Economies of Scale and the Size of Exporters*

Roc Armenter
Federal Reserve Bank of Philadelphia

Miklós Koren
Central European University, IEHAS and CEPR

December 1, 2008

Abstract

Exporters are few—less than one fifth among U.S. manufacturing firms—and are larger than non-exporting firms—they have about 4-5 times more total sales on average. These facts are often cited as support for models with economies of scale and firm heterogeneity as in Melitz (2003). We find that the basic Melitz model cannot simultaneously match the size and share of exporters given the observed distribution of total sales. Instead exporters are expected to be between 90 and 100 times larger than non-exporters. To reconcile the model with the data we need a lot of variation independent of firm size. We show that the augmented model has markedly different implications in the event of a trade liberalization. Crucially, the extensive margin plays a minor role; and productivity gains due to reallocation are halved.

*PRELIMINARY. We thank George Alessandria for helpful comments as well as the participants of the CEU seminar. The views expressed here do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

1 Introduction

Exporters represent less than one fifth of all the U.S. manufacturing firms, and are markedly larger than non-exporters—on average their total sales are about 4-5 times higher.[1] These stylized facts are often cited as support for models with economies of scale and firm heterogeneity. In Melitz (2003), only larger, more productive firms generate enough revenues in the foreign market to cover the fixed costs associated with exporting. A key implication is that, in the event of a trade liberalization, resources will be reallocated towards the more productive firms, raising the average productivity in the industry.

In this paper we evaluate the *quantitative* implications of the Melitz model for the share and size of exporters. According to the basic Melitz model we should observe strict sorting by size, that is, the smallest exporter should be larger than the largest non-exporter. The set of exporters is thus easily characterized with a cut-off rule in terms of total sales. We use this equilibrium restriction to derive the quantitative implications of the theory without having to specify the full model.

We find that the basic Melitz model cannot simultaneously match the share and size of exporters in the data given the large skewness observed in the distribution of total sales.[2] Given that exporters are roughly one fifth of all firms, strict sorting suggests that exporters should be concentrated in the top quintile of the firm size distribution. We can thus obtain the model prediction for the size of exporters by comparing the top quintile firms with the rest. We find that the firms below the top quintile are quite small (an average of $740,000 total sales), while the firms in the top quintile are much larger ($70 million on average). Hence the basic Melitz model greatly overpredicts the size premium of exporters: exporters should be between 90 and 100 times larger than non-exporters.

Of course we did not expect the strict sorting of exporters to hold *exactly* in the data. Melitz (2003) certainly does not intend to preclude the importance of other idiosyncratic factors, unrelated to size, in the firm's decision to export. For example, the costs of exporting—variable or fixed—may vary from firm to firm.

We proceed to reconcile the model with the data by introducing further firm-level heterogeneity. At first pass we take a latent variable approach so there is no need to specify what are the additional sources of variation in the export decision—only that these factors are independent of the firm size. We find that we need a lot of size-independent variation in order to bridge the large gap between the model and the data. This indicates that size plays only a small role in the determination of the export status of a firm.

Our main finding is that the large amount of independent variation "waters down" the core mechanism in the Melitz model. To illustrate this we set up a very simple, partial equilibrium model based on Melitz (2003). We then explore two versions of the model. In

[1] Both facts are documented in Bertrand, Jensen, Redding and Schott (2007) for the universe of U.S. manufacturing firms operating in 2002.

[2] We use the distribution of total sales for manufacturing firms (NAICS code 31-33) as given by the 2002 Statistics of U.S. Businesses of the Census.

the first version productivity is the only source of variation across firms. The model displays then the strict sorting property and, as we discussed above, overstates the size of exporters. In the second version we assume firms face heterogeneous fixed costs. We then calibrate the distribution of fixed costs to match the share and size of exporters.[3] From the previous discussion, we know we need a huge amount of dispersion in the fixed costs. We then compare the models in the event of a trade liberalization, with trade costs falling by half.

Critically, the augmented model has only a minor role for the extensive margin. First, there is much less entry than in the standard Melitz model in response to the fall in trade costs. We find that the growth rate of the number of exporters is more than 60% in the strict sorting version, but only 15% when all the latent heterogeneity is accounted for. Second, new and existing exporters have relatively similar size in the latent heterogeneity model. This stands in marked contrast with the standard Melitz model. In the latter, new exporters are ten times smaller than the average exporter prior to the trade liberalization. The differences across models are more striking given that both models have very similar implications for the aggregate trade response.

We seek to quantify further the role of the extensive margin in both models. To this end we decompose total export growth (normalized by total sales) in an intensive and extensive margin. In the standard version the extensive margin accounts for more than 60% percent. Once we match the size of exporters, the extensive margin accounts for less than 20%. The muted response of export participation is quite surprising given the Melitz model is by now considered the workhorse model for an expanding literature on the extensive margin.[4]

We also look at productivity gains due to the reallocation effect across both models. We find that the productivity gains due to the trade liberalization are halved in the augmented model. This is perhaps not surprising given the previous results. However, it must be noted that we miss on the productivity gains due to exit in the domestic market since our model is partial equilibrium and the wage rate—or more generally, unit costs—is taken as given. Our results prove to be robust to the choice of parameters and distributions.

Finally, we explore alternative specifications for the additional heterogeneity. First we consider the possibility that firms face heterogeneous variable costs associated with exporting. We argue that we should not expect big differences: the model is essentially log-linear and how the heterogeneity enters the model is close to irrelevant. We also ask whether the latent heterogeneity stems from differences across broadly-defined industries. If that were the case we would find that the size premium of exporters within each industry would be closer to the data. Instead we find that for each 3-digit manufacturing industry the implied size premium remains implausibly large.

We then move to the possibility that the fixed costs of exporting are sunk. In this case there is exporter hysteresis and the particular history of each firm can be the source of heterogeneity at play. More importantly, this form of heterogeneity would be compatible

[3] The distribution of fixed costs is the only difference between the two models. We use a common value for all remaining parameters, including those governing the distribution of productivity.

[4] See Baldwin and Harrigan (2007) for example.

with large adjustments along the extensive margin. We show that sunk costs can easily match the share and size of exporters. However, it is now the size of the *new* exporters that presents a puzzle: new exporters should be significantly larger than incumbent exporters.

The paper is organized as follows. Section 2 derives the size of exporters in the basic Melitz model. Section 3 introduces a latent variable to reconcile the model with the data. The implications of the latent heterogeneity in a structural model are explored in Section 4. We explore some alternative specifications for the heterogeneity in Section 5. We also include a brief discussion on the empirical literature on the extensive margin. Section 7 concludes.

2 On the Size of Exporters

In the Melitz model, firm i will export only if its foreign sales, net of the associated variable costs, would cover the fixed costs associated with exporting,

$$r_i^* - c_i^* \geq f. \tag{1}$$

Firms differ in their labor productivity or in the quality of goods produced. Either way, more efficient firms can generate more net income abroad and are thus more likely to be exporters. Hence the model is unambiguous on its prediction that exporters are more productive and sell higher quality goods than non-exporters. Incidentally, more efficient firms also sell more in the domestic market, so exporters are also larger than non-exporters in terms of total sales. These qualitative predictions are borne in the data and often cited as support for the Melitz model.

We seek to explore the *quantitative* predictions of the model on the exporters' characteristics. The first step is to rewrite the entry condition (1) in terms of total sales. Firm-level data on total sales are easily accessible and, more importantly, we can observe a firm's total sales independently of whether the firm exports or not. In the basic Melitz model, there is a tight relationship between a firm's total sales and its underlying efficiency parameter: a firm with higher productivity will always have more total sales in equilibrium. Since net income abroad is also strictly increasing in productivity, we have an increasing, monotone relationship between total sales and net income abroad—even if the latter is a counterfactual because the firm does not export.[5]

We can thus summarize the model's predictions for the set of exporters with a simple threshold rule in place of (1): firm i will export only if its total sales r_i are above some level t,

$$r_i \geq t. \tag{2}$$

It is immediate that exporters and non-exporters are strictly sorted by size, that is, the smallest exporting firm has more total sales than the largest non-exporting firm. The value

[5] The same reasoning applies if firms differ in the quality of their products—as long as the quality is constant in each market they serve.

of the threshold level t is determined in equilibrium and is bound to depend on the model's parameters.

We can, though, easily derive the model's predictions using the share of exporters and total sales distribution observed in the data. The threshold condition (2) implies that the share of exporters is equal to the fraction of firms with total sales equal or larger than t. We have thus that

$$s_x = 1 - \Psi_r(t),$$

where Ψ_r is the empirical c.d.f. of total sales, and s_x is the fraction of firms with positive export sales. In other words, we can solve for the $(1 - s_x)$th percentile in the distribution of total sales and obtain the threshold t consistent with the observed share of exporters. We can then easily compute the truncated mean,

$$E\{r_i | r_i \geq t\} = \int_t^\infty r_i \frac{d\Psi_r(r_i)}{s_x},$$

so we can compare the model's implications for the average total sales for exporters and non-exporters.

Bertrand, Jensen, Redding and Schott (2007) report that only 18% of the U.S. manufacturing firms had positive sales abroad in 2002.[6] Exporters, thus, are relatively rare. For the distribution of total sales we look at the 2002 Statistics of U.S. Businesses of the Census. Table 1 summarizes the data for manufacturing firms.[7] As it is well known, there is an enormous amount of skewness in the size distribution of firms. The average firm sells \$13.2 million and yet 45% of the firms sell less than \$500,000. In short, there are many many small firms and a few very large ones.

Size bin	Frequency	Cumulative Frequency	Average sales
0–\$100,000	0.145	0.145	\$55,600
\$100,000–\$500,000	0.305	0.450	\$257,000
\$500,000–\$1 million	0.144	0.594	\$718,000
\$1–5 million	0.257	0.851	\$2.26 million
\$5–10 million	0.060	0.911	\$6.84 million
\$10–50 million	0.063	0.974	\$19.3 million
\$50–100 million	0.010	0.984	\$56.4 million
over \$100 million	0.015	1.000	\$670 million
All Firms			*\$13.2 million*

Table 1: **The distribution of firm sales in manufacturing** – *Census*

[6] Exporters are similarly scarce if we look at plants or establishments rather than firms—see Bernard, Eaton, Jensen, and Kortum (2003) for example. The scarcity of exporters has also been confirmed in a variety of countries.

[7] NAICS code 31-33.

From Table 1 we see that the 82nd percentile falls somewhere between \$1 and \$5 million sales, definitively closer to the latter. This already suggests that non-exporters are expected to be quite small under strict sorting: firms below \$1 million sales would constitute over 70% of all non-exporters and average less than $\$320,000$ in sales. At the same time, exporters will be pretty large. Firms above \$1 million average \$82 million in total sales and would represent over 80% of all the exporters. We want to be a little bit more precise than this, though. For the firms within the range $\$1 - \$5$ million we do not know the exact distribution: we assume these firms follow a two-side truncated Pareto distribution, parameterized to match the average total sales in the range (\$2.26 million).[8]

We find that strict sorting by size implies that exporters should sell, on average, between 90 and 100 times more than non-exporters. The exporter size premium remains very large for whatever distribution one assumes for the firms in the range $\$1 - \$5$ million. As a check we computed a lower bound on the exporter's size by taking all firms with sales between \$1 million and \$5 million to be identical, with total sales equal to the bin's average (\$2.26 million). Even in this case strict sorting predicts exporters are more than 80 times larger than non-exporters.

How does the implied exporter size premium compare with the data? Bertrand et al. (2007) report that exporters are 4 to 5 times larger than non-exporters.[9] Strict sorting thus greatly overpredicts the size differences between exporters and non-exporters. Table 2 compares the size of exporters and non-exporters in the model with the data for U.S. manufacturing firms in 2002. Under the hypothesis of strict sorting, exporters should have, on average, \$70.1 million in total sales—double of what we actually observe. In the model non-exporters are expected to be very small, just $\$740,000$ in total sales. In the data, though, there are clearly some non-exporters that are large enough to bring their average total sales above \$8 million.

	Data	Strict Sorting
Average Sales - Non-exporters	\$8.1 million	\$740,000
Average Sales - Exporters	\$36.4 million	\$70.1 million
Exporter Size Premium	4.5	95

Table 2: **Exporters and non-exporters in the model and the data**

The magnitude of the gap between the basic model and the data already suggests that size cannot be the sole determinant of the export status of a firm. As a matter of fact the

[8] There is no need to assume anything about the distribution of firms outside this size bin: they are all squarely in the exporter or non-exporter category.

[9] Table 3 in Bertrand et al. (2007) states that the difference in average log shipments between exporters and non-exporters is 1.48 for the same set of firms we used. The finding that exporters are larger than non-exporters has also been confirmed for plants and establishments, as well as for other countries. In this Section we work with total sales rather than log total sales so we can use the information in Table 1 without any further assumption.

opposite hypothesis—that is, that firm size has nothing to do with its export status—is closer to the data. The ratio of export sales to total sales is about .16 for manufacturing as a whole. So even if we select 18% of the firms at random as exporters, their total sales will amount to almost 90% more than non-exporters, or an exporter size premium of almost 2.[10]

The model cannot match the size of exporters and non-exporters even if we allow some leeway in the fraction of exporters. Table 3 reproduces the previous exercise for different assumptions on the fraction of exporters, s_x. As we increase the share of exporters, the model's prediction for the size premium declines but does not get anywhere close to the actual number. For example, say we take exporters to be twice as frequent as they actually are, $s_x = .35$. The average total sales of exporters is then very close to the data. Non-exporters, though, are then very small and the size premium is still 15 times larger than in the data.

	Data	Strict Sorting		
Share of exporters s_x	.18	.15	.25	.35
Average Sales - Non-exporters	$8.1 million	$746,000	$637,000	$491,000
Average Sales - Exporters	$36.4 million	$83.9 million	$51.0 million	$36.8 million
Exporter Size Premium	4.5	114	89	75

Table 3: **Exporters and non-exporters given the share of exporters s_x**

3 Introducing Latent Heterogeneity

Of course we did not expect the strict sorting of exporters to hold *exactly* in the data. Melitz (2003) certainly does not intend to preclude the importance of other idiosyncratic factors, unrelated to size, in the firm's decision to export. For example, the costs of exporting—variable or fixed—may vary from firm to firm. It is possible then a low productivity firm finds exporting profitable simply because it is quite cheap to transport its product. Alternatively we could think that whatever makes a product successful in the domestic market does not always translate into more sales in the foreign market. Clearly once we consider these additional factors the strict sorting of exporters no longer holds.

Here we proceed to reconcile the model with the data by introducing the necessary firm-level heterogeneity in the export decision. We rewrite the threshold condition (2) in terms of a latent variable, t_i. A firm i exports if its total sales satisfy

$$r_i \geq t_i. \tag{3}$$

[10] Armenter and Koren (2008) show that the split between exporters and non-exporters is not explained by the sparsity of the data, a potential source of randomness in the exporter status. The difference is that in the exercise above we have taken the share of exporters as given.

The latent variable is identically and independently distributed across firms, with a c.d.f. Ψ_t over support $\Re_+$.

With the latent variable condition (3) we can capture all the underlying heterogeneity without having to specify the sources of variation. Indeed the only structure imposed on the latent variable is the independence with firm size. As in the previous Section, we can postpone laying out and solving the structural model while seeing through the model's implications given the data.

It is now useful to fit the empirical distribution of total sales with a parametric distribution. We use a lognormal distribution with mean $\mu_r = 6.3$ and standard deviation $\sigma_r = 2.6$—so we reproduce the average total sales (in thousands of dollars) as well as the approximated location of the 82th percentile. The use of the lognormal distribution for firm size has a long tradition in economics. See Aitchison and Brown (1969) for a complete treatment of the distribution.

We also assume that the latent variable t_i follows a lognormal distribution. This is mainly a choice of convenience: we want a flexible two-parameter distribution defined over positive numbers. We pick the mean μ_t and standard deviation σ_t such that the model reproduces the share and size of exporters. That is, we solve for μ_t and σ_t such that equations

$$s_x = \int \Psi_t(r)\, d\Psi_r(r), \tag{4}$$

$$E\{\log(r_i)\,|r_i \geq t_i\} = \int \log(r)\, \Psi_t(r)\, d\Psi_r(r)\,/s_x, \tag{5}$$

reproduce the observed values for the share of exporters, $s_x = .18$, and the average log total sales of exporters, $E\{\log(r_i)\,|r_i \geq t_i\} = 5.66$.[11] We use Ψ_t to denote the c.d.f. of the latent variable distribution.

The latent variable is very dispersed: we find that the parameter values matching moments (4)-(5) are $\mu_t = 13.73$ and $\sigma_t = 7.7$. These parameters imply that the coefficient of variation of the latent variable is more than 10^{10} times the respective number for total sales! We explored an array of parameters for the distribution of total sales and found always that we need a huge dispersion for the latent variable in order to reproduce the share and size of exporters. More precisely, we consider values μ_r in the range $5.5 - 7.5$ and σ_r in the range $2 - 3$. The resulting parameters for μ_t and σ_t were always above 10 and 5 respectively.

The results are perhaps not that surprising: after all we have to make up for a large gap between the model and the data in the size premium of exporters. In order to reduce the size of exporters we need the latent variable to take very large values with high probability, so some large firms do not export. Simultaneously, some other firms must draw a low realization of the latent variable and export independently of the size.

[11] Here we use the log total sales instead of total sales, so the exporter size premium is now given by the difference in average log total sales between exporters and non exporters. The change of units has virtually no implication for the parameters of the latent variable distribution—but it turns out to be very convenient for the calibration of the model later.

The huge dispersion of the latent variable indicates that size plays only a small role in the determination of the export status of a firm. We can illustrate this point with a simple exercise. Note that given a firm size r_i, the probability that firm i exports is

$$\begin{aligned} \Pr\left(i \text{ exports } |r_i\right) &= \Pr(r_i \geq t_i|r_i) \\ &= \Psi_t\left(r_i\right). \end{aligned}$$

Given our results for Ψ_t, a firm of median size exports with probability .167, very close to the unconditional probability of .18.[12] In other words, a firm of median size could be taken as representative of the industry as a whole. In contrast, a firm with the median latent variable will export only with probability .0016. Table 4 repeats the exercise with the $25th$ and $75th$ percentiles for total sales and the latent variable.

	Percentile		
Ordered	$25th$	$50th$	$75th$
By total sales	.1179	.1674	.2286
By latent variable	10^{-7}	.0016	.1881

Table 4: **Conditional probability of exporting**

In the next Section we will explore the implications of the additional heterogeneity in the context of a structural model. Here we anticipate our main result with a simple exercise that illustrates how the predictions for the extensive margin crucially depend on whether we match the exporter size in the data or not.

The exercise is as follows: we increase the total sales of *all firms* by a constant percentage δ. We then look at the predictions for the growth rate of the number of exporters for both models: the strict sorting model given by (2) and the latent variable model given by (3). It must be emphasized that there is no change either in the threshold or the distribution of the latent variable.

Table 5 reports the findings for the growth rate in the number of exporters for revenue increases of 5%, 10% and 20%. The strict sorting model predicts a relatively large fraction of firms change status from non-exporter to exporters. For a 10% increase in total sales the model predicts 2.69% growth rate in exporters. The distribution of total sales is still pretty thick just below the 82nd percentile. Because size is the only determinant of the export status of a firm, all the firms whose total sales were originally δ below the threshold t export after the revenue increase.

In contrast the latent variable model predicts much smaller growth rates for the number of exporters. Even after a 10% increase in total sales the growth rate is smaller than under a 5% increase in the strict sorting. This just illustrates that size is not a key determinant of the export status once all the heterogeneity is accounted for. Loosely, the gain in size must

[12] The median size is given by $\exp\left(\mu_r\right)$ or approximately \$600,000.

	Revenue increase δ		
Model	$\delta = 5\%$	$\delta = 10\%$	$\delta = 20\%$
Strict Sorting	2.69%	5.29%	10.26%
Latent Variable	0.67%	2.23%	3.67%

Table 5: **Growth rate of the number of exporters**

be quite large in order to overcome the other determinants of the exporting decision. Hence even when total sales increase by 20% the growth rate of exporters is only 3.67%.

4 Trade Liberalization and Exporters

In this Section we set up a small model with economies of scale and firm-heterogeneity. The model is simpler than Melitz (2003) in that it is partial equilibrium, taking the wage rate as given. As a result the model abstracts from entry and exit in the domestic market.

We explore the implications of the additional heterogeneity for a trade liberalization episode. To this end we consider two versions of the model. In the first one there is strict sorting, so the model does not match the size of exporters. In the second one we introduce the needed variation as a random fixed cost. The rest of the parameters are common across models.

4.1 A simple model of exports and exporters

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 that 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 $y^d(\omega)$ denotes 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 its product. The profit-maximization problems 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 focused on the export growth.

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

$$r^d(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}}\right)^{\theta-1} R^d \tag{6}$$

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.[13] We will still make use of the relationship between productivity and domestic sales given by (6).

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

$$Y^f = \left[\int_{\Omega_x} \left(y^f(\omega)\right)^\rho d\omega\right]^{1/\rho} \tag{7}$$

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

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

[13]Briefly, $R^d = r^d(\tilde{\varphi})$ by (6), 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

$$P^f = \left[\int_{\Omega_x} \left(p^f(\omega)\right)^{1-\theta} d\omega\right]^{\frac{1}{1-\theta}}. \tag{8}$$

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

$$Y^f = Y^* \left(P^f\right)^{-\nu}, \tag{9}$$

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 any given 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,

$$r^f(\varphi) = \left(\frac{\varphi}{\tilde{\varphi}_x}\right)^{\theta-1} \frac{R^f}{M_x} \tag{10}$$

where R^f is total export sales and

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

is the average productivity among exporters. Note that the set of exporters Ω_x affects the export revenues of each firm (10) 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 only find profitable to export if its net income abroad would cover the fixed expenses,

$$\frac{1}{\theta} r^f(\varphi(\omega)) \geq f(\omega), \tag{11}$$

where we have used that export net income, that is, export revenues minus costs, can be 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 (11) holds. However, export revenues depend themselves on the set of exporters, so in equilibrium exports and exporters are determined simultaneously.

4.2 Strict sorting and latent heterogeneity

We now consider two versions of the model above. In the first we stick to the basic Melitz model and shut down the heterogeneity in the fixed costs. As a result exporters and non-exporters are strictly sorted by size and the model inherits the inability to match the size of exporters as we documented in Section 2. We name this first version of the model after the strict sorting property. In the second calibration we use the dispersion on fixed costs to reproduce the latent variable distribution we worked out in Section 3. By construction the model then matches the share and size of exporters. We label this calibration as the "latent heterogeneity" model.

It must be emphasized that the *only* difference between the two models is the distribution of fixed costs. The models will share the same parameter values for the elasticities, trade costs, and the distribution of productivity. We will also set the exogenous variables such that both models match the same facts in the data.

We start thus with the common parameters. We set the elasticity of substitution ρ across exported goods such that the price elasticity of each good is $\theta = 8$. This number is essentially in the middle of the range of estimates surveyed by Anderson and Van Wincoop (2004). For the aggregate demand for home exports (9) we set the price-elasticity of $\nu = 6$. Our baseline trade costs are set at 50%, $\tau = 1.5$. This is the midpoint between the estimated trade costs in 1987 and 2002 by Alessandria and Hoi (2008). The levels of the foreign income and the wage rate are set to reproduce the ratio of total exports to total sales in the industry.

The last common parameter is the distribution of productivity, G. For the calibration we want to use the observed distribution for total sales. Unfortunately the mapping from the productivity distribution to the total sales distribution is not the same for both versions of the model.[14] It is thus not possible to have thus a common parametrization for G that matches the observed distribution of total sales in both models. However, we do not want any difference in the calibration of the models to govern the results—other than the size of exporters, that is. We thus adopt the following compromise: we set the productivity distribution to capture *all* the variation in total sales. More precisely, we take G to be a lognormal distribution with standard deviation equal to $\sigma_\varphi = \sigma_r/(\theta - 1)$. The location parameter can be set such that the average productivity $\tilde{\varphi}$ among all producers is normalized to one.

Table 6 summarizes all the parameters common to both models.

Finally we get to what makes the strict sorting and latent heterogeneity models different. For the strict sorting we shut down all the variation in the fixed cost, so H is a degenerate distribution at some point f. The value of the fixed cost f is set such that the share of exporters is 18% as in the data. For the latent heterogeneity model we calibrate the distribution of fixed costs H such that we capture all the variation in the latent variable t_i. For this it is very convenient to use a log normal distribution for fixed costs so the mapping is

[14] Both models share the mapping from G to the distribution of *domestic* total sales. The distribution of export sales, though, are different for each model.

Parameter		Value
Elasticity of Substitution	θ	8
Price-Elasticity exports	ν	6
Trade Cost	τ	1.5
Std.Dev. Log-productivity	σ_ψ	.37
Mean Log-productivity	μ_ψ	-0.48

Table 6: **Calibration - Common Parameters**

simple. By construction the latent heterogeneity model replicates the observed exporter size premium; by appropriate choice of the location parameter we match the share of exporters as well. The resulting parameters are $\mu_f = .98$ and $\sigma_f = 7.69$.

4.3 Trade liberalization

We compare the two model's predictions for export growth in response to a fall in trade costs. The exercise is labeled as a "trade liberalization" but there is no distinction in our model whether it is a tariff reduction or an improvement in the shipping technology.[15] We consider a range of trade liberalizations up to a halving of the trade costs—a fall of 25 percentage points.

We find that both models have virtually identical implications for trade volume. Figure 1 plots the growth rate of different variables (as a percentage) as a function of the fall in trade costs (in percentage points). The solid and dashed line correspond to the strict sorting and latent heterogeneity models respectively. The top left panel displays the growth rate of exports. Both models predict that trade will approximately triple once trade costs are cut in half. The models' predictions are so similar that the lines are on the top of each other for most of the range of trade costs. The top right panel in Figure 1 displays the growth rate in the export price index, as given by (8). Again there are no differences between both models, and the price index falls with trade costs at the same rate for both models. Hence the profile for export revenues is very similar (bottom left panel). Finally the bottom right panel in Figure 1 displays total employment in the industry. The trade liberalization results in a sizeable expansion of employment, about 10% for the largest fall in trade costs considered.

As a matter of fact, the path for total exports in both models is undistinguishable from the predictions of a representative firm model.[16] This is quite obvious in Figure 2, where we have added the representative firm model, indicated with a dotted line. We plot export

[15] Or, for that matter, an exogenous change in the real exchange rate.

[16] The representative firm model is just a single equation

$$log(Y^f) = -\nu log(\tau) + cte. \tag{12}$$

which can be easily derived from the foreign demand given by (9). We maintain the same value for the price-elasticity ν given in Table 6.

growth (top) and the export price growth (bottom) as a function of the fall in trade costs. Clearly all three models are very similar in both prices and quantities. Why firm-level heterogeneity does not make a difference for aggregate variables? The straight answer is that, even after a fall of 25 percentage points in trade costs, the vast majority of exports are still made by the firms that were exporting initially. We will return to this point later once we have explored the models in depth.

Behind the similarities for trade volume the models display marked differences for the number of exporters. The top panel in Figure 3 plots the growth rate in the number of exporters for each model. In the strict sorting model the number of exporters grows very fast. Entry drives up the number of exporters up by almost 60% when trade costs fall by 25 percentage points. Even with a small drop in trade costs like 5 percentage points the number of exporters grow by more than 10%, suggesting that even at relatively high frequencies entry could play an important role.

In contrast, there is not much entry in the foreign market in the latent heterogeneity model. The growth rate of the number of exporters barely gets over 15%, one fourth of the growth in the strict sorting model. For smaller trade costs reductions like 5%, the number of exporters is very close to flat. The similarities for total trade volume only make the differences in entry even more striking.

We seek to quantify further the role of export entry in both models. We decompose the growth rate of exports in the change of export intensity and participation,

$$\Delta log(Y^f) = \Delta log(r^f) + \Delta log(M_x). \tag{13}$$

The first term on the right hand side is the intensive margin, that is, the growth rate on the average export revenues per exporter; and the second is the extensive margin, or the growth in the number of exporters. We normalize export growth by total sales, and express each margin as a percentage of the total.[17] Table 7 collects the results for trade cost reductions of 5, 15, and 25 percentage points. In the strict sorting model the extensive margin is more than 60% of the growth rate in exports. However entry plays a much minor role in the latent heterogeneity model, just below 20%. These numbers are very similar across the range of trade cost decreases. They are also quite robust to alternative parameterizations for elasticities. This leads us to the conclusion that once the Melitz model is at odds with a large role for the extensive margin once it is augmented to account for the share and size of exporters.

The reader may be wondering how can the two models have such different entry rates and yet very similar export growth. The bottom panel of Figure 3 clues us in. Despite having much lower entry rates, total employment by exporters actually grows much faster in the latent heterogeneity model. The reason is that new exporters are very different in the two models. In the strict sorting model new exporters are very small in comparison with

[17] We are following Alessandria and Hoi (2008) closely, although we do not distinguish between export intensity and premium as they do.

	Trade Cost Decrease		
Model	5	15	25
Strict Sorting	63.3%	63.3%	63.5%
Latent Heterogeneity	19.2%	19.8%	20.4%

Table 7: The Role of the Extensive Margin. *Trade Costs decrease in percentage points.*

the incumbent exporters, about 10 times smaller than the average exporter size prior to the trade liberalization.[18] This is a direct consequence of the strict sorting: since exporters are almost 100 times larger than non-exporters, the firm at the threshold is still quite large compared with non-exporters but, more importantly, it is very small compared with exporters. Returning to Table 2 in Section 2, the *smallest* exporter has about 5 times more total sales than the average non-exporter, but 18 times less total sales than the average exporter.

In contrast new exporters in the latent heterogeneity model are still smaller than existing exporters, but not by much. For a fall in trade costs of 5 percentage points, new exporters are just 40% smaller than the average exporter size prior to the trade liberalization. If we go all the way to a halving of trade costs, new exporters are actually 30% larger than the average exporter size prior to the trade liberalization. In short, new exporters are not very different from the average firm in the industry.

We now evaluate the main mechanism in Melitz (2003), namely, the gains in average productivity in the industry due to the trade liberalization. We follow Melitz (2003) and define

$$\hat{\varphi} = \left(\tilde{\varphi}^{\theta-1} + M_x \left(\tilde{\varphi}_x/\tau\right)^{\theta-1}\right)^{\frac{1}{\theta-1}}$$

as the aggregate productivity.[19] It must be noted that our model can only do a partial evaluation. Since we take the wage rate as given, the set of non-exporters does not change so there are no productivity gains from the exit of the least productivity firms as in Melitz (2003).

We find the aggregate productivity grows significantly less in the latent heterogeneity model than in the strict sorting model. Figure 4 plots the results. Aggregate productivity growth in the latent heterogeneity is about half the one in the strict sorting model—a ratio that is approximately constant across the range of trade cost decreases.

That productivity gains are smaller is not surprising given our previous results. Given that we abstract from exit in the domestic market, it can be said that the mechanism in Melitz (2003) works through the selection and entry of exporters. First, exporters are more productive than the average firm, so anything that expands the total employment of exporters will lead to productivity gains due to composition effects. Second, new exporters

[18]Because previous exporters grow rapidly with the trade liberalization, the new exporters are 50 times smaller when compared with the average size the incumbent exporter after the trade costs decrease.

[19]Melitz (2003) refers to is as the combined average productivity and shows it completely summarizes the effects of the distribution of productivity levels on the aggregate outcome.

experience a large jump on their output and, since they are still more productive than most firms in the economy, also induce gains in the average productivity.

Both selection and entry are much weaker in the latent heterogeneity model than in the strict sorting model. We have seen that strict sorting greatly overpredicts the size of exporters, that is, their productivity. The latent heterogeneity reconciles the model and data: exporters are still more productive than non-exporters, but only modestly so. Moreover there is much less entry in the latent heterogeneity model, so the second source of productivity gains is weaker too.

We conclude our analysis with a look at how the average exporter changes with the trade liberalization. Figure 5 displays the average among exporters for export revenues, export output, average productivity, and total employment. It must be emphasized that the set of exporters changes as we cut the trade costs. First we note that export revenues and output grow much faster in the latent heterogeneity model. This is, of course, the flip side of the results on the extensive margin documented in Table 7. Second, the average productivity for exporters falls in both models but by different amounts, as shown in the bottom left panel in Figure 5. In both models new exporters are less productive than incumbent exporters; but in the strict sorting they are much less so. As a result the exporter average productivity drops by a staggering 6% while only a modest 1% in the latent heterogeneity.

The differences in the average output and productivity of exporters combine for opposite predictions with respect to the total employment for exporter. In the strict sorting model exporters, on average, employ less workers as trade costs fall. New exporters do employ more workers than before entering the foreign market. However, they are so small compared with the incumbent exporters that they bring the average down by an astounding 25%. Recall that the number of exporters grows by more than 50% so entry has a big impact on averages. In contrast total employment for exporter grows in latent heterogeneity, as the weaker selection and entry effects cannot overturn the employment gains due to overall expansion of exports.

4.4 Robustness

We briefly discuss here alternative specifications for the common structure of both models. We start with our choice of the lognormal distribution for firm productivity. In particular, theory work has favored instead the Pareto distribution for its tractability. It has also been argued that the Pareto distribution is a very good approximation of the tail of the firm-size distribution.[20]

Unfortunately the Pareto distribution proves to be a very restrictive choice for our purposes. First and foremost, a Pareto distribution for firm productivity implies that the intensive margin completely shuts down under strict sorting. We can easily see this by examining

[20] See Axtell (2001) and Luttmer (2007).

the entry condition (11). Strict sorting implies there is a least productive exporter, which we denote $\hat{\varphi}_x$, for which the entry condition holds with strict equality:

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

The CES demand system implies that we can relate the revenues of any firm to average revenues,

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

Here it is where the assumption of Pareto distribution comes in. If x is a random variable distributed with a Pareto distribution with slope k over support $x \geq x_0$ then $E(x|x \geq x_1) = k/(k-1)x_1$ for any $x_1 \geq x_0$. It is easy to show that if φ is distributed according to a Pareto with slope k, then $\varphi^{theta-1}$ is distributed with a Pareto with slope $k(\theta-1)$. We can thus rewrite the entry condition simply as

$$\frac{k(\theta-1)-1}{k(\theta-1)} \frac{R^f}{M_x} = \theta f.$$

It follows that the average export revenues per firm are *not* a function of trade costs. Thus as trade costs fall all the adjustment must occur through the extensive margin, that is, the number of exporters.[21] We do not regard the Pareto distribution assumption as a fundamental piece of the theory in Melitz (2003) so we do not think its extreme implications for the extensive margin merit further consideration.[22] Instead we favor a more flexible specification that allows for both margins to be active.

The Pareto distribution also proves to be problematic for the calibration of the latent heterogeneity. The simple reason is that exporters are not concentrated in the tail of the firm-size distribution, so the Pareto distribution is a poor approximation of the size distribution of exporters. Returning to Table 1 in Section 2, we find that a Pareto distribution cannot encompass the distribution of firms above $ 1 million properly. For estimates of the slope parameter in Axtell (2001) and Luttmer (2007), we find that the exporter size premium implied by strict premium is above 120.

Finally we also explore some alternative parameter specifications. Table 8 reports the growth rate for exports, the number of exporters, and aggregate productivity in the event of a trade cost reduction of 10 percentage points for both the strict sorting (SS) and the latent heterogeneity (LH) model. In addition to the baseline calibration, we consider two alternative parametrizations. In the first we set the elasticity of substitution between exports to 12. There is no significant change for both models predictions for the growth rate of exports

[21] It is important to recall that we have abstracted from general equilibrium effects on the wage rate which, in turn, could have affected the fix costs associated with exporting. This channel is operative in Melitz (2003) and leads to adjustment along the intensive margin.

[22] This view is not shared by many. See Chaney (2008) for example.

	Baseline		$\theta = 12$		$\nu = 3$	
	SS	LH	SS	LH	SS	LH
Total Exports	52.8	55.8	52.3	54.1	23.3	23.7
Number Exporters	20.3	6.1	20.1	6.0	7.7	2.4
Agg. Productivity	0.5	0.2	0.1	0	0.5	0.2

Table 8: Robustness Analysis. *All variables are growth percentage rates after a fall of 10 percentage points in trade.*

and exporters, but there is now virtually no aggregate productivity growth. This shows that the "love of variety" effect is the main driver of aggregate productivity. We also look at a second parametrization with a very low price elasticity of aggregate exports. Naturally total trade growth is smaller as demand does not react to the fall in costs. Interestingly aggregate productivity displays similar gains.

5 Other Theories of Heterogeneity

5.1 Other sources of variation

There are many idiosyncratic factors other than fixed costs that influence the firm's decision to export. For example, firms may have different *variable* costs associated with exports. Perhaps the model's results are tied to the assumption that fixed costs explain all the variation in the data.

There are good reasons to think this is not the case. First, if the additional heterogeneity enters the exporter revenues multiplicatively, then the resulting model—once it is appropriately re-calibrated—is actually isomorphic to ours. The reason is that the model is essentially loglinear. Consider the previous example: on the top of the trade costs τ, the marginal cost of firm's ω exports is $(1+\eta(\omega))$ times more the marginal cost for output sold at the domestic market. We could then rewrite (10) as

$$r^f(\omega) = \left(\frac{\varphi(\omega)}{\tilde{\varphi}_x}\right)^{\theta-1} (1+\eta(\omega))^{-\theta} \frac{R^f}{M_x}.$$

The entry condition (11) would now be

$$\left(\frac{\varphi(\omega)}{\tilde{\varphi}_x}\right)^{\theta-1} (1+\eta(\omega))^{-\theta} \frac{R^f}{M_x} \geq \theta f(\omega).$$

We could, though, simply define $\tilde{f}(\omega) = f(\omega)(1+\eta(\omega))^{\theta}$ and the equilibrium condition is identical.[23] Quantitatively we would calibrate the distribution of both η and f to reproduce

[23]This requires, of course, that the additional heterogeneity is normalized such that it does not change the baseline calibration or the policy experiment.

the variation in the latent variable—in other words, we would calibrate the variable $\tilde{f}(\omega)$ as we did with $f(\omega)$ in Section 4. So as long η and f remain independent of φ, the results would be unchanged.

We do not expect big differences even if the conditions for the isomorphism do not apply. Say we consider firm variation in the trade costs *per unit*, that is, the marginal cost of a firm is now $\eta(\omega) + \tau w/\varphi(\omega)$. In this case the entry condition (11) is not loglinear. We can, though, think of a nth order log-approximation to (10) around the representative firm type, $\tilde{\omega}$. The higher order terms may be important for large reductions of trade costs: they do, though, only on the measure that the variation η interacts with the productivity parameter, φ. This interaction is not captured by the latent variable in Section 3—so non-multiplicative factors affect results only on the measure that they induce departures from the assumption of independence.

5.2 Industry-level Heterogeneity

Here we explore whether the latent heterogeneity stems from between-industry variation. Manufacturing include goods as diverse as tobacco products and machinery. So it is quite reasonable to think that sectors face very different trade costs, both fixed and variable. If the required heterogeneity is present at a very aggregate level then we may be able to capture it with a simple two-sector specification.[24]

We compute the size premium implied by strict sorting for each three-digit NAICS industry code. The procedure is the same we used in Section 2 for manufacturing as a whole. Bernard et al. (2007) report the share of firms exporting in each sector for the year 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 Computer 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 1—provided by the Census.

Table 9 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: the predicted size premium *for each sector* are very large. The reason is that the firm size distribution within a sector is still very skewed, so any strict sorting exercise is bound to return large size premia. 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. As we saw in Table 3 the skewness in total sales is such that even if exporters were way more common that actually observed the implied size premium would remain very large.

[24] This is often done in open macroeconomic models in other to replicate home bias or deviations from the law of one price.

			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 9: Lower Bound for Size Premium Predictions, by Industry.

5.3 Sunk Costs

So far we have considered sources of variation in the export decision in a static framework. However it is well-known that the data displays a fair amount of export hysteresis, giving some support to the presence of sunk costs.[25] While we are not interested in the dynamics of exporting per se, export hysteresis has important implications for the cross-sectional of exporters. In the presence of sunk costs, the exporter status of a firm is not determined alone by its current productivity. An exporter may have no reason to stop exporting for productivity level ψ as it can still cover the per-period fixed costs. The same productivity level, though, may not be high enough to convince a firm to start exporting as it will not

[25] Baldwin(1988) and Baldwin and Krugman (1989) are credited with developing the first theories of export hysteresis. Roberts and Tybout (1997) evaluate empirically the role of sunk costs through reduced form models. More recent work has estimated structural models with sunk costs: see Das et al. (2007) and Ruhl and Willis (2008).

cover the sunk costs. Indeed, there seems to be a basis for this possibility. Alessandria and Hoi (2007, 2008) calibrate models of sunk export costs that match the size and share of exporters, as well as the distribution of firm size in manufacturing. In these papers, though, there are many other sources of heterogeneity so it is not clear the quantitative role of export hysteresis in explaining the cross-sectional facts.

It is thus possible that the history of each firm provides the necessary heterogeneity to match the exporter size premium. A full analysis of this hypothesis is behind the scope of this paper. We include here a simple exercise that suggests that yes, sunk costs can explain the share and size of exporters, but only by shifting the size puzzle from exporters to *new* exporters.

Consider the following variation of the reduced-form model in Section 3. Firm i 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 from 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 $\Psi_r(t_0)$ of previous exporters exit the exporter market, while a fraction $1-\Psi_r(t_1)$ of the previously non-exporters start exporting now. If the share of exporters is constant, we must have that

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

With some manipulation we have that equation (14) gives the share of exporters as a function of the starter and stopper points,

$$s_x = \frac{1-\Psi_r(t_1)}{1-\Psi_r(t_1)+\Psi_r(t_0)}. \tag{15}$$

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 then the average total sales by exporters as

$$\mu_x = (\Psi_r(t_1)-\Psi_r(t_0))\, E\{r|r\in[t_0,t_1]\} + (1-\Psi_r(t_1)+\Psi_r(t_0))\, E\{r|r\geq t_1\}\,. \tag{16}$$

We then solve for the two equations (14) and (16) 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%.[26]

Note that while the sunk cost model breaks the strict sorting between exporters and non-exporters, it now features strict sorting between *new* exporters and non-exporters. Because the entry rate is so small, the threshold for entry is way deep in the tail of the distribution (the 98*th* percentile to be precise). The size premium between new exporters and non-exporters is thus even larger, on the neighborhood of 200. All empirical evidence point to new exporters being significantly smaller than existing exporters—so it looks like to reconcile the sunk cost model with the size of new exporters we will need, again, a lot of independent variation in the entry decision.

6 The Extensive Margin in the Data

The empirical literature has not come to a consensus on the quantitative importance of the extensive margin for aggregate trade patterns. This reflects, in part, that there is no unique concept of the extensive margin: one can define entry and exit at the level of the firm, plant, or product.

Two papers are well-known for arguing that the extensive margin is an important dimension of aggregate data. Hummels and Klenow (2005) find that the extensive margin accounts for 60 percent of the cross-country differences in trade. Broda and Weinstein (2006) find large welfare gains associated with the expansion in import variety for the U.S over the last three decades. Both papers built upon the analysis in Feenstra (1994) and share the focus on the long-run.

In our exercise we have used the measure of the extensive margin in Alessandria and Hoi (2008) so we take some time to discuss their results. Using the census of manufacturers, Alessandria and Hoi (2008) look at the increase in export participation by plants and find that it accounts for half of the export growth in the U.S. from 1987 to 2002. Interestingly, this is quite close to the prediction of the standard Melitz model for the extensive margin. We cannot ignore, though, that the standard version of the model allocates an important role to the extensive margin only by greatly exagerating the role of the economies of scale.

Recently, several papers have taken the position that the extensive margin contributes little to aggregate trade patterns. Besedes and Prusa (2008) argue that new exporting relationships have little impact on long-run export growth because they tend to be very

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

short-lived. Arkolakis, Demidova, Klenow, and Rodriguez-Clare (2008) document a sizeable increase in variety in Costa Rica from 1986 to 1992. They argue, though, that the increase did not translate into large welfare gains because new varieties were imported in small quantities. Armenter and Koren (2008) show that several well-known facts about the extensive margin are the result of the sparsity of the data.

Summarizing, there is no definitive evidence of the quantitative importance of the extensive margin in the determination of aggregate trade patterns. While we may record much activity along the extensive margin—say many new products are exported in a given year—this does not necessarily translate into aggregate patterns—either because these new products are dropped the next year or constitute a very small fraction of total exports.

7 Conclusions

Since Melitz (2003), models with economies of scale and firm-level heterogeneity have become very popular in international trade. These models have been used to explain the characteristics of multi-product exporters and exporters-importers, among others. Most researchers are contempt to show the model qualitative predictions are as documented in the data.

In this paper we have shown that a simple Melitz model cannot match *quantitatively* even the most basic cross-sectional facts about exporters, namely, their frequency and size. The model can be easily reconciled with the data by introducing enough additional sources of variation. However, this is not without implications. In the event of a trade liberalization, the augmented model has a minor role for the extensive margin, and the productivity gains due to reallocation of resources are smaller than in the standard model.

Given the attention the literature has given to the extensive margin, it is quite surprising that the calibrated version of the Melitz model features only small changes in export participation. Our results seem quite robust on this end. In particular, the source of variation at work is not important as long as it is exogenous and independent of size. The solution to this "puzzle" would have to match the weak relationship between size and exporters in the cross-section distribution and yet have export participation responding to the fall in trade costs. We have also seen that while sunk costs can do both, it also implies that new exporters should be much larger than existing exporters.

We find the model in Melitz (2003) compelling so we ask instead whether we are measuring export participation correctly. One particular hypothesis of interest is that the Melitz model fits well the behavior and characteristics of large exporters, but the facts on the extensive margin are driven by a set of small exporters.

References

[1] Alessandria, G. and Choi, H.: 2007, Do Sunk costs of Exporting Matter for Net Export Dynamics? *Quarterly Journal of Economics* **122**(1), 289-336.

[2] Alessandria, G. and Choi, H.: 2008, Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization. Working Paper, Federal Reserve Bank of Philadelphia.

[3] Alessandria, G. and Choi, H.: 2008, Do Falling Iceberg Costs Account for Recent US Export Growth? Working Paper, Federal Reserve Bank of Philadelphia.

[4] Arkolakis, C., Demidova, S., Klenow, P.J. and Rodriguez-Clare, Andres:2008 Endogenous Variety and the Gains from Trade, *American Economic Review* **98**(4), 444-450.

[5] Armenter, R. and Koren, M.: 2008, A Balls-and-Bins Model of Trade. Working Paper, Federal Reserve Bank of New York.

[6] Axtell, R. L.: 2001, Zipf Distribution of U.S. Firm Sizes, *Science* **293**(5536), 1818–1820.

[7] Baldwin, R.: 1998, Hysteresis in Import Prices: The Beachead Effect *American Economic Review* **78**(4), 773–85.

[8] Baldwin, R., and Krugman, P.:1989 Persistent Trade Effects of Large Exchange Rate Shocks *Quarterly Journal of Economics* **104**(4), 635-54.

[9] Baldwin, R. and Harrigan, J.: 2007, Zeros, Quality and Space: Trade Theory and Trade Evidence, NBER Working Paper No. 13214.

[10] Besedes, T. and Prusa, T.:2008 The Role of Extensive and Intensive Margin and Export Growth, Working Paper, Rutgers University.

[11] Bernard, A. B., Eaton, J., Jensen, J. B. and Kortum, S.: 2003, Plants and Productivity in International Trade, *American Economic Review* **93**(4), 1268–1290.

[12] Bernard, A. B. and Jensen, J. B: 1999, Exceptional Exporter Performance: Cause, Effect, or Both?, *Journal of International Economics* **47**(1), 1–25.

[13] Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K.: 2007, Firms in International Trade, *Journal of Economic Perspectives* **21**(3), 105–130.

[14] Bernard, A. B., Jensen, J. B. and Schott, P. K.: 2007, Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods, *in* Dunne, J.B. Jensen and M.J. Roberts (eds.), Producer Dynamics: New Evidence from Micro Data.

[15] Chaney, T.: 2008, Distorted Gravity: the Intensive and Extensive Margins of International Trade, *American Economic Review* **98**(4), 1707-1721.

[16] Das, S., Roberts, M. and Tybout, J.:2007 Market Entry Costs, Producer Heterogeneity, and Export Dynamics *Econometrica* **75**(3), 837-873.

[17] Eaton, J., Eslava, M., Kugler, M. and Tybout, J.: 2007, Export Dynamics in Colombia: Firm-Level Evidence, NBER Working Paper No. 13531.

[18] Eaton, J., Kortum, S. and Kramarz, F.: 2004, Dissecting Trade: Firms, Industries, and Export Destinations, *American Economic Review* **94**(2), 150–154.

[19] Eaton, J., Kortum, S. and Kramarz, F.: 2007, An Anatomy of International Trade: Evidence from French Firms, Working Paper.

[20] Feenstra, R.:1994 New Product Varieties and the Measurament of International Prices, *American Economic Review* **84**(1), 157-177.

[21] Halpern, L., Koren, M. and Szeidl, A.: 2007, Imports and Productivity, Working Paper.

[22] Helpman, E., Melitz, M. and Rubinstein, Y.: 2007, Estimating Trade Flows: Trading Partners and Trading Volumes, *Quarterly Journal of Economics*, forthcoming.

[23] Hopenhayn, H: 1992, Entry, Exit, and Firm Dynamics in Long Run Equilibrium, *Econometrica*, **60**, 1127-1150.

[24] Hummels, D., Klenow, P. J.: 2005, The Variety and Quality of a Nation's Exports, *American Economic Review* **95**(3), 704–723.

[25] Johson, N. L., Kepm, A. W., and Kotz, S.: 2005, *Univariate Discrete Distributions*, John Wiley & Sons.

[26] Mayer, T. and Ottaviano, G.: 2007, The Happy Few: The Internationalization of European Firms, Bruegel Blueprint Series. Volume III.

[27] Melitz, M. J.: 2003, The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity, *Econometrica* **71**(6), 1695–1725.

[28] Roberts, M. and Tybout, J.:1997 The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs *American Economic Review* **87**(4), 545-564.

[29] Ruhl, K. and Willis, J.:2008 New Export Dynamics *NYU Working Paper*.

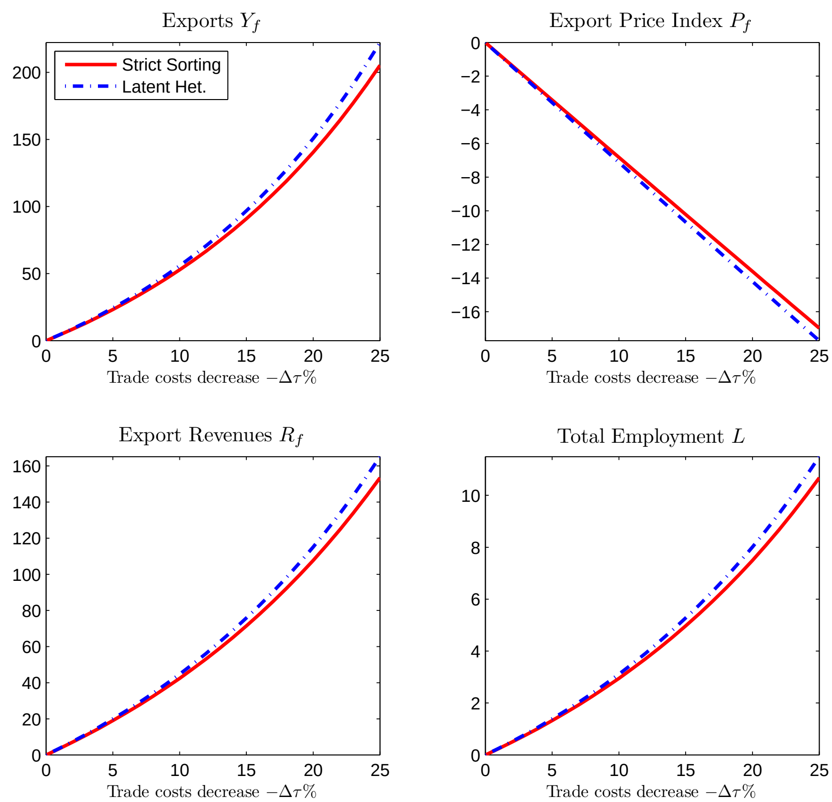

Figure 1: Aggregate Exports *Trade costs decrease in percentage points. All variables in growth rates*

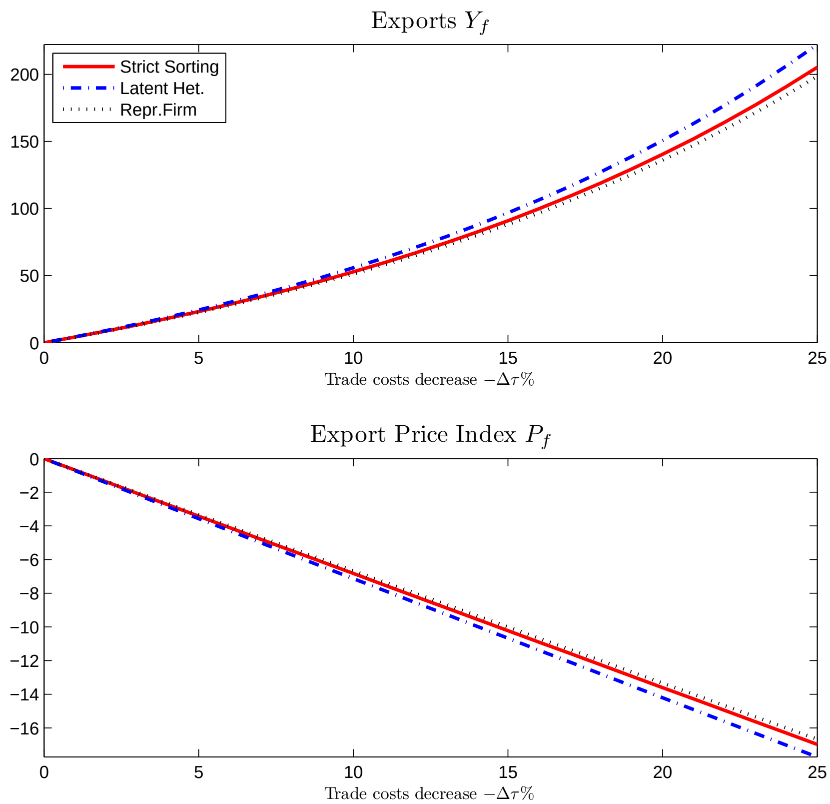

Figure 2: Aggregate Exports - Comparison with the Representative Firm Model *Trade costs decrease in percentage points. All variables in growth rates*

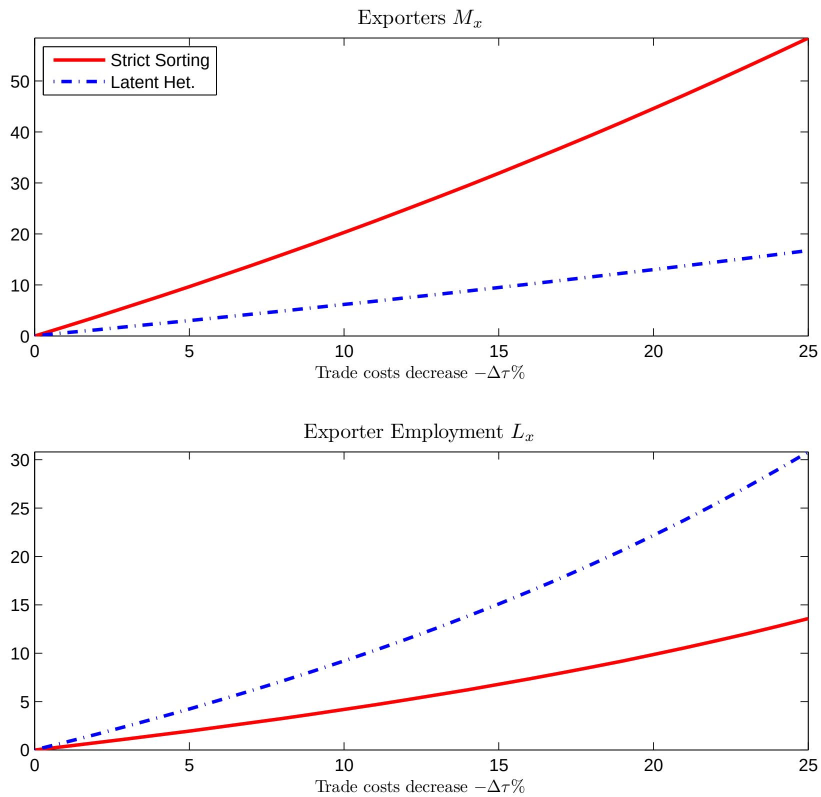

Figure 3: Entry and Employment Growth by Exporters *Trade costs decrease in percentage points. All variables in growth rates*

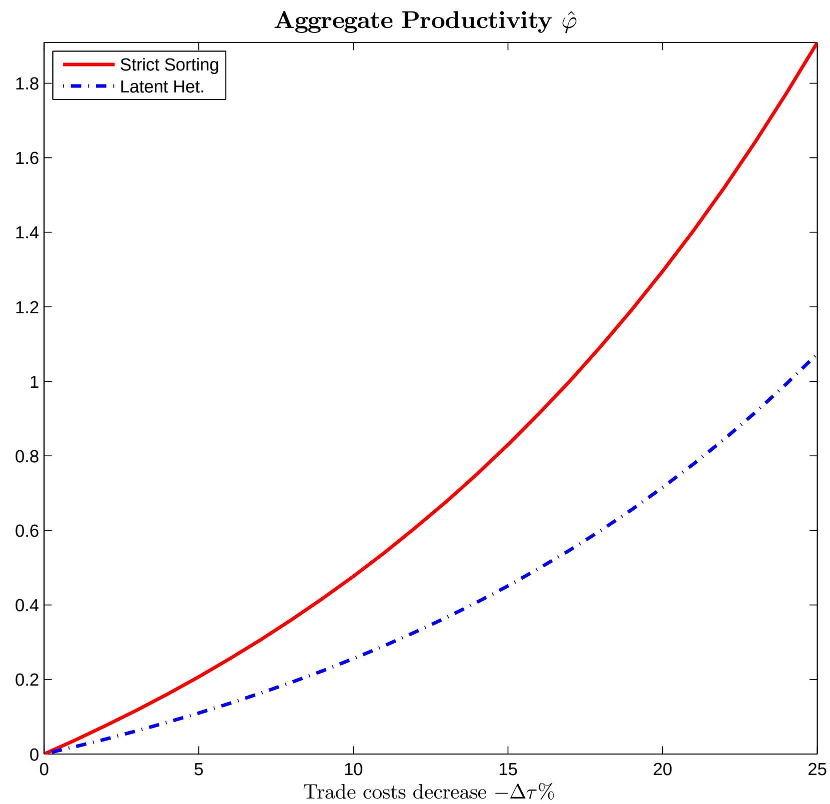

Figure 4: Aggregate Productivity *Trade costs decrease in percentage points. All variables in growth rates*

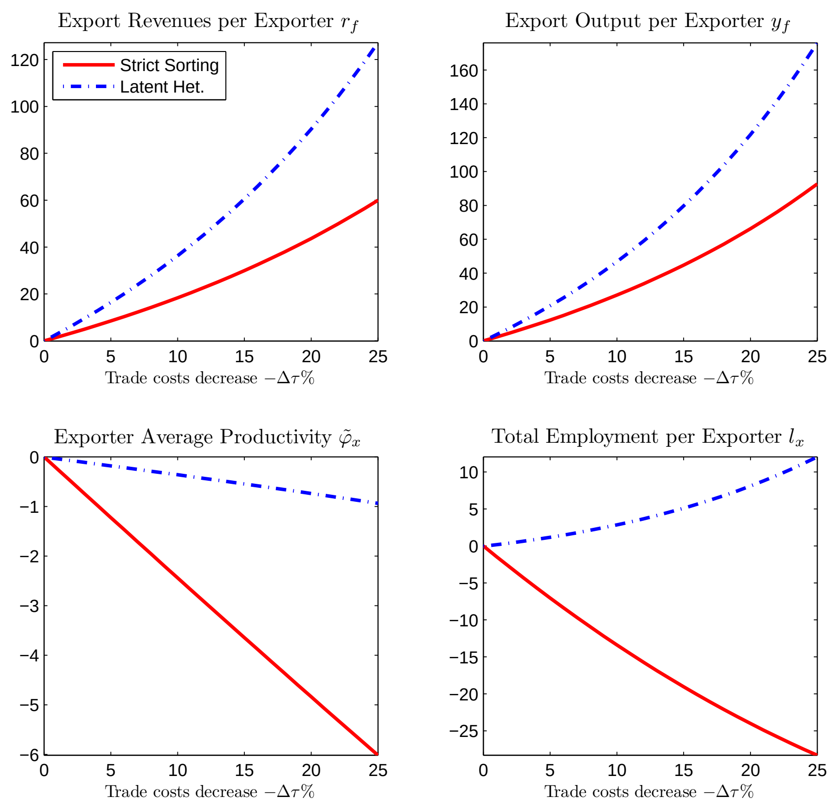

Figure 5: Exporters *Trade costs decrease in percentage points. All variables in growth rates*