Two-piece	3.005	1.119	0.929	0.059	0.505	0.246	0.034	0.035
	0.014	0.009	0.002					
Log-normal	0.607	-0.710		0.080	0.318	0.159	0.102	0.216
	0.002	0.000						
Pareto	2.419	0.351	0.269	1.576	1.001	0.099	0.209	
	0.013	0.001						
B. Pareto	2.012	0.320	4.447	0.236	1.487	0.917	0.123	0.183
	0.033	0.002	0.296					

Table notes: In the case of the Two-piece distribution, parameter (I) refers to the shape parameter, α, (II) and (III) refer to the scale parameters, θ and ρ, respectively; in the case of the Log-normal distribution, (I) and (II) refer to the scale and location parameters; in the case of the Pareto, (i) and (II) refer to the shape and scale parameters; in the case of the Bounded Pareto distribution, (I) refers to the shape parameter and (II) and (III) to two location parameters. All parameters are estimated using 100,000 quantile data points.

Table: Estimation Results under Alternative Weighting