Welfare and Trade Without Pareto

Keith Head, Thierry Mayer and Mathias Thoenig

Online Appendix

A1. Welfare and the share of domestic trade

Here we derive equation (2), showing welfare changes as a function of changes in the domestic share and the mass of domestic entrants. This equation resembles an un-numbered equation in Arkolakis, Costinot and Rodriguez-Clare (2012), p. 111. However, it reduces the determinants of welfare to just changes in own trade and changes in the mass of entrants. Along the way, we set up the model in general terms: C asymmetric countries, and general distribution functions, which provides equation (2) and other useful results fo the calibration.

Bilateral trade can be expressed as the product of M^e_{ni}, the mass of entrants from i into destination n, and the mean export revenues of exporters from i serving market n.

$$X_{ni} = G(\alpha^*_{ni}) M^e_i \int_0^{\alpha^*_{ni}} \frac{x_{ni}(\alpha) g(\alpha) d\alpha}{G(\alpha^*_{ni})}, \tag{A1}$$

where α^*_{ni} is the cutoff cost over which firms in i would make a loss in market n.

With demand being CES (denoted σ), equilibrium markups ($\bar{m} = \sigma/(\sigma-1)$) being constant, and trade costs (τ_{ni}) being iceberg, the export value of an individual firm with productivity $1/\alpha$ is given by

$$x_{ni}(\alpha) = (\bar{m}\alpha w_i \tau_{ni})^{1-\sigma} P_n^{\sigma-1} Y_n, \tag{A2}$$

with Y_n denoting total expenditure and P_n the price index of the CES composite.

Following Helpman, Melitz and Rubinstein (2008), it is useful to define

$$V_{ni} = \int_0^{\alpha^*_{ni}} \alpha^{1-\sigma} g_i(\alpha) d\alpha. \tag{A3}$$

Now we can re-express aggregate exports from i to n as

$$X_{ni} = M^e_i Y_n (\bar{m} w_i \tau_{ni})^{1-\sigma} P_n^{\sigma-1} V_{ni}, \qquad \text{with} \qquad P_n^{1-\sigma} \equiv \sum_\ell M^e_\ell (\bar{m} w_\ell \tau_{n\ell})^{1-\sigma} V_{n\ell}. \tag{A4}$$

Since market clearing and balanced trade imply $Y_i = w_i L_i$, we can replace w_i with Y_i/L_i. We also divide X_{ni} by Y_n to obtain the expenditure shares, π_{ni} for importer n on exporter i:

$$\pi_{ni} = M^e_i L_i^{\sigma-1} Y_i^{1-\sigma} (\bar{m}\tau_{ni})^{1-\sigma} V_{ni} P_n^{\sigma-1}, \tag{A5}$$

with

$$P_n^{1-\sigma} = \sum_\ell M^e_\ell L_\ell^{\sigma-1} Y_\ell^{1-\sigma} (\bar{m}\tau_{n\ell})^{1-\sigma} V_{n\ell}. \tag{A6}$$

Gross profits in the CES model are given by x_{ni}/σ. Hence, assuming that fixed costs are paid using labor of the origin country, the cutoff cost such that profits are zero is determined by $x_{ni}(\alpha^*) =$

$\sigma w_i f_{ni}$. Combined with $w_i = Y_i/L_i$ we obtain:

$$\alpha^*_{ni} = \sigma^{1/(1-\sigma)} \left(\frac{L_i}{Y_i}\right)^{\sigma/(\sigma-1)} \left(\frac{Y_n}{f_{ni}}\right)^{1/(\sigma-1)} \frac{P_n}{\bar{m}\tau_{ni}}. \tag{A7}$$

Welfare in this model is given by real income. Inverting equation (A7), welfare can be expressed in terms of the domestic cutoff:

$$\mathbb{W}_i \equiv \frac{Y_i}{P_i} = \left(\frac{L_i}{\sigma}\right)^{\sigma/(\sigma-1)} \frac{\sigma-1}{\tau_{ii} f_{ii}^{1/(\sigma-1)}} \frac{1}{\alpha^*_{ii}}. \tag{A8}$$

This is equation (1) in the main text. Since α^*_{ii} is the sole endogenous variable, a change in international trade costs implies that $\frac{d\mathbb{W}_i}{\mathbb{W}_i} = -\frac{d\alpha^*_{ii}}{\alpha^*_{ii}}$. The next step is to relate changes in the cutoff to changes in trade shares. To do this we divide both sides of equation (A6) by $P_n^{1-\sigma}$, and differentiate, to obtain:

$$\sum_\ell \pi_{n\ell} \left[\frac{dM^e_\ell}{M^e_\ell} + (1-\sigma)\frac{d\tau_{n\ell}}{\tau_{n\ell}} + (1-\sigma)\frac{dY_\ell}{Y_\ell} + \frac{dV_{n\ell}}{V_{n\ell}} + (\sigma-1)\frac{dP_n}{P_n}\right] = 0 \tag{A9}$$

Analyzing the dV/V term first, we can see from the definition in equation (A3) that it is the product of the elasticity of V with respect to the cutoff times the percent change in the cutoff. We follow ACR in denoting the first elasticity as γ; it is given by

$$\gamma_{ni} \equiv \frac{d\ln V_{ni}}{d\ln \alpha^*_{ni}} = \frac{\alpha^{*2-\sigma}_{ni} g(\alpha^*_{ni})}{\int_0^{\alpha^*_{ni}} \alpha^{1-\sigma} g(\alpha) d\alpha}. \tag{A10}$$

From the definition of V and equilibrium cutoffs in (A7), we can write the change in V as

$$\frac{dV_{n\ell}}{V_{n\ell}} = \gamma_{n\ell}\frac{d\alpha^*_{n\ell}}{\alpha^*_{n\ell}} = \gamma_{n\ell}\left[\frac{1}{\sigma-1}\frac{dY_n}{Y_n} - \frac{\sigma}{\sigma-1}\frac{dY_\ell}{Y_\ell} + \frac{dP_n}{P_n} - \frac{d\tau_{n\ell}}{\tau_{n\ell}}\right]. \tag{A11}$$

Combining (A9) and (A11) leads to

$$\sum_\ell \pi_{n\ell}\left[\frac{dM^e_\ell}{M^e_\ell} + (1-\sigma-\gamma_{n\ell})\left(\frac{d\tau_{n\ell}}{\tau_{n\ell}} - \frac{dP_n}{P_n}\right) + \left(1-\sigma-\frac{\sigma\gamma_{n\ell}}{\sigma-1}\right)\frac{dY_\ell}{Y_\ell} + \frac{\gamma_{n\ell}}{\sigma-1}\frac{dY_n}{Y_n}\right] = 0 \tag{A12}$$

Differentiating bilateral trade shares in equation (A5),

$$\frac{d\pi_{n\ell}}{\pi_{n\ell}} = \frac{dM^e_\ell}{M^e_\ell} + (1-\sigma)\frac{d\tau_{n\ell}}{\tau_{n\ell}} + (1-\sigma)\frac{dY_\ell}{Y_\ell} + \frac{dV_{n\ell}}{V_{n\ell}} + (\sigma-1)\frac{dP_n}{P_n}, \tag{A13}$$

$$\frac{d\pi_{nn}}{\pi_{nn}} = \frac{dM^e_n}{M^e_n} + (1-\sigma)\frac{dY_n}{Y_n} + \frac{dV_{nn}}{V_{nn}} + (\sigma-1)\frac{dP_n}{P_n}. \tag{A14}$$

Hence, the difference in those share changes gives

$$\frac{d\pi_{n\ell}}{\pi_{n\ell}} - \frac{d\pi_{nn}}{\pi_{nn}} + \frac{dM^e_n}{M^e_n} = \frac{dM^e_\ell}{M^e_\ell} + (1-\sigma)\frac{d\tau_{n\ell}}{\tau_{n\ell}} + (1-\sigma)\left[\frac{dY_\ell}{Y_\ell} - \frac{dY_n}{Y_n}\right] + \frac{dV_{n\ell}}{V_{n\ell}} - \frac{dV_{nn}}{V_{nn}}. \tag{A15}$$

Let us focus now in the difference in V term. From (A11), we can write:

$$\begin{aligned}\frac{dV_{n\ell}}{V_{n\ell}}-\frac{dV_{nn}}{V_{nn}}&=\gamma_{n\ell}\frac{d\alpha^*_{n\ell}}{\alpha^*_{n\ell}}-\gamma_{nn}\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}\\&=\gamma_{n\ell}\left[\frac{1}{\sigma-1}\frac{dY_n}{Y_n}-\frac{\sigma}{\sigma-1}\frac{dY_\ell}{Y_\ell}-\frac{d\tau_{n\ell}}{\tau_{n\ell}}+\frac{dP_n}{P_n}\right]\\&\quad-\gamma_{nn}\left[-\frac{dY_n}{Y_n}+\frac{dP_n}{P_n}\right].\\&=(\gamma_{n\ell}-\gamma_{nn})\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}+\gamma_{n\ell}\left[\frac{\sigma}{\sigma-1}\left(\frac{dY_n}{Y_n}-\frac{dY_\ell}{Y_\ell}\right)-\frac{d\tau_{n\ell}}{\tau_{n\ell}}\right].\end{aligned}\tag{A16}$$

We then plug (A16) into (A15) to obtain

$$\begin{aligned}\frac{d\pi_{n\ell}}{\pi_{n\ell}}-\frac{d\pi_{nn}}{\pi_{nn}}+\frac{dM^e_n}{M^e_n}-(\gamma_{n\ell}-\gamma_{nn})\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}=&\frac{dM^e_\ell}{M^e_\ell}+(1-\sigma-\gamma_{n\ell})\frac{d\tau_{n\ell}}{\tau_{n\ell}}\\&+\left(1-\sigma-\frac{\sigma\gamma_{n\ell}}{\sigma-1}\right)\left[\frac{dY_\ell}{Y_\ell}-\frac{dY_n}{Y_n}\right].\end{aligned}\tag{A17}$$

Therefore the term in square brackets inside (A12) is equal to

$$\frac{d\pi_{n\ell}}{\pi_{n\ell}}-\frac{d\pi_{nn}}{\pi_{nn}}+\frac{dM^e_n}{M^e_n}-(\gamma_{n\ell}-\gamma_{nn})\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}+(1-\sigma-\gamma_{n\ell})\left[\frac{dY_n}{Y_n}-\frac{dP_n}{P_n}\right].\tag{A18}$$

After replacing $\frac{dY_n}{Y_n}-\frac{dP_n}{P_n}=-\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}$, and canceling out the terms involving $\gamma_{n\ell}$, we can substitute the result into (A12) to obtain

$$\sum_\ell\pi_{n\ell}\left[\frac{d\pi_{n\ell}}{\pi_{n\ell}}-\frac{d\pi_{nn}}{\pi_{nn}}+\frac{dM^e_n}{M^e_n}+(\sigma-1+\gamma_{nn})\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}\right]=0\tag{A19}$$

Noting that only $d\pi_{n\ell}/\pi_{n\ell}$ terms depend on ℓ we can re-arrange as

$$-(\sigma-1+\gamma_{nn})\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}=-\frac{d\pi_{nn}}{\pi_{nn}}+\frac{dM^e_n}{M^e_n}+\sum_\ell\pi_{n\ell}\frac{d\pi_{n\ell}}{\pi_{n\ell}}\tag{A20}$$

Using $\sum_\ell\pi_{n\ell}\frac{d\pi_{n\ell}}{\pi_{n\ell}}=0$, we can finally express the welfare change as

$$\frac{d\mathbb{W}_n}{\mathbb{W}_n}=-\frac{d\alpha^*_{nn}}{\alpha^*_{nn}}=\frac{-d\pi_{nn}/\pi_{nn}+dM^e_n/M^e_n}{(\sigma-1+\gamma_{nn})},\tag{A21}$$

which after defining $\epsilon_{nn}=1-\sigma-\gamma_{nn}$, is equation (2) in the text.

*A2. How **M1** (entry share) affects welfare in the symmetric model*

Under the trading regime, our micro-data calibration procedure is characterized by the two equilibrium relationships (3) and (4), the two moment conditions $\mathbf{M1}-G(\alpha^*_d)=0$ and $\mathbf{M2}-G(\alpha^*_x)/G(\alpha^*_d)=0$, and four unknowns $(\alpha^*_d,\alpha^*_x;f^E,f_x)$.

Differentiating the two moment conditions with respect to **M1** we obtain

$$\frac{d\alpha^*_d/\alpha^*_d}{d\mathbf{M1}/\mathbf{M1}}=\frac{G(\alpha^*_d)}{\alpha^*_dG'(\alpha^*_d)}>0,\tag{A22}$$

$$\text{(A23)} \qquad \frac{d\alpha_x^*/\alpha_x^*}{d\mathbf{M1}/\mathbf{M1}} = \frac{G(\alpha_x^*)}{\alpha_x^* G'(\alpha_x^*)} > 0,$$

Simple manipulations of the differentiated system also yields

$$\text{(A24)} \qquad \frac{df_x}{f_x} = (\sigma - 1) \times \left[\frac{G(\alpha_d^*)}{\alpha_d^* G'(\alpha_d^*)} - \frac{G(\alpha_x^*)}{\alpha_x^* G'(\alpha_x^*)} \right] \times \frac{d\mathbf{M1}}{\mathbf{M1}},$$

$$\text{(A25)} \qquad \frac{df^E}{f^E} = A_1^+ \frac{d\alpha_d^*}{\alpha_d^*} + A_2^+ \frac{d\alpha_x^*}{\alpha_x^*} + A_3^+ \frac{df_x}{f_x},$$

where (A_1^+, A_2^+, A_3^+) are positive parameters. Looking at the Pareto version of definition (5), it is clear that $\frac{G(\alpha_d^*)}{\alpha_d^* G'(\alpha_d^*)} - \frac{G(\alpha_x^*)}{\alpha_x^* G'(\alpha_x^*)} = 0$, which means that the right hand side of (A24) is zero under Pareto. Therefore, a change of **M1** is i) not related to changes in f_x, ii) affecting all cutoffs in the same way, leaving export propensity, but also gains from trade unaffected. Under log-normal on the contrary, $\frac{G(\alpha_d^*)}{\alpha_d^* G'(\alpha_d^*)} - \frac{G(\alpha_x^*)}{\alpha_x^* G'(\alpha_x^*)} > 0$ (see (5)). Hence in the LN case,

$$\text{(A26)} \qquad \frac{df_x/f_x}{d\mathbf{M1}/\mathbf{M1}} \geqslant 0.$$

Combined with (A22), equations (A23), (A24) and (A25) thus imply that

$$\text{(A27)} \qquad \frac{df^E/f^E}{d\mathbf{M1}/\mathbf{M1}} > 0.$$

Let us consider now the domestic cutoff in autarky, characterized by $G(\alpha_{dA}^*)\left[\mathcal{H}(\alpha_{dA}^*) - 1\right] = f^E$. Differentiating this relationship we get

$$\text{(A28)} \qquad \frac{d\alpha_{dA}^*/\alpha_{dA}^*}{d\mathbf{M1}/\mathbf{M1}} > 0$$

We conclude from the previous computations that an increase in **M1** leads to an increase in both α_d^* and α_{dA}^*, namely a less selective domestic market both in autarky and in the trading equilibrium.

The change in trade gains is equal to

$$\text{(A29)} \qquad \frac{d\mathbb{T}}{\mathbb{T}} = \left[\frac{d\alpha_{dA}^*}{\alpha_{dA}^*} - \frac{d\alpha_d^*}{\alpha_d^*} \right] \times \frac{d\mathbf{M1}}{\mathbf{M1}}$$

The sign of the previous relationship cannot be characterized algebraically and we consequently rely on our quantitative procedure to show that it is positive under log-normal.

A3. *Distribution parameters for Chinese exports to Japan*

Table A1 replicates Table 1 for the case of Chinese exports to Japan in 2000.

A4. *Distributions of total sales*

Some of the prior literature asserting Pareto is based on firm size distribution, rather than looking at the distribution of export sales from one origin in a particular importing country (which is also done in Eaton et al. (2011)).

The mapping between productivity distribution parameters and sales distributions is less clear when considering total sales of firms (domestic sales plus exports to all destinations). Chaney

TABLE A1—PARETO VS LOG-NORMAL: QQ REGRESSIONS (CHINESE EXPORTS TO JAPAN IN 2000).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	all	top 50%	top 25%	top 5%	top 4%	top 3%	top 2%	top 1%
Obs:	24832	12416	6208	1241	993	745	496	248
	Log-normal: RHS $= \Phi^{-1}(\hat{F}_i)$, coeff = $\tilde{\nu}$							
$\Phi^{-1}(\hat{F}_i)$	2.558a	2.125a	1.950a	1.936a	1.934a	1.929a	1.910a	1.970a
R^2	0.986	0.995	0.999	0.998	0.998	0.997	0.995	0.992
ν	0.853	0.708	0.650	0.645	0.645	0.643	0.637	0.657
	Pareto: RHS $= -\ln(1-\hat{F}_i)$, coeff = $1/\tilde{\theta}$							
$-\ln(1-\hat{F}_i)$	2.194a	1.239a	0.946a	0.718a	0.698a	0.674a	0.640a	0.618a
R^2	0.725	0.930	0.971	0.990	0.991	0.992	0.995	0.994
θ	1.367	2.422	3.170	4.175	4.296	4.452	4.688	4.854

Notes: the dependent variable is the log exports of Chinese firms to Japan in 2000. The standard deviation of log exports in this sample is 2.576, which should be equal to $\tilde{\nu}$ if x is log-normally distributed and to $1/\tilde{\theta}$ if distribution if Pareto. ν and θ are calculated using $\sigma = 4$. Standard errors still have to be corrected.

(2013) and Di Giovanni et al. (2011) are examples using total exports and sales, respectively, for French firms. Both papers truncate the samples. In Tables A2 and A3, and figure A1, we corroborate the evidence in favor of log-normality of total sales of French and Spanish firms. We also show that the superior performance of log-normal is not driven by exports of intermediaries. For both the French and Chinese export samples, restricting to non-intermediaries yields similar results.

TABLE A2—PARETO VS LOG-NORMAL: QQ REGRESSIONS (FRENCH FIRMS TOTAL SALES IN 2000).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	all	top 50%	top 25%	top 5%	top 4%	top 3%	top 2%	top 1%
Obs:	92988	46494	23247	4649	3719	2789	1860	930
	Log-normal: RHS $= \Phi^{-1}(\hat{F}_i)$, coeff = $\tilde{\nu}$							
$\Phi^{-1}(\hat{F}_i)$	1.790a	2.076a	2.330a	2.579a	2.586a	2.603a	2.610a	2.586a
R^2	0.984	0.990	0.996	0.999	0.998	0.998	0.997	0.992
ν	0.597	0.692	0.777	0.860	0.862	0.868	0.870	0.862
	Pareto: RHS $= -\ln(1-\hat{F}_i)$, coeff = $1/\tilde{\theta}$							
$-\ln(1-\hat{F}_i)$	1.658a	1.251a	1.143a	0.955a	0.932a	0.906a	0.869a	0.806a
R^2	0.844	0.988	0.991	0.991	0.991	0.990	0.990	0.989
θ	1.809	2.398	2.624	3.140	3.220	3.312	3.452	3.723

Notes: the dependent variable is the log exports of French total sales in 2000. The standard deviation of log exports in this sample is 1.805, which should be equal to $\tilde{\nu}$ if x is log-normally distributed and to $1/\tilde{\theta}$ if distribution if Pareto. ν and θ are calculated using $\sigma = 4$. Standard errors still have to be corrected.

A5. Comparison of QQ estimator to other methods

One alternative to the QQ estimators is to use method of moments. In this case, we infer the distributional parameters from the means and standard deviations of log sales. We can use equations (6) and (7) to obtain an idea of what those coefficients should be. With log of sales distributed Normal, they have a mean value of $\tilde{\mu}$, and a standard deviation of $\tilde{\nu}$. In the Pareto case, the log of sales have a mean value of $\ln \underline{\tilde{\varphi}} + 1/\tilde{\theta}$, and a standard deviation of $1/\tilde{\theta}$. In this sample, the standard deviation of log sales is 2.393, hence predicted coefficients in Table 1 are 2.393 for Log-Normal

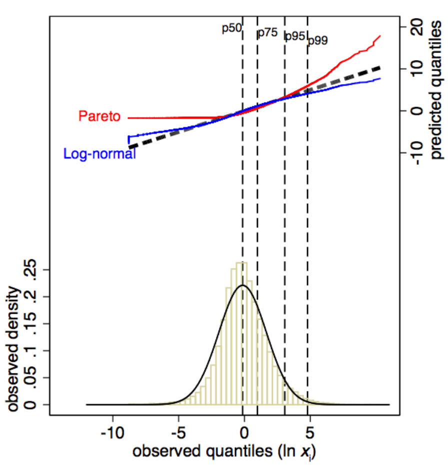
TABLE A3—PARETO VS LOG-NORMAL: QQ REGRESSIONS (SPANISH FIRMS TOTAL SALES IN 2000).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	all	top 50%	top 25%	top 5%	top 4%	top 3%	top 2%	top 1%
Obs:	87998	43999	21999	4400	3520	2640	1760	880
	Log-normal: RHS $=\Phi^{-1}(\hat{F}_i)$, coeff = $\tilde{\nu}$							
$\Phi^{-1}(\hat{F}_i)$	1.588a	1.859a	2.095a	2.419a	2.435a	2.462a	2.510a	2.599a
R^2	0.986	0.988	0.992	0.998	0.997	0.996	0.995	0.991
ν	0.529	0.620	0.698	0.806	0.812	0.821	0.837	0.866
	Pareto: RHS $=-\ln(1-\hat{F}_i)$, coeff = $1/\tilde{\theta}$							
$-\ln(1-\hat{F}_i)$	1.489a	1.122a	1.032a	0.899a	0.880a	0.861a	0.840a	0.814a
R^2	0.866	0.990	0.995	0.995	0.996	0.997	0.997	0.996
θ	2.015	2.674	2.907	3.337	3.409	3.486	3.573	3.687

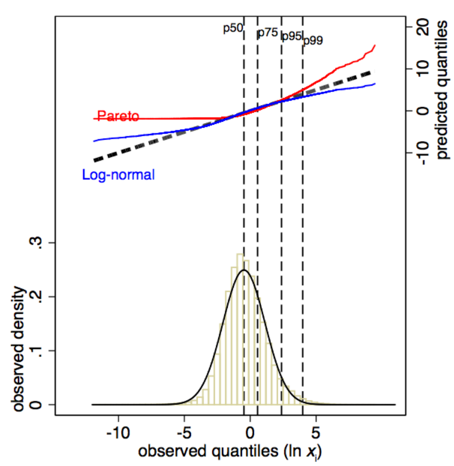
Notes: the dependent variable is the log exports of Spanish total sales in 2000. The standard deviation of log exports in this sample is 1.599, which should be equal to $\tilde{\nu}$ if x is log-normally distributed and to $1/\tilde{\theta}$ if distribution if Pareto. ν and θ are calculated using $\sigma = 4$. Standard errors still have to be corrected.

FIGURE A1. QQ GRAPHS ON TOTAL SALES

(a) French firms



(b) Spanish firms



and Pareto independently of truncation. The un-truncated sample estimate almost exactly matches that prediction for the log-normal case, when most estimates of Pareto case are quite far off.

There is a close relationship between the QQ estimator for the Pareto and the familiar log rank-size regressions examined by Gabaix and Ioannides (2004) since both rank, $1 + (n - i)$, and one minus the empirical CDF are linear in i. This closely resembles the QQ estimator since, following the suggestion of Bury (1999), we estimate the empirical CDF as $\hat{F}_i = (i - 0.3)/(n + 0.4)$. Thus, the empirical CDF is an affine transformation of the rank. The coefficient on log sales is $-\tilde{\theta} = -\frac{\theta}{\sigma-1}$. Eaton et al. (2011), Di Giovanni et al. (2011) are recent examples that pursue this approach and it is also referred to by Melitz and Redding (2013) in their parameterization of **M3**.

A6. Macro-data simulations

In this section, we adopt the $\mathbf{M3}'$ approach where the underlying micro parameters ν and θ are calibrated to match the international trade elasticity, ϵ_x. Under the Pareto distribution $\epsilon_x = \epsilon_d = -\theta$. Thus, we calibrate the Pareto heterogeneity parameter as $\theta = -\mathbf{M3}'$. Under log-normal

$$\mathbf{M3}' = 1 - \sigma - \frac{1}{\nu} h\left(\frac{\ln \alpha_x^* + \mu}{\nu} + (\sigma - 1)\nu\right),$$

where $h(x) \equiv \phi(x)/\Phi(x)$, the ratio of the PDF to the CDF of the standard normal. In this case, the calibration procedure will therefore select values for f^E, f_x and ν such that target values for $\mathbf{M1}$, $\mathbf{M2}$, and $\mathbf{M3}'$ are matched.

The most obvious empirical target value for $\mathbf{M3}'$ (recommended by Arkolakis, Costinot and Rodriguez-Clare (2012)) comes from estimates of the gravity literature regressing trade flows on bilateral applied tariffs. Head and Mayer (2014) survey this literature and report a median estimate of -5.03, which we take as our target for both Pareto and log-normal. The left panel of figure A2 plots the GFT as in figures 2 and 3, and the right panel graphs the three relevant trade elasticities: ϵ^{P} for Pareto, constant at -5.03, ϵ_x^{LN} and ϵ_d^{LN}, the international and domestic elasticities for the log-normal case. By construction, ϵ_x^{LN} coincides with Pareto at the benchmark trade cost ($\tau = 1.83$). As τ declines, the elasticity falls in absolute value. The domestic elasticity, ϵ_d^{LN}, is uniformly smaller in absolute value than ϵ_x^{LN}. It rises with increases in τ because higher international trade costs make the domestic market easier in relative terms.

Despite this large heterogeneity in trade elasticities between Pareto and log-normal, gains from trade happen to be very proximate in this symmetric country calibration. While the GFT are very similar for this set of parameters, they are not identical, as the zoomed-in box reveals. Second, they can be much more different when one changes some parameter targets, in particular the share of exporters. Third, this calibration searches for parameters in order to fit a unique trade elasticity (the international one), while the LN version of the model features *two* elasticities that depend crucially on ν. Calibrating the model to fit an average of the two trade elasticities in figure A3, the Pareto and log-normal GFT again diverge from each other.

A7. Generative processes for log-normal and Pareto

Because the Pareto distribution has been thought to characterize a large set of phenomena in both natural and social sciences, much effort has gone into developing generative models that predict the Pareto as a limiting distribution. The building block emphasized in the literature, see especially Gabaix (1999), is Gibrat's law of proportional growth. Applied to sales of an individual firm i in period t, Gibrat's Law states that $X_{i,t+1} = \Gamma_{it} X_{it}$. The key point is that the growth rate from period to period, $\Gamma_{it} - 1$ is independent of size. A confusion has arisen because it is straightforward to show that the law of proportional growth delivers a log-normal distribution. In

FIGURE A2. WELFARE GAINS CALIBRATED ON TRADE ELASTICITY

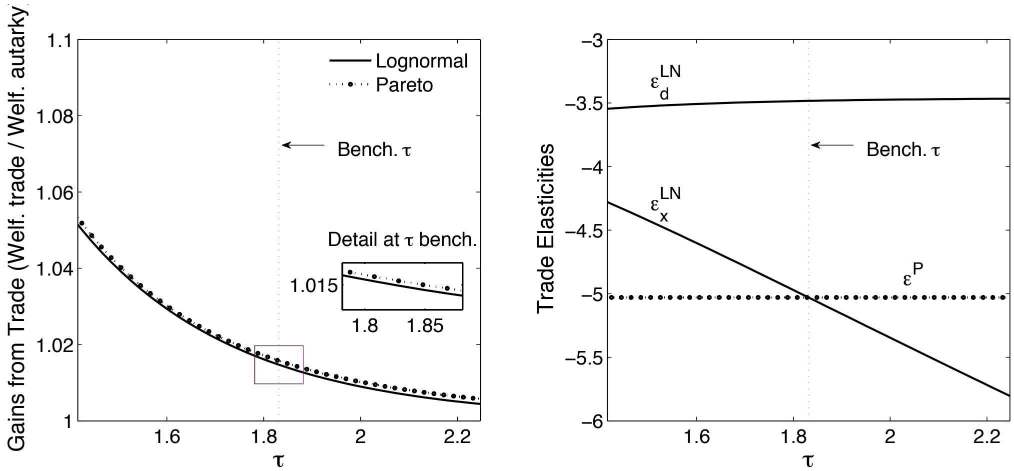
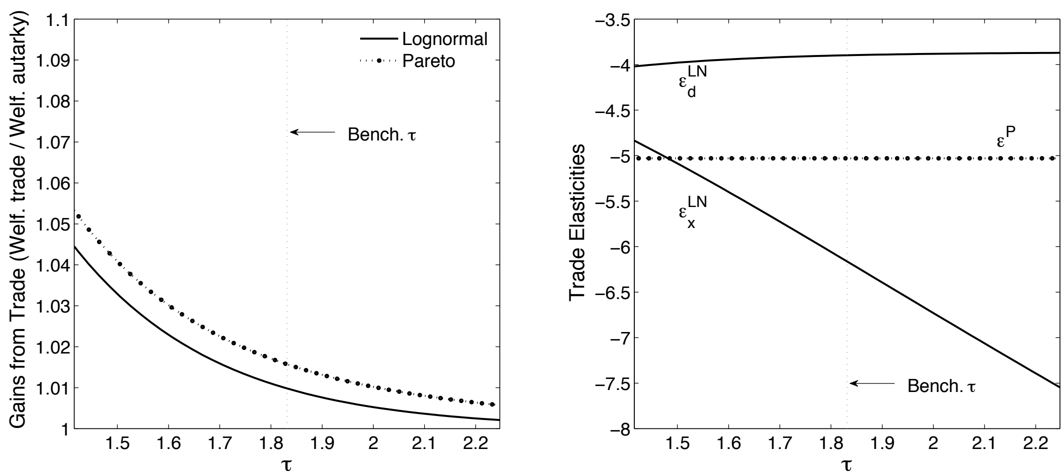


FIGURE A3. WELFARE GAINS CALIBRATED ON AVERAGE TRADE ELASTICITY



period T size is given by

$$X_{iT} = \exp(\ln X_{i0} + \sum_{t=1}^{T} \ln \Gamma_{it})$$

The central limit theorem implies for large T,

$$\sqrt{T}\left(\frac{\sum_t \ln \Gamma_{it}}{T} - \mathbb{E}[\ln \Gamma_{it}]\right) \sim \mathcal{N}(0, \mathbb{V}[\ln \Gamma_{it}]),$$

where $\mathbb{E}$ and $\mathbb{V}$ are the expectation and variance operators. Rearranging and, for convenience only, initializing sizes at $X_{i0} = 1$, $\ln X_{it}$ is normally distributed with expectation $T\mathbb{E}[\ln \Gamma_{it}]$ and variance $T\mathbb{V}[\ln \Gamma_{it}]$. This implies X_{iT} is log-normal with log-mean parameter $\tilde{\mu} = T\mathbb{E}[\ln \Gamma_{it}]$ and log-SD parameter $\tilde{\nu} = \sqrt{T\mathbb{V}[\ln \Gamma_{it}]}$.

This demonstration that Gibrat's Law implies a limiting distribution that is log-normal echoes similar arguments by Sutton (1997) for firms and Eeckhout (2004) for cities. The problem with this formulation is that it is only valid for large T and yet as T grows large, the distribution exhibits some perverse behavior. Assume that sizes are not growing on average, i.e. $\mathbb{E}[\Gamma_{it}] = 1$. By Jensen's Inequality, $\mathbb{E}[\ln \Gamma_{it}] < \ln(\mathbb{E}[\Gamma_{it}]) = 0$. Since the median of X_{iT} is $\exp(\tilde{\mu}) = \exp(T\mathbb{E}[\ln \Gamma_{it}])$, the median should decline exponentially with time. The mode, $\exp(\tilde{\mu} - \tilde{\nu}^2) = \exp[T(\mathbb{E}[\ln \Gamma_{it}] - \mathbb{V}[\ln \Gamma_{it}])]$ should decline even more rapidly with time. Thus, as T becomes large, Gibrat's law with $\mathbb{E}[\Gamma_{it}] = 1$ implies a distribution with a mode going to zero while the variance is becoming infinite. Evidently something must be done to rescue Gibrat's law from generating degeneracy.

A variety of modifications to Gibrat's Law have been investigated. Kalecki (1945) specifies growth shocks that are negatively correlated with the level. This allows for a log-normal with stable variance to emerge. Gabaix (1999) shows in an appendix that a simple change to the growth process, $X_{i,t+1} = \Gamma_{it} X_{it} + \varepsilon$ with $\varepsilon > 0$ (the Kesten process) is enough to solve the problem of degeneracy. But the resulting stable distribution is Pareto, not log-normal. Reed (2001) instead assumes finite-lived agents with exponential life expectancies. This leads to a double-Pareto distribution.

*

Appendix References

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012). New trade models, same old gains? *American Economic Review 102*(1), 94–130.

Bury, K. (1999). *Statistical distributions in engineering*. Cambridge University Press.

Chaney, T. (2013). The gravity equation in international trade: An explanation. Working Paper 19285, National Bureau of Economic Research.

Di Giovanni, J., A. A. Levchenko, and R. Ranciere (2011). Power laws in firm size and openness to trade: Measurement and implications. *Journal of International Economics 85*(1), 42–52.

Eaton, J., S. Kortum, and F. Kramarz (2011). An anatomy of international trade: Evidence from French firms. *Econometrica 79*(5), 1453–1498.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, 1429–1451.

Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly Journal of Economics 114*(3), 739–767.

Gabaix, X. and Y. M. Ioannides (2004). The evolution of city size distributions. *Handbook of regional and urban economics 4*, 2341–2378.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In E. Helpman, G. Gopinath, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4. Elsevier.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics 123*(2), 441–487.

Kalecki, M. (1945). On the Gibrat distribution. *Econometrica: Journal of the Econometric Society*, 161–170.

Melitz, M. J. and S. J. Redding (2013). Firm heterogeneity and aggregate welfare. Working Paper 18919, National Bureau of Economic Research.

Reed, W. J. (2001). The Pareto, Zipf and other power laws. *Economics Letters 74*(1), 15–19.

Sutton, J. (1997). Gibrat's legacy. *Journal of economic Literature 35*(1), 40–59.