Online Appendix for

"Economies of Scale and the Size of Exporters"

Roc Armenter

Federal Reserve Bank of Philadelphia

Miklós Koren

Central European University, MTA KRTK and CEPR

March 1, 2014

A Appendix to Section 2

A.1 Measurement error

There may have been some measurement error in the data, with some actual exporters coded as non-exporters, and the other way around. We explore here several possibilities regarding misclassification. Let p be the probability that an exporter was recorded as a non-exporter, that is, $\Pr\left(r_i^f > 0 | \hat{r}_i^f = 0\right)$, where variables with a hat denote the data and variables without a hat denote the actual value; and let q be the probability a non-exporter was recorded as an exporter, or $\Pr\left(r_i^f = 0 | \hat{r}_i^f > 0\right)$. Other than in relationship to the exporter status of a firm, we assume that equivocation errors are independent of size.

Given values for p and q, we can recover the actual share of exporters and exporter size premium from their recorded values. Let s_x be the true share of exporters, with μ_x and μ_{nx} the true average sales of exporters and non-exporters, respectively; and let $\hat{s}_x$ be the share of exporters as observed in the data, and $\hat{\mu}_x$ and $\hat{\mu}_{nx}$ the average sales of exporters and non-exporters, respectively, again as measured in the data.

The true share of exporters relates to the observed share of exporters as

$$\hat{s}_x = (1-p)s_x + q(1-s_x),$$

that is, observed exporters are made up of the correctly-coded exporters plus the misclassified non-exporters. Solving for s_x we obtain

$$s_x = \frac{\hat{s}_x - q}{1-p-q}.$$

Note that the observed share of exporters limits the equivocation error q. For the exporter size premium, the observed average sales for exporter and non-exporters relate to their actual values according to:

$$\hat{\mu}_x = \frac{(1-p)s_x}{\hat{s}_x}\mu_x + \frac{q(1-s_x)}{\hat{s}_x}\mu_{nx}, \tag{1}$$

$$\hat{\mu}_{nx} = \frac{ps_x}{1-\hat{s}_x}\mu_x + \frac{(1-q)(1-s_x)}{1-\hat{s}_x}\mu_{nx}. \tag{2}$$

Inverting the system above we recover the true average sales of exporter and non-exporters, respectively.

In Table 1 we report the true share and size premium of exporters for a wide range of error rates. The last column reports the model's predicted size premium given the true share of exporters implied by the error rate.[1] The first row, in boldface, are the baseline values in the main text, which assume there was no error in recording the export status of firms.

As reported in Table 1, measurement error may have biased the estimated exporter size premium downward. However, the corrected values are nowhere close to the model's

[1] Recall the share of exporters is all is needed to obtain a predicted size premium in the strict sorting model.

Measurement errors		Share exporters	Exporter size premium	
$X \to NX$	$NX \to X$	s_x	Data	Model
$p = .00$				
	$q = .00$	**.18**	**4.5**	**95**
	$q = .01$	.17	4.6	97
	$q = .05$	.14	5.6	105
	$q = .10$	.09	8.1	116
$p = .01$				
	$q = .00$	.18	4.5	95
	$q = .01$	.17	4.7	97
	$q = .05$	.14	5.6	105
	$q = .10$	.09	8.1	116
$p = .05$				
	$q = .00$	.19	4.7	93
	$q = .01$	.18	4.8	95
	$q = .05$	.14	5.8	105
	$q = .10$	.09	8.4	116
$p = .1$				
	$q = .00$	.20	4.9	92
	$q = .01$	.19	5.1	93
	$q = .05$	.15	6.1	102
	$q = .10$	.10	8.8	115
$p = .2$				
	$q = .00$	.22	5.6	91
	$q = .01$	.21	5.8	92
	$q = .05$	.17	7.0	97
	$q = .10$	.11	10.1	114

Table 1: **Implications of measurement error in data and model**

implications. Assuming 10% and 20% error rates of non-exporters and exporters, respectively, bring the actual exporter size premium to just 10—and these are very high error rates.

It should also be noted that, unless actual exporters were substantially more likely to be misclassified than actual non-exporters, the true share of exporters will lie below the observed share. For example, for symmetric error rates $p = q$, the actual share of exporters would .17, .14, and .10 for error rates of 1%, 5%, and 10% respectively. The reason is that there are many more non-exporters than exporters in the data, so a higher (symmetric) error rate decreases the estimate of the true share of exporters. As this happens, the correspondent predicted exporter size premium in the model *increases*, so the gap between data and model, if anything, widens further.

A.2 Pareto distribution

When firm sales follow a Pareto distribution, we can obtain a simple analytic formula for the exporter size premium. In this case, we have that the cutoff t is given by $t = (s_x)^{-1/k}$ where k is the slope parameter of the Pareto distribution. The unconditional average firm size μ can be decomposed in the conditional averages

$$\mu = s_x \mu_x + (1 - s_x)\mu_{nx}$$

where μ_x and μ_{nx} are the average total sales for exporter and non-exporters respectively. For a Pareto distribution, μ_x/μ is equal to the cutoff t and we have that

$$\frac{\mu_{nx}}{\mu} = \frac{1 - (s_x)^{1-1/k}}{1 - s_x}.$$

With some algebra, we obtain an exporter size premium equal to

$$\frac{\mu_x}{\mu_{nx}} = \frac{1 - s_x}{s_x} \left(\frac{s_x^{1-1/k}}{1 - s_x^{1-1/k}} \right).$$

The exporter size premium is thus pinned down by the share of exporter and the slope parameter alone.

B Appendix to Section 4

B.1 Main framework

There is a set Ω of firms that produce and sell in the home country. Firms are heterogeneous in their productivity, denoted φ, and the fixed costs they face if they start exporting, denoted f. Productivity and fixed costs are independently distributed over $\Re_+$ with c.d.f. G and H, respectively. We summarize a firm by its type $\omega = \{\varphi, f\}$. The set of firms Ω (and their distribution) is taken as a given.

Each firm is the single producer of a differentiated good with technology

$$y(\omega) = \varphi(\omega)\, l(\omega)$$

where $l(\omega)$ is the labor demanded by firm ω. Consumers in the home country aggregate the differentiated goods according to

$$Y^d = \left[\int_\Omega \left(y^d(\omega)\right)^\rho d\omega\right]^{1/\rho}$$

where $\rho \in (0,1)$ and $y^d(\omega)$ denote the output of firm ω sold in the home country. The demand for each good ω is given by

$$y^d(\omega) = \left(p^d(\omega)/P^d\right)^{-\theta} Y^d$$

where $\theta = (1-\rho)^{-1}$ is the price elasticity, $p^d(\omega)$ the price set by firm ω, and the price index P^d is given by

$$P^d = \left[\int_\Omega \left(p^d(\omega)\right)^{1-\theta} d\omega\right]^{\frac{1}{1-\theta}}.$$

Firms are monopolistic competitors and internalize the downward sloping demand for their products. The profit-maximization problem leads to the familiar price equation

$$p^d(\omega) = \frac{1}{\rho}\frac{w}{\varphi(\omega)}$$

where w is the wage rate. We take the wage rate as exogenously given, so ours is a partial equilibrium model.

It is clear that only the productivity parameter φ will determine domestic sales. We can thus write $p^d(\varphi)$ and $y^d(\varphi)$. The c.e.s. demand structure also implies that firm φ revenues from domestic sales can be expressed

$$r^d(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}}\right)^{\theta-1} R^d \tag{3}$$

where R^d are total sales revenues in the domestic market, and

$$\tilde{\varphi} = \left[\int_0^\infty \varphi^{\theta-1} dG(\varphi)\right]^{\frac{1}{\theta-1}}$$

is the average productivity defined as in Hopenhayn (1992) and Melitz (2003). Since we take both the wage and the distribution of firms in the home country as given, total domestic sales R^d are also exogenously determined.[2] We will still make use of the relationship between productivity and domestic sales given by (3).

We now move to the determination of exports and exporters. Not all firms export: let Ω_x denote the set of firms that do and M_x its measure. We normalize the measure of all firms to one, so M_x is also the share of exporters. Consumers in the foreign country aggregate the subset of exported goods according to

$$Y^f = \left[\int_{\Omega_x} \left(y^f(\omega)\right)^\rho d\omega\right]^{1/\rho} \tag{4}$$

[2]Briefly, $R^d = r^d(\tilde{\varphi})$ by (3), and $P^d = w/(\tilde{\varphi}\rho)$ by substituting the price for each good, $p^d(\varphi)$. We do not model the import decision of the domestic households.

where $y^f(\omega)$ is the output of firm ω sold in the foreign country. The foreign demand for exporter good ω is given

$$y^f(\omega) = \left(p^f(\omega)/P^f\right)^{-\theta} Y^f$$

where

$$P^f = \left[\int_{\Omega_x} \left(p^f(\omega)\right)^{1-\theta} d\omega\right]^{\frac{1}{1-\theta}}. \tag{5}$$

Finally, we assume there is an aggregate demand for exports, given by

$$Y^f = Y^* \left(P^f\right)^{-\nu}, \tag{6}$$

where Y^* is the (exogenously given) income of the foreign country, and ν is the price elasticity of aggregate exports of the home country. We assume that $\nu < \theta$; that is, exports of the home country are closer substitutes of each other than they are of a good produced elsewhere.

Let us first solve for the export revenues of a firm, taking as given the set of exporters Ω_x. Profit maximization implies that

$$p^f(\omega) = \frac{1}{\rho} \frac{\tau w}{\varphi(\omega)}$$

where $\tau > 1$ is an iceberg trade cost associated with exports. It is clear again that, conditional on exporting, only the productivity parameter φ determines sales.

The c.e.s demand system allows us to write a firm's export revenues as a function of the average export revenues within exporters,

$$r^f(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}_x}\right)^{\theta-1} \frac{R^f}{M_x} \tag{7}$$

where R^f is total export sales and

$$\tilde{\varphi}_x = \left[\frac{1}{M_x} \int_{\Omega_x} (\varphi(\omega))^{\theta-1} d\omega \right]^{\frac{1}{\theta-1}}$$

is the average productivity among exporters. Note that the set of exporters Ω_x affects the export revenues of each firm (7), both through the share of exporters M_x and the productivity distribution within the set.

Last but most certainly not least, we get to the determination of the set of exporters Ω_x. A firm that exports incurs in a per period fixed cost. As a result, a firm ω will find it profitable to export only if its net income abroad would cover the fixed expenses,

$$\frac{1}{\theta} r^f (\varphi(\omega)) \geq f(\omega), \tag{8}$$

where export net income, that is, export revenues minus costs, is expressed as $r^f (\varphi(\omega)) / \theta$. Thus the set of exporters Ω_x is the set of firms $\omega \in \Omega$ such that the entry condition (8) holds. However, export revenues depend themselves on the set of exporters, so in equilibrium exports and exporters are determined simultaneously.

B.2 Three analytic models

While our analysis is based on the two models just described, we find it useful to consider three simpler models for comparison purposes: a representative firm model, a homogeneous firm model, and a heterogeneous-productivity model with a Pareto distribution. In these three models we can derive the relationship between total exports and trade costs in closed form. This helps build an intuition about the mechanisms at play in the richer full model.

Representative firm model. There is a representative firm that exports under all circumstances. In terms of the framework above, we encompass the representative firm model by having a degenerate distribution of firm productivity and a fixed cost equal to zero.

There are, thus, no changes in the measure of firms exporting or the average productivity of exporters. We express the price of exports (5) in terms of logs,

$$\log(P_f) = \log(\tau) + const.$$

where all terms that are constant are collected in *const*. Substituting in the aggregate demand for exports (6) we have the simplest model of exports,

$$\log(R_f) = (1 - \nu)\log(\tau) + const. \tag{9}$$

Homogeneous firm model. We introduce entry by assuming a positive fixed cost but abstract from all heterogeneity: all firms have identical labor productivity and face the same fixed cost of exporting. Some firms, though, export, while some others do not. This can be an equilibrium only if firms are actually indifferent between exporting or not, so the entry condition (8) holds with equality. Since there is no firm heterogeneity, the entry condition (8) effectively pins down the average export sales,

$$\log(R_f/M_x) = const.$$

Thus all of the adjustment in this model occurs through entry, since firms do not change their export intensity in response to trade costs.

Since all firms have identical productivity, there is no change in the average productivity of exporters either. The export price (5) is thus a function of the trade costs and the share of exporters,

$$\log(P_f) = \frac{-1}{\theta - 1}\log(M_x) + \log(\tau) + const.$$

We substitute the export price in the aggregate demand for exports (6) to obtain

$$\log(R_f) = \frac{\nu - 1}{\theta - 1}\log(M_x) + (1-\nu)\log(\tau) + const.$$

Finally, we use that $\log(R_f) = \log(M_x) + const$ to obtain total export revenues as a function of the trade cost,

$$\log(R_f) = -\frac{(\theta-1)(\nu-1)}{\theta-\nu}\log(\tau) + const. \tag{10}$$

The assumption that $\theta > \nu$ guarantees that the export revenues increase with a fall in trade costs, as we would expect

Heterogeneous-productivity firms with a Pareto distribution. The last of our models features heterogeneous firms and an identical fixed cost of exporting. The only difference with the strict sorting model discussed above is that productivity follows a Pareto distribution. This allows us to solve for the model analytically but has stark implications for the margin of adjustment, as will be clear very soon.

We now have to simultaneously solve for the export demand (6) and the entry condition (8). We start with the latter. Assuming that there is a subset of firms that do not export, the entry condition must hold with equality.[3] The firm's export revenues are strictly increasing in productivity, φ, so we can characterize the set of exporters with a simple threshold rule. Let $\hat{\varphi}_x$ hold

$$\frac{1}{\theta}r^f(\hat{\varphi}_x) = f.$$

The set of exporters Ω_x is then given by the set of firms $\varphi \geq \hat{\varphi}_x$ and thus $M_x = 1 - G(\hat{\varphi}_x)$. Using the c.d.f. for the Pareto distribution we have that

$$\log(M_x) = -k\log(\hat{\varphi}_x) \tag{11}$$

[3] Since the support for the Pareto distribution is unbounded, there is always a firm that finds it profitable to export.

where k is the slope parameter associated with the Pareto distribution.

We now get to use the key property of the Pareto distribution that allows for an analytical solution. It is easy to show that if φ is distributed according to a Pareto distribution with slope k, then we have a linear relationship between $\tilde{\varphi}_x$ and $\hat{\varphi}_x$,

$$\tilde{\varphi}_x = \left[\int_{\varphi \geq \hat{\varphi}_x} \varphi^{\theta-1} \frac{dG(\varphi)}{1-G(\hat{\varphi}_x)}\right]^{\frac{1}{\theta-1}} = \left[\frac{k}{k+1-\theta}\right] \hat{\varphi}_x.$$

We write the export revenues of firm $\hat{\varphi}_x$ in terms of the average export revenues using (7),

$$\left(\frac{\hat{\varphi}_x}{\tilde{\varphi}_x}\right)^{\theta-1} \frac{R_f}{M_x} = \theta f$$

and it is immediate that the average export sales are pinned down by the entry condition,

$$\log\left(R_f/M_x\right) = const.$$

Thus average export sales do not respond to trade costs, so the growth rate in total revenues equals the growth rate in the number of exporters.[4] This stark implication does not hold for other productivity distributions.

We are set to solve the model. The export price (5) is

$$\log\left(P_f\right) = \frac{-1}{\theta-1} \log\left(M_x\right) + \log\left(\tau\right) - \log\left(\tilde{\varphi}_x\right) + const.$$

Using the relationships between $\hat{\varphi}_x$ and $\tilde{\varphi}_x$ as well as (11),

$$\log\left(P_f\right) = \left[-\frac{1}{\theta-1} + \frac{1}{k}\right] \log\left(M_x\right) + \log\left(\tau\right) + const.$$

[4]Note that the export sales per firm increase with a fall in trade costs; however, new exporters are smaller than incumbents and drive the average down.

Now we substitute in the aggregate demand for exports to obtain

$$\log\left(R_f\right) = -\left[\frac{1}{\nu-1} - \frac{1}{\theta-1} + \frac{1}{k}\right]^{-1} \log\left(\tau\right) + const \tag{12}$$

where the average exports are constant.

C Appendix to Section 6

C.1 Sunk costs

Consider the following variation of the latent-heterogeneity model in Section 3. Firm i's total sales at date d, denoted r_{id}, are an *iid* random variable, with distribution Ψ_r. Since we do not evaluate the model at any frequency, the lack of any persistence is not particularly worrisome and allows for an easy characterization. There are two thresholds, t_0 and t_1 with $t_0 \leq t_1$, that determine the transition in and out of the foreign market as follows:

- An exporter at date $d-1$ exports at date d if $r_{id} \geq t_0$.
- A non-exporter at date $d-1$ exports at date d if $r_{id} \geq t_1$.

All readers familiar with sunk cost models will recognize the thresholds t_0 and t_1 as the stopper and starter points.

We now briefly show how we map the observations on the share and size of exporters, given the distribution of total sales, to pin down the stopper and starter points. The first equation to use is the steady-state condition on the share of exporters. Clearly, a fraction $\Psi_r\left(t_0\right)$ of previous exporters exit the exporter market, while a fraction $1-\Psi_r\left(t_1\right)$ of the previously non-exporters start exporting now. If the share of exporters is constant, we must have that

$$s_x = \left(1-\Psi_r\left(t_0\right)\right) s_x + \left(1-\Psi_r\left(t_1\right)\right)\left(1-s_x\right). \tag{13}$$

With some manipulation we have that equation (13) gives the share of exporters as a function of the starter and stopper points,

$$s_x = \frac{1 - \Psi_r(t_1)}{1 - \Psi_r(t_1) + \Psi_r(t_0)}. \tag{14}$$

Note the measure of exporters between t_0 and t_1 is simply $(\Psi_r(t_1) - \Psi_r(t_0))s_x$, so their proportion among exporters themselves is just $\Psi_r(t_1) - \Psi_r(t_0)$. We can then compute the average total sales by exporters as

$$\mu_x = (\Psi_r(t_1) - \Psi_r(t_0)) E\{r|r \in [t_0, t_1]\} + (1 - \Psi_r(t_1) + \Psi_r(t_0)) E\{r|r \geq t_1\}. \tag{15}$$

We then solve for the two equations (13) and (15), with t_0 and t_1 as the two unknowns.

We have no problem finding values for t_0 and t_1 that match the share and size of exporters. As simple as the set-up is, we can actually compute the entry and exit rates. We find that both are quite small (2.5% and 11.5%, respectively) indicating that there is a lot of persistence in the export status of the firm. Unfortunately, it is difficult to compare the numbers with the data, since we did not specify at what frequency the model is operating. Given our assumption that total sales are *iid*, we should take a period to be at very least five years. At that frequency the exit rate is very low compared with the data, as Bernard and Jensen (1999) report an annual stopper rate of 17%.[5] Because the entry rate is so small, the threshold for entry is way deep in the tail of the distribution (the $98th$ percentile to be precise). The size premium between new exporters and non-exporters is thus even larger, in the neighborhood of 200.

C.2 Zero without economies of scale

We take a look at the predictions for the size of exporters in a version of the model introduced in Bernard, Jensen, Eaton and Kortum (*American Economic Review*, vol. 93,

[5]The rate is for U.S. establishments in the period 1984-1992.

n.4, (2003), pp. 1268-1290), BJEK henceforth. For simplicity, we consider only two countries, the U.S. and the rest of the world (ROW). We assume a substantial familiarity with the BEJK model and attempt to follow their notation whenever possible.

C.2.1 Abridged model description

There is a continuum $[0, 1]$ of goods, indexed by j. There are two countries: (1) home (H), standing for the U.S.; and abroad (A), standing for the rest of the world. Demand everywhere combines these goods with a constant elasticity of substitution $\sigma > 0$. There are iceberg trade costs $\tau > 0$ between both countries but no fixed costs of exporting.

Each country has multiple potential producers of each good, with different productivity levels. Let $Z_{ki}(j)$ be the productivity of the k-th most efficient producer in country i of good j. Producers of the same good will compete head-to-head a la Bertrand in each country/market. Thus there is a single supplier of good j in each country i: It may be a domestic or a foreign firm. Because of the trade costs, it is possible that the best supplier for good j in each country is a domestic firm, i.e., there is no trade between the countries in good j. Thus firms split between exporters and non-exporters even if there are no fixed costs. Note that the productivity of the firm abroad producing the same good is effectively an additional source of heterogeneity, unrelated to the domestic's firm size as measured in domestic sales.[6] In this sense it fits our exercise in Section 3.

For a firm to be an exporter, it must have a high efficiency in order to out-price the best producer in the foreign market. As a result, exporters tend to be more productive and larger. However, there is no strict sorting in this model: a high-productivity firm may be unlucky and face an even more productive firm abroad producing the same good, and thus fail to have any exports; while a low-productivity firm may be lucky to have a very low productivity foreign counterpart and manage to sell in there.

[6] In addition, there are inactive firms, namely, those that are beaten in their own market by a foreign firm. However, those are dropped for the simulation as in BEJK.

Since firms compete a la Bertrand, the price in each market is set to be equal to the second most efficient supplier or the monopoly markup $\bar{m} = \sigma/(\sigma - 1)$, whatever is smaller. As a result, we only need to know the two most efficient producers of each good in each country, $Z_{1i}(j)$ and $Z_{2i}(j)$. BEJK assume these two productivity levels are drown from a joint Frechet distribution with c.d.f.

$$F_i(z_1, z_2) = \left[1 + T_i \left(z_2^{-\theta} - z_1^{-\theta}\right)\right] e^{-T_i z_2^{-\theta}}$$

where T_i is the country-specific location parameter and θ controls the dispersion in the distribution. See BEJK for a discussion of the underlying assumptions.

C.2.2 Solving the model

BEJK show how many of the model's implications can be easily computed using bilateral trade shares.[7] The latter are summary variable for all the cross-country heterogeneity in trade costs, income, and average productivity. Let π_{ij} be the share of sales in country i supplied by country j. It turns out π_{ij} is also the fraction of firms from country j that sell in country i. In our two-country setting, the share of exporters (among active firms) in the U.S. is given by

$$s_x = \frac{\pi_{AH}}{\pi_{HH}}. \tag{16}$$

That is, π_{AH} is the share of goods/firms provided by U.S. firms (country H) abroad. The denominator, π_{HH}, is the share of U.S. firms that are active, that is, they manage to beat the competitor from abroad, $\pi_{HH} = 1 - \pi_{HA}$.

Furthermore, BEJK show how which firm supplies a good j in each market n is simply given by

$$i^*(j) \quad = \quad \arg\min_i \left\{ \frac{U_{1i}}{\pi_{ni}} \right\}$$

[7]See Section IV.A. in their paper.

where $U_{1i}(j) = T_i Z_{1i}(j)^{-\theta}$ is a transformation of the efficiency term can be drawn from a parameter-free distribution. Using the same transformation for the second-most efficient producer in each country, $U_{2i}(j)$, BEJK show how to derive the markup, $M_i j$, at which the good j is supplied in country i. In order to obtain sales in market i for good j, we use the CES demand

$$X_i(j) = X_i \left(M_i(j)\right)^{1-\sigma} \left(\min_n \left\{ \frac{U_{1n}}{\pi_{in}} \right\} \right)^{(1-\sigma)/\theta}$$

where X_i is a combination of terms that are specific to the country/market, which can be derived in general equilibrium if one wishes to do so. By keeping track of which firms exports and which do not, we can retrieve the distribution of domestic and foreign sales. As in BEJK, we simulate the model with 10^6 firms, discarding firms that are not active domestically.

C.2.3 Quantitative evaluation

We need values for the country-specific demand intercepts, X_H, X_A and absorption rates π_{HH}, π_{AA}, as well as the elasticity of substitution σ and the dispersion parameter for the distribution, θ. For the U.S. absorption rate we have a direct counterpart in the data: We use gross manufacturing output minus exports of goods, against imports of goods (minus petroleum imports), as provided by the NIPA. This results in an approximate value $\pi_{HH} = .72$.[8]

Following the same steps as in the main text, we ensure that the model matches the share of exporters in the data. From equation (16), we have a single parameter left to do so, π_{AH}, which of course it is just the converse of the absorption rate in the ROW, $\pi_{AH} = 1 - \pi_{AA}$. In order to match the data, $s_x = .18$, we need to set $\pi_{AA} = .87$. This seems a reasonable estimate for the trade-weighted average U.S. market share abroad. In any case, the model is only sensitive to the *ratio* π_{AA}/π_{HH}.

[8] Not surprisingly, the absorption rate is not constant over time: We set it roughly to the average over the decade 1998-2008. Fortunately, the results are not sensitive to π_{HH}.

As in the main text, we also want to ensure the model matches the average total and domestic sales in the data. The demand intercepts X_H, X_A are exactly the two degrees of freedom we need to do so. There is no analytic formula for doing this. Instead we simulate the model and then scale domestic and foreign sales to deliver the correct averages.

We are left with two parameters, σ and θ. Ideally we would like to use both to match the distribution of total sales across firms, as we did in our choice of log-normal parameters in the main text. Unfortunately, these two parameters are poorly identified with the data at hand, as both govern the dispersion in sales in very similar ways. We instead rely on BEJK estimate for the elasticity of substitution, $\sigma = 3.8$, which lies well within the estimates elsewhere in the literature, including work that exploits price data. And then we use the dispersion in productivity, θ, to match the median total sales in the data, effectively capturing the right amount of skewness.[9] We obtain an estimate of $\theta = 2.6$.[10]

C.2.4 Benchmark results

We find the predicted exporter size premium, in total sales, to be about 75—only mildly below the implication of strict sorting and still way above the data counterpart of 4 to 5. Exporters in the model are substantially larger than in the data: In the model, they sell an average \$64 million domestically while in the data the exporter's average domestic sales is just \$31 million. As in strict sorting, non-exporters are way too small, selling less than \$1 million while they average in excess of \$8 million in the data. So while there is no strict sorting in BEJK, the relationship between size and exporter status is still too tight.

A closer look at the distribution of total sales implied by the benchmark calibration shows that the model is somewhat amiss of the data. Table 2 compares the distribution of total sales in the data and the model under the benchmark calibration described above. By construction, both average and median are matched.

[9] In the main text we targeted the standard deviation in log total sales. It turns out the BEJK model does not stand a chance if we do so: See the discussion in the next two sections.

[10] BEJK also estimate θ using U.S. plant data in a multi-country model, obtaining $\theta = 3.6$. We will return to these estimates below.

	Data		Model	
Size bin	Frequency	Average sales	Frequency	Average sales
0–$100,000	0.145	$55,600	0.006	$86,600
$100,000–$500,000	0.305	$257,000	0.404	$300,600
$500,000–$1 million	0.144	$718,000	0.240	$708,000
$1–5 million	0.257	$2.26 million	0.265	$2.09 million
$5–10 million	0.060	$6.84 million	0.040	$6.96 million
$10–50 million	0.063	$19.3 million	0.035	$20.4 million
$50–100 million	0.010	$56.4 million	0.005	$69.5 million
over $100 million	0.015	$670 million	0.005	$2,117 million
All Firms		*$13.2 million*		*$13.2 million*

Table 2: **The distribution of firm sales : Data and benchmark model**

The fit of the model is quite good at the center of the distribution but we run into trouble when it comes to both tails. First, there are too few really small firms. The reason is that U.S. firms with low productivity are likely to lose out against the foreign firm in the *domestic* market, and thus are inactive overall. Second, the right tail in the model is too long but too thin too early. This is a result of the Frechet distribution governing firms with very high productivity, which also sell a lot abroad. The resulting standard deviation of log total sales is just 1.27, way below the counterpart in the data, 2.6. Unfortunately these problems are endemic to the model: Any attempt to correct them by adjusting parameters θ, σ actually worsens the overall fit. For example, if we set θ to match the standard deviation of log sales, we obtain a median sales of barely $100,000 and completely unrealistic right tail—as well as an exporter size premium in excess of 200.

C.2.5 Robustness results

We explore some additional values for the dispersion parameter, θ, on the account it may be able to relax the sorting of exporters. And indeed we find it can, but only when the total sales distribution in the model has no resemblance with the data.

First we explore the parameter value estimated by BEJK, $\theta = 3.6$. The exporter size premium is drastically reduced to 8, only twice the value in the data. Exporters are still substantially too large compared with the data, but now non-exporters sell an average of \$5 million, much closer to the actual number. Unfortunately, the model approaches the exporter size premium because the total sales distribution is completely at odds with the data. The median firm in the model sells in excess of \$5 million, compared to \$600,000 in the data. There are virtually no firms with sales below \$1 million, while these amount to close to 60% of the firms observed in the data. By getting rid of small firms overall the model mechanically increases the average sales of non-exporters and thus reduces the exporter size premium.

Interestingly, BEJK obtained its estimate by matching the exporter size premium (for U.S. plants) in the data. As reported in Table 4, page 1283 of BEJK, this resulted in a very large fraction of exporters, 51%. As a check, we target a share of exporters of .51 with $\theta = 3.6$ and we did indeed obtain the correct exporter size premium. Unfortunately, the fit of the distribution of total sales actually worsens further, with the median firm now selling close to \$6 million and very few firms selling less than \$3 million.

Finally we explore different values for π_{HH}, the absorption share in the U.S. For each value, we re-estimate θ to match the median total sales. Increasing the absorption share all the way to .9 results in a slightly higher exporter size premium of about 85 and brings a quite modest improvement in the fit of the sales distribution. Reducing the absorption share all the way to .5 reduced the exporter-size premium somewhat, to 71—mainly by once again depleting the left tail of the distribution of total sales.

C.3 Industry-level heterogeneity

We compute the size premium implied by strict sorting for each three-digit NAICS industry code. The procedure is the same as the one we used in Section 2 for manufacturing as a whole. Bernard et al. (2007) report the share of firms exporting in each sector for 2002. As noted by Bernard et al. (2007), there is a large variation in the share of exporters. In Printing only 5% of the firms exports, while in Computers and Electronic Products almost

40% of the firms do. We also have the summary of the distribution of total sales for each sector—as in Table 2—provided by the Census. Unfortunately we do not have access to the same data for finer levels of dis-aggregation.

			Size Premium Prediction	
Industry	*NAICS code*	*Share Exporters*	*Lower Bound*	*Estimate*
Food	311	.12	115.3	118
Beverage and tobacco	312	.23	193.7	246
Textile mills	313	.23	49.6	69
Textile product mills	314	.12	57.6	68
Apparel	315	.08	55.8	56
Leather	316	.24	38.8	43
Wood product	321	.08	30.2	32
Paper	322	.24	48.7	53
Printing	323	.05	41.0	43
Petroleum and coal	324	.18	164.4	165
Chemical	325	.36	100.7	176
Plastics	326	.28	30.9	31
Nonmetallic mineral	327	.09	36.2	38
Primary metal	331	.30	69.9	70
Fabricated metal	332	.14	28.0	45
Machinery	333	.33	33.0	43
Computer and electronic	334	.38	72.4	97
Electrical equipment	335	.38	48.3	67
Transportation equipment	336	.28	190.5	298
Furniture	337	.07	48.4	50
Miscellaneous	339	.02	87.8	88
Aggregate Manufacturing	31-33	.18	81.2	95

Table 3: **Estimates for size premium predictions, by industry.**

Table 3 reports the size premium as predicted by strict sorting for each three-digit NAICS code. We compute a lower bound by assuming that all firms within any given size bin are identical; we also report a point estimate based on a fitted Pareto distribution. Both deliver the same message: *for each sector* the predicted size premium is very large. The reason is that the firm size distribution within a sector remains very skewed, so any strict sorting exercise is bound to return large size premiums. The differences on the share of exporters does create a lot of variation in the implied size premiums across sectors. However, it does not get them in the range observed in the data. Heavily traded sectors, like computer or electrical equipment, have implied size premiums of a magnitude comparable to the manufacturing sectors. Machinery (another commonly traded) has a somewhat lower size premium, but it is still well above the differences documented in the data.