DISCUSSION PAPER SERIES

No. 7773

POWER LAWS IN FIRM SIZE AND OPENNESS TO TRADE: MEASUREMENT AND IMPLICATIONS

Julian di Giovanni, Andrei A. Levchenko and Romain Rancière

INTERNATIONAL MACROECONOMICS and INTERNATIONAL TRADE AND REGIONAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP7773.asp

ISSN 0265-8003

POWER LAWS IN FIRM SIZE AND OPENNESS TO TRADE: MEASUREMENT AND IMPLICATIONS

Julian di Giovanni, International Monetary Fund
Andrei A. Levchenko, University of Michigan
Romain Rancière, International Monetary Fund, PSE and CEPR

Discussion Paper No. 7773
April 2010

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7773

April 2010

ABSTRACT

Power Laws in Firm Size and Openness to Trade: Measurement and Implications

Existing estimates of power laws in firm size typically ignore the impact of international trade. Using a simple theoretical framework, we show that international trade systematically affects the distribution of firm size: the power law exponent among exporting firms should be strictly lower in absolute value than the power law exponent among non-exporting firms. We use a dataset of French firms to demonstrate that this prediction is strongly supported by the data. While estimates of power law exponents have been used to pin down parameters in theoretical and quantitative models, our analysis implies that the existing estimates are systematically lower than the true values. We propose two simple ways of estimating power law parameters that take explicit account of exporting behavior.

JEL Classification: F12 and F15
Keywords: firm size distribution, international trade and power laws

Julian di Giovanni
Research Department
International Monetary Fund
700 19th Street N.W.
Washington, DC 20431
USA

Email: jdigiovanni@imf.org

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=163552

Andrei A. Levchenko
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, MI 48109
USA

Email: alev@umich.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=161834

Romain Rancière
International Monetary Fund
Research Department
700 19th Street NW
Washington, DC 20431
USA

Email: RRanciere@imf.org

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=153503

Submitted 30 March 2010

We are grateful to Xavier Gabaix for detailed suggestions, and Michael Aubourg for superb research assistance. This paper was completed while Romain Rancière was visiting INSEE whose hospitality – and in particular the help of Pauline Givord – is gratefully acknowledged. Romain Rancière thanks CEPREMAP for financial support. The views expressed in this paper are those of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management.

1 Introduction

Many relationships in economics appear to be governed by power laws. A distributional power law is a relationship of the type: $\Pr(X > x) = Cx^{-\zeta}$ where $\Pr(X > x)$ is the probability that a random variable X is greater than x, and C and ζ are constants. Power laws arise in a variety of contexts, such as the distribution city size (Zipf 1949), income (Champernowne 1953), firm size (Axtell 2001), and sectoral trade flows (Hinloopen and van Marrewijk 2006, Easterly, Reshef and Schwenkenberg 2009).[1]

The literature has emphasized the importance of the precise value of the power law exponent, ζ. For instance, for the distribution of firm size, Axtell (2001) reports a range of estimates between 0.996 and 1.059, very precisely estimated with standard errors between 0.054 and 0.064.[2] The literature has sought to both explain why ζ is close to 1 – a phenomenon known as Zipf's Law – and to explore its implications in a variety of contexts. It has been argued that Zipf's Law will arise when the variable of interest – be it city, or firm size – follows a geometric Brownian motion (Gabaix 1999, Luttmer 2007, Rossi-Hansberg and Wright 2007). At the same time, the precise magnitude of the power law exponent has been shown to matter for such different phenomena as macroeconomic fluctuations (Gabaix 2009b, di Giovanni and Levchenko 2009a), regulation of entry (di Giovanni and Levchenko 2009b), and executive compensation (Gabaix and Landier 2007).

This paper revisits the power law in the distribution of firm size in the context of international trade. We first set up a simple version of the Melitz (2003) model of production and trade, adopting the common assumption that the distribution of firm productivities is Pareto. This model is naturally suited to studying the firm size distribution because of its emphasis on heterogeneous firms. The Melitz-Pareto framework delivers a power law in firm size. However, it also predicts that in the presence of international trade, the power law exponent in the distribution of firm size is not constant. Because larger firms are more likely to export, and the more productive the firm, the more markets it serves, we would expect the estimated power law exponent to be lower in absolute value among exporting firms compared to the non-exporting ones. In other words, in the presence of international trade, power law estimates that do not take into account international trade could be misleading regarding the deep parameters of the economy.[3]

We evaluate these predictions of the Melitz-Pareto model using the data on production and exports for a large sample of French firms. In the full sample that includes all firms, the power law in firm size is strikingly similar to what Axtell (2001) found for the census of U.S. firms. The estimated power law exponent and the fit of the relationship are both nearly identical. However, when we

[1] See Gabaix (2009a) for a recent survey.

[2] The fit of this relationship is typically very close. The R^2's reported by Axtell (2001) are in excess of 0.99.

[3] This paper focuses on power law estimation because power laws appear to be the best description of observed firm size distributions (Luttmer 2007). However, the qualitative mechanisms we highlight apply to any other underlying distribution of firm size.

separate the firms into exporting and non-exporting ones, it turns out that in the exporting sample, the power law coefficient is consistently lower, while in the non-exporting sample, consistently higher than in the full sample of firms. This difference is present across all estimators, and highly statistically significant.

We then provide several pieces of supporting evidence that international trade changes the distribution of firms size in ways predicted by theory. First, we show that the power law exponent for exporting firms converges to the power law exponent for domestic firms as we restrict the sample to larger and larger exporters. And second, the power law coefficients exhibit the same pattern at the disaggregated industry level as well. Furthermore, at sector level the differences between power law coefficients are larger in sectors that are more open to trade, a striking regularity that is consistent with the theoretical intuition developed in the paper. All of these pieces of evidence lend empirical support to the main idea of the paper: international trade systematically changes the distribution of firm size, and inference that does not take that into account will likely lead to biased estimates.

One of the reasons empirical power law estimates are important is that they can be used to pin down crucial parameters in calibrated heterogeneous firms models (see, among many others, Helpman, Melitz and Yeaple 2004, Chaney 2008, Buera and Shin 2008, di Giovanni and Levchenko 2009a). At the same time, quantitative results often depend very sharply on the precise parameter values that govern the distribution of firm size. Di Giovanni and Levchenko (2009b) show that welfare gains from reductions in fixed costs are an order of magnitude lower, and gains from reductions in variable costs an order of magnitude higher in a model calibrated to Zipf's Law compared to the counterfactual case in which $\zeta = 2$ instead. We return to the Melitz-Pareto model, and propose two alternative ways of estimating the power law parameters that are internally consistent with the canonical heterogeneous firms model of trade. The first is to use a sample of only non-exporting firms. The second is to use only domestic sales to estimate the power law parameter.

We are not the first to provide parameter estimates for the firm size distribution that explicitly account for international trade. Eaton, Kortum and Kramarz (2008) set up a multi-country heterogeneous firms model and estimate a set of model parameters with Simulated Method of Moments using the data on French firms.[4] The advantage of our approach is simplicity. The alternative estimation strategies proposed here are very easy to implement and do not require any additional modeling or estimation techniques. All they rely on is an appropriate modification of the sample or variables used in estimation. Our approach thus substantially lowers the barriers to obtaining reliable power law estimates, and can be applied easily in many contexts.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework.

[4]See Arkolakis (2008, 2009) for related theoretical treatments.

Section 3 describes the dataset used in the analysis and the methodology for estimating power laws. Section 4 describes the results. Section 5 concludes.

2 Theoretical Framework

Consider a model in the spirit of Melitz (2003). Consumers in market n maximize

$$\begin{aligned}
&\max \left[\int_{J_n} c_{ni}^{\frac{\varepsilon-1}{\varepsilon}} di \right]^{\frac{\varepsilon}{\varepsilon-1}} \\
&s.t. \\
&\int_{J_n} p_{ni} c_{ni} di = Y_n,
\end{aligned}$$

where c_{ni} is consumption of good i in country n, p_{ni} is the price of this good, Y_n is total expenditure in the economy, and J_n is the mass of varieties consumed in country n, coming from all countries. It is well-known that demand for variety i in country n is equal to

$$c_{ni} = \frac{Y_n}{P_n^{1-\varepsilon}} p_{ni}^{-\varepsilon}, \tag{1}$$

where P_n is the ideal price index in this economy, $P_n = \left[\int_{J_n} p_{ni}^{1-\varepsilon} di \right]^{\frac{1}{1-\varepsilon}}$.

Each country has a mass $\bar{I}_n$ of potential (but not actual) entrepreneurs. For what we wish to illustrate, it does not matter whether $\bar{I}_n$ solves a free entry condition as in Melitz (2003) and Helpman et al. (2004), or $\bar{I}_n$ is simply a fixed endowment as in Eaton and Kortum (2005) and Chaney (2008). Each potential entrepreneur can produce a unique CES variety, and thus has some market power. There are both fixed and variable costs of production and trade. At the beginning of the period, each potential entrant $i \in [0, \bar{I}_n]$ in each market n learns its type, which is the marginal cost a_i. On the basis of this cost, each entrepreneur in country n decides whether or not to pay the fixed cost of production f, and which, if any, export markets to serve. Let ω_n be the price of the input bundle in country n. Technology is linear in the input bundle: a firm with marginal cost a_i must use a_i units of the input bundle to produce one unit of its final output. As an example, if there is only one factor of production, labor, then the cost of the input bundle is simply the wage $\omega_n = w_n$. Alternatively, in an economy with both labor and capital and a Cobb-Douglas production function, the cost of the input bundle is (up to a constant) $\omega_n = w_n^{\alpha} r_n^{1-\alpha}$, where r_n is the return to capital in country n.

Firm i from country n selling to its domestic market faces a demand curve given by (1), and has a marginal cost $\omega_n a_i$ of serving this market. As is well known, the profit maximizing price is a constant markup over marginal cost, $p_{ni} = \frac{\varepsilon}{\varepsilon-1} \omega_n a_i$, and the quantity supplied is equal to $\frac{Y_n}{P_n^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon-1} \omega_n a_i \right)^{-\varepsilon}$. Domestic sales D_i are given by:

$$D_i = \frac{Y_n}{P_n^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon - 1} \omega_n a_i \right)^{1-\varepsilon} = M_n \times B_i,$$

where $M_n \equiv \frac{Y_n}{P_n^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon-1}\omega_n\right)^{1-\varepsilon}$ is a measure of the size of domestic demand, which is the same for all firms, and $B_i \equiv a_i^{1-\varepsilon}$ is the firm-specific (but not market-specific) productivity-cum-sales term. Conveniently, the total variable profits are a constant multiple of D_i:

$$\pi_D^V(a_i) = \frac{D_i}{\varepsilon}.$$

The firm sells domestically only if its variable profits cover the fixed costs of setting up production: $D_i/\varepsilon \geq f$. This defines the minimum size of the firm observed in this economy, $\underline{D} = \varepsilon f$, as well as the cutoff marginal cost above which the firm does not operate:

$$a_{nn} = \left(\frac{M_n}{\varepsilon f}\right)^{\frac{1}{\varepsilon-1}}.$$

To start exporting from country n to country m, firm i must pay the fixed cost κ_{mni} that varies by firm, and an iceberg per-unit cost of $\tau_{mn} > 1$.[5] We normalize the iceberg cost of domestic sales to one. It is easy to verify that export sales by firm i can be expressed as $M_m^* B_i$, where $M_m^* = \frac{Y_m}{P_m^{1-\varepsilon}} \left(\frac{\varepsilon}{\varepsilon-1}\tau_{mn}\omega_n\right)^{1-\varepsilon}$ is m's market size *from the perspective of the firm that exports there from* n, and B_i is defined above. As before, variable profits from exporting are equal to $M_m^* B_i/\varepsilon$, and firm i exports only if

$$\frac{M_m^* B_i}{\varepsilon} \geq \kappa_{mni}. \tag{2}$$

The fixed cost of exporting to a foreign market is stochastic and varies from firm to firm, as in Eaton et al. (2008). Thus, there will not be a single "exporting cutoff," above which all firms export, and below which none do. Instead, our formulation delivers both exporting and non-exporting firms with the same exact productivity, or equivalently, domestic sales.

The production structure of the economy is pinned down by the number of firms from each country that enter each market. Using this information, one can express P_n in terms of Y_n and ω_n, for all countries. The model can be closed by solving for the Y_n's and ω_n's. To do this, we must impose balanced trade and factor market clearing for each country, as well as a free entry condition if we assume that $\bar{I}_n$ is endogenous. We do not pursue the full solution to this model here, since it is not necessary to derive our main predictions (for an example, see di Giovanni and Levchenko 2009b). Instead, we describe the analytical power law relationship in the distribution of firm size.

2.1 Power Law in Firm Size, With and Without International Trade

In this section, we first demonstrate the power law in an autarkic economy, and then discuss how the distribution of firm size is affected by international trade. Firm sales, S_i, in the economy follow

[5] That is, the firm in country n must ship $\tau_{mn} > 1$ units to country m in order for one unit of the good to arrive there.

a power law if their distribution is described by:

$$\Pr(S_i > s) = Cs^{-\zeta}. \tag{3}$$

We postulate that B_i follows a Pareto distribution with exponent ζ. Under some conditions (e.g., Gabaix 1999, Luttmer 2007), B_i comes from a random growth model, which yields a value of ζ close to 1. It turns out that in our model this is equivalent to assuming that firm productivity is Pareto, but with a different exponent. To see this, suppose that firm productivity, $1/a$ ~Pareto(b, θ), and thus its cdf is given by: $\Pr(1/a < y) = 1 - \left(\frac{b}{y}\right)^{\theta}$. In the autarkic economy, where $S_i = D_i$, the power law follows:

$$\Pr(S_i > s) = \Pr(M_n B_i > s) = \Pr\left(a_i^{1-\varepsilon} > \frac{s}{M_n}\right) =$$

$$\Pr\left(\left(\frac{1}{a}\right)^{\varepsilon-1} > \frac{s}{M_n}\right) = \Pr\left(\frac{1}{a} > \left(\frac{s}{M_n}\right)^{\frac{1}{\varepsilon-1}}\right) = \left(b^{\varepsilon-1}M_n\right)^{\frac{\theta}{\varepsilon-1}} s^{-\frac{\theta}{\varepsilon-1}}, \tag{4}$$

satisfying (3) for $C = \left(b^{\varepsilon-1}M_n\right)^{\frac{\theta}{\varepsilon-1}}$ and $\zeta = \frac{\theta}{\varepsilon-1}$. The model-implied distribution of sales is depicted in Figure 1. In addition, this calculation shows that S_i ~Pareto$(b^{\varepsilon-1}M_n, \frac{\theta}{\varepsilon-1})$.

This relationship – the power law exponent constant and equal to $\frac{\theta}{\varepsilon-1}$ – holds true in autarky, and also among non-exporting firms in the trade equilibrium. But how does exporting behavior change the firm size distribution? We describe two mechanisms by which exporting tilts the power law relationship systematically to make it flatter (more right-skewed). The first relies on entry into progressively more foreign markets. The second, on stochastic export market entry costs that vary by firm. In the second case, it is possible to obtain a number of analytical results about the distribution of firm sales, and show that it is systematically affected by exporting behavior.

Consider first the case of multiple export markets. For simplicity, let $\kappa_{mni} = \kappa_{mn}\ \forall i$. In the presence of firm heterogeneity and fixed costs of entering export markets, there is a hierarchy of firms in their export market participation. For each potential export market m, equation (2) defines a partition of firms according to how many markets the firm serves. From this expression it is clear that a firm will first serve a market that is closer (low τ_{mn} and κ_{mn}), has larger size (high Y_m), and lower competition (high P_m). We can order potential export destinations according to how productive a firm needs to be in order to export there. This is illustrated in Figure 2, which orders firms according to marginal cost, with more productive firms closer to the origin. Since each firm in the home country faces the same aggregate conditions and trade costs in each trading partner, if a firm exports to any market, it also exports to all markets served by the less productive firms.

What this implies for the distribution of firm sales is illustrated in Figure 1. For all the firms that only sell to the domestic market, the power law is still a straight line with the same slope as what we derived for autarky, $\frac{\theta}{\varepsilon-1}$. However, participation in export markets results in a series of parallel shifts in this cumulative distribution function, one for each additional export market that firms might enter. Because the more productive a firm is, the more markets it sells to, the distribution of firm size becomes more fat-tailed.

The second mechanism that tilts the power law in firm size is the stochastic fixed costs of exporting that vary across firms. To obtain a number of analytical results, we assume that there is only one export market m. Clearly, this should be thought of as a composite of the potential for global sales of the company. This framework would thus apply particularly well if there is a considerable fixed cost common to entering any and all foreign markets, but once a firm exports to one country, it finds it much easier to export to others.[6] The notation will be simplified considerably if we define ϕ to be the ratio of the foreign market size relative to the domestic one:

$$\phi \equiv \frac{M_m^*}{M_n} = \tau_{mn}^{1-\varepsilon} \frac{Y_m}{Y_n} \left(\frac{P_n}{P_m} \right)^{1-\varepsilon}.$$

In that case, the exporting decision condition (2) can be written as a function of domestic sales:

$$\frac{\phi D_i}{\varepsilon} \geq \kappa_i,$$

where to streamline notation we omit the mn-subscripts: $\kappa_i = \kappa_{mni}$. Denote by H the pdf of $\kappa / \left(\frac{\phi}{\varepsilon} \right)$:

$$H(x) = \Pr\left(\kappa_i \leq \frac{\phi x}{\varepsilon} \right).$$

We will call $H(x)$ the "export probability function": a firm with domestic size D_i exports with probability $H(D_i)$, which is weakly increasing in D_i. If the firm exports, the exports are $X_i = M_n^* B_i = \phi D_i$. Hence,

$$X_i = \begin{cases} 0 & \text{if } \frac{\phi D_i}{\varepsilon} < \kappa_i; \quad \text{probability } 1 - H(D_i) \\ \phi D_i & \text{if } \frac{\phi D_i}{\varepsilon} \geq \kappa_i; \quad \text{probability } H(D_i). \end{cases} \tag{5}$$

The total (worldwide) sales of the firms are $S_i = D_i + X_i$, which implies that

$$S_i = \begin{cases} D_i & \text{if } \frac{\phi D_i}{\varepsilon} < \kappa_i; \quad \text{probability } 1 - H(D_i) \\ (1+\phi) D_i & \text{if } \frac{\phi D_i}{\varepsilon} \geq \kappa_i; \quad \text{probability } H(D_i). \end{cases} \tag{6}$$

The distributions of domestic sales, export sales, and total sales are described in the following proposition. The proof is presented in the Appendix.

Proposition 1 *The densities of domestic sales D_i, exports X_i (when they are nonzero), and worldwide sales S_i are:*

$$p_D(x) = k x^{-\zeta-1} 1_{x > \underline{D}}$$

[6] Hanson and Xiang (2008) use U.S. motion picture exports to provide empirical evidence that fixed exporting costs are global rather market-specific.

$$p_X(x) = Kx^{-\zeta-1}H\left(\frac{x}{\phi}\right)1_{x>\phi\underline{D}} \tag{7}$$

$$p_S(x) = kx^{-\zeta-1}\left[1 - H(x) + H\left(\frac{x}{1+\phi}\right)(1+\phi)^{\zeta}\right]1_{x>(1+\phi)\underline{D}} + kx^{-\zeta-1}1_{\underline{D}<x<(1+\phi)\underline{D}}, \tag{8}$$

where $k = \zeta\underline{D}^{\zeta}$, K is a constant ensuring $\int p_X(x)\,dx = 1$, and $1_{\{.\}}$ is the indicator function.

In other words, when the underlying distribution of productivity, and therefore domestic sales, is Pareto, the presence of exporting behavior implies that the distribution of total sales, as well as export sales, is systematically different. Thus, the standard practice of estimating power laws in firm size based on total firm sales will not yield reliable estimates of the underlying power law parameter, ζ, which is in turn often used to calibrate the model parameter combination, $\theta/(\varepsilon-1)$. As is evident from equation (8) that describes the distribution of total sales, fitting instead the simple power law relationship (3) will not yield the correct estimate of the power law exponent.

As an example, suppose that the distribution of fixed exporting cost κ_i is itself Pareto, with some upper truncation: $H(x/\phi) = k''x^{\alpha}$, for $x < x^*$ and some k'', and $H(x/\phi) = k''(x^*)^{\alpha}$ for $x > x^*$. Then, equation (7) implies that the distribution of export sales is given by:

$$p_X(x) \propto \begin{cases} x^{-\zeta-1+\alpha} & \text{for } x < x^* \\ x^{-\zeta-1} & \text{for } x \geq x^*, \end{cases}$$

and thus the power law exponent of X is:

$$\zeta_X(x) = \begin{cases} \zeta - \alpha & \text{for } x < x^* \\ \zeta & \text{for } x \geq x^*. \end{cases}$$

When H has a high slope, the Pareto exponent of X is lower than that of domestic sales: there are fewer small exports, due to the selection effect coming from the fixed cost of exporting. However, in the region where the H function "saturates," the local Pareto exponent of exports converges to the exponent of domestic sales. Such a possibility is depicted in Figure 3. We will provide evidence consistent with this auxiliary prediction of the model in Section 4.1 below.

In summary, we have now described two mechanisms by which the estimated slope of the power law among exporting firms is systematically lower in absolute value than the slope among the firms that sell only to the domestic market, at least in some regions. The preceding discussion also identifies two ways of estimating the power law parameter that take explicit account of exporting behavior. The first, suggested by Figures 1 and 3, is to use only non-exporting firms in the estimation of ζ. The second, suggested by equation (4) and Proposition 1, is to estimate the power law exponent on the firms' domestic sales only. We now test the theoretical predictions using a comprehensive dataset on sales and exports of French firms, and present the results of implementing the two simple alternative ways of estimating ζ that correct for the exporting behavior.

3 Data and Empirical Methodology

3.1 Data and Estimation Sample

To carry out the empirical analysis, we start with a comprehensive dataset that covers the entire universe of French firms for 2006. The data are based on the mandatory reporting of firms' income statements to tax authorities. What is crucial for our analysis is that each firm has to report the breakdown of its total sales between domestic sales and exports. In total, the dataset includes 2,182,571 firms, of which 194,444 (roughly 9%) are exporters.

The exhaustive nature of the dataset implies that there are many observations for very small firms, whose economic activity is practically nil, and are thus both uninformative and uninteresting to the researcher. For our purposes, this is important because it is widely recognized that many power laws fit the data well only above a certain minimum size threshold (Axtell 2001, Luttmer 2007). Thus, a power law may not be a good description of the size distribution of very small firms. To address this problem, we follow the common practice in the literature and pick the lower cutoff based on visual inspection (Gabaix 2009a). The minimum sales cutoff we select is Euro750,000 of annual sales.[7] This cutoff is quite low: though it results in the removal of a large number of firms from the dataset, the dropped firms account for only 7.7% of total sales. Our results are robust to a variety of cutoffs, including sales cutoffs as low as Euro100,000.

Since the focus of the paper is on how exporting behavior changes the firm size distribution, we also omit non-tradeable sectors. The conventional approach to isolate the tradeable sector is to focus exclusively on manufacturing. However, many sectors in agriculture, mining, and services report non-trivial export sales as well. To broaden the definition of the tradeable sector, we only drop industries for which total exports are less than 5% of total sales.[8] The remaining sample (of industries with exports of at least 5% of total sales) includes all of the manufacturing, agricultural, and natural resource sectors, as well as some services such as consulting and wholesale trade. The non-tradeable sector accounts for 51% of the total sales in the universe of French firms, so this is a large reduction in the sample. However, all of our main results are robust to including the non-traded sector in the estimation.

Table A1 provides descriptive statistics for the final sample of domestic sales and exports. The final sample includes 157,084 firms, 67,078 of which are exporters (42.7%). An average exporter's total sales are approximately four times larger than those of an average non-exporter (Euro 24.2

[7]This value also has an institutional justification: below this sales threshold, firms have the option of filing their information according to simplified reporting requirements, while above it the firms must provide comprehensive accounting data to tax authorities.

[8]This leads to dropping of Construction, Retail Trade, Real Estate, Financial Services, Post and Telecommunications, Business Services, Hotels and Restaurants, Recreational, Cultural, and Athletic Activities, Health and Social Action, Public Administration, and Unions and Extra-Territorial Activities, most of which are plausibly non-tradeable sectors.

million vs. Euro 6.4 million).

3.2 Empirical Methodology

In order to obtain reliable estimates, this paper uses three standard methods of estimating the slope of the power law ζ. We now describe each in turn. The first method, based on Axtell (2001), makes direct use of the definition of the power law (3), which in natural logs becomes:

$$\ln\left(\Pr(S_i > s)\right) = \ln\left(C\right) - \zeta \ln\left(s\right). \tag{9}$$

For a grid of values of sales s, the estimated probability $\Pr(S_i > s)$ is simply the number of firms in the sample with sales greater than s divided by the total number of firms. We then regress the natural log of this probability on $\ln(s)$ to obtain our first estimate of ζ. Following the typical approach in the literature, we do this for the values of s that are equidistant from each other on log scale. This implies that in absolute terms, the intervals containing low values of s are narrower than the intervals at high values of s. This is done to get a greater precision of the estimates: since there are fewer large firms, observations in small intervals for very high values of s would be more noisy.

The second approach starts with the observation that the cdf in (3) has a probability density function

$$f(s) = C\zeta s^{-(\zeta+1)}. \tag{10}$$

To estimate this pdf, we divide the values of firm sales into bins of equal size on the log scale, and compute the frequency as the number of firms in each bin divided by the width of the bin. Since in absolute terms the bins are of unequal size, we regress the resulting frequency observations on the value of s which is the geometric mean of the endpoints of the bin (this approach follows Axtell 2001). Note that the resulting coefficient is an estimate of $-(\zeta+1)$, so when reporting the results we subtract 1 from the estimate to recover a value of ζ.

Alternatively, we simply regress the natural log of $(\text{Rank}_i - 1/2)$ of each firm in the sales distribution on log of its sales:

$$\ln\left(\text{Rank}_i - \frac{1}{2}\right) = \text{Constant} + \widehat{\zeta}_{LR} \ln S_i + \epsilon_i,$$

where i indexes firms, and S_i are sales. This is the estimator suggested by Gabaix and Ibragimov (2008). The power law coefficient $\widehat{\zeta}_{LR}$ has a standard error of $|\widehat{\zeta}_{LR}|(N/2)^{-1/2}$, where N is the sample size. This standard error is reported throughout for this estimator. As we show below, in practice the three estimators deliver remarkably similar results.

4 Results

We begin by replicating the conventional results in the literature that pool all firms and do not consider exporting behavior. Table 1 reports estimates of the power law in firm size using the three methods outlined above. We can see that the fit of the data is quite similar to that reported in existing studies: the R^2's are all 99% or above. The point estimates of ζ are close to what Axtell finds for the U.S.. Figure 4 presents the results graphically. The two panels show the cdf and pdf of the power law in firm size. For convenience, each plot also reports the regression equation, the fit, and the number of firms underlying the estimation. It is clear that in these data on French firms, the power law holds about as well as in the existing studies.

We now present the main results of the paper, which show that the power law in fact differs significantly between exporting and non-exporting firms as implied by theory. Table 2 reports estimates of the power law for exporting and non-exporting firms separately. Columns (1), (3), and (5) present the results of estimating power laws on exporting firms only, while columns (2), (4), and (6) for non-exporting firms only. Once again, we report the estimate, standard error, R^2, and the number of firms for each estimate. Comparing the columns for exporters and non-exporters, we can see the main result clearly. In every case, the exponent of the power law among exporting firms is lower than that for non-exporting firms. In other words, the size distribution for exporting firms is systematically more fat-tailed than the size distribution of the non-exporting ones. The difference is highly statistically significant. Column (7) reports the t-statistic for the difference between the coefficients in columns (5) and (6), and shows that they are in fact significantly different.[9] As we argue above, while this is exactly what theory predicts, this aspect of the firm size distribution has until now been ignored in the empirical literature.

A couple of other features of the results are worth noting. First, for all estimation methods, the full-sample coefficient from Table 1 is always between the exporting and the non-exporting sample coefficients in Table 2. This is exactly what we would expect. Second, we can see that the fit of the power law relationships in both subsamples is still quite good. The range of the R^2's is 0.972 to 0.999. Once again, as predicted by theory, the simple power law describes the distribution of exporting firms slightly less well, with R^2's systematically lower for exporting than non-exporting firms.

Figure 5 presents the results graphically. Panels (a) and (b) report the cdf and the pdf, respectively. The cdf for exporting firms is everywhere flatter and above the cdf for non-exporting ones. At each size cutoff, there are more larger firms that export than those that do not. The pdf plot

[9] We do not report the t-tests for whether the CDF and PDF coefficients are statistically significantly different from each other, as those are "heuristic" estimators without a well-developed statistical model of standard errors. Nonetheless, a simple t-test based on the coefficients and standard errors reported in Table 2 always rejects equality between the CDF and PDF coefficents in Columns (1) compared to (2) and (3) compared to (4).

conveys the same message.

4.1 Size Distribution For Exporting Firms: Nonlinearity and Saturation

One of the predictions of the two trade mechanisms outlined above is that for big enough exporting firms, the power law coefficient will converge back to the "true" ζ. This would be the case either because eventually all the big firms will have entered all markets, as in the lower right-hand corner of Figure 1, or because the stochastic fixed exporting costs have some upper bound, so that (nearly) every firm above a certain productivity cutoff finds it profitable to export. To check for this possibility, we re-estimated the power law in total sales for exporting firms while moving the lower cutoff. Theory predicts that as we constrain the sample to bigger and bigger exporting firms, the power law coefficient will get larger and larger in absolute value, converging eventually to ζ, the coefficient for domestic firms. Figure 6 depicts the estimates of the power law in firm size for exporting firms as a function of the lower sales cutoff. That is, as we move to the right on the horizontal axis, the power law is estimated on subsamples of bigger and bigger exporting firms.

In line with theory, as we move the cutoff upwards, the power law coefficient becomes larger and larger in absolute value, converging to 1.029, the non-exporting firms coefficient. The bottom panel of the figure reports the R^2 of the power law fit. In the sample of exporting firms, as the estimated coefficient converges to the conceptually correct value, the fit of the power law estimate improves as well, from 0.97 to more than 0.99. This is once again consistent with theory: because exporting behavior induces deviations from a precise power law in the sample of exporting firms, the fit is less tight initially. However, a power law becomes a better and better description of the exporters' size distribution as we constrain the sample to larger and larger firms.

4.2 Sector-Level Evidence

The model in Section 2 can be interpreted as describing an individual sector rather than the whole economy. Thus, we should expect to see at sector level the same patterns found above for the aggregate. Examining the predictions of the heterogeneous firms model at the industry level can also reassure us that the results are not driven by compositional effects. In addition, we can exploit differences in trade openness by sector to provide further evidence consistent with theory.

Table 3 reports the results of estimating the power laws in firm size for exporters and non-exporters by sector, the industry-level equivalent of Table 2. For compactness, we only report the log-rank-log-size estimates, though the PDF and the CDF estimators deliver virtually identical results. It is clear that the effects we illustrate in the economy-wide data are present at the sector level. In every one of the 25 tradeable sectors, the power law coefficient for non-exporting firms is larger in absolute value than the coefficient for exporting firms. In 22 out of 25 of these sectors, this difference is statistically significant. In fact, if anything the difference between the two coefficients is slightly more pronounced at the sector level: while in the aggregate results of Table 2 the difference between the two coefficients is 0.29, at sector level the mean difference between these coefficients is 0.35.

An auxiliary prediction of theory is that the exporting behavior will induce greater deviations in the value of power law exponents in sectors that are more open to trade. Figure 7(a) investigates this prediction. On the vertical axis is the percentage difference between the power law coefficient for domestic sales and the traditional power law coefficient as estimated on the total sales of all firms in the sector: $\frac{\zeta_{dom}-\zeta_{total}}{\zeta_{dom}}$, where ζ_{dom} is the coefficient obtained from fitting a power law on domestic sales only, and ζ_{total} is the exponent of the power law for total sales. On the horizontal axis is the overall sector openness: the ratio of exports to total sales in the sector. For ease of comparison, the non-tradeable sectors are denoted by hollow dots, and the tradeable sectors by solid dots. The figure also reports the OLS fit through the data. The underlying power law coefficients and exports to sales ratios are reported in Appendix Table A2.

The relationship is striking: the more open the sector, the greater is the difference between the conventional power law coefficient – estimated on total sales – and the power law coefficient estimated on domestic sales only. The simple bivariate correlation between these two variables is a remarkable 0.92. Similarly, Figure 7(b) plots the relationship between sector openness and the difference between the conventional power law coefficient and the coefficient estimated on non-exporting firms only: $\frac{\zeta_{nonex}-\zeta_{total}}{\zeta_{nonex}}$, where ζ_{nonex} is the power law coefficient estimated on the sales of nonexporting firms only, reported in column 5 of Table 3. Once again, the positive and significant relationship is quite pronounced: the correlation between the two variables is 0.72.

The variation in the impact of exporting behavior on power law coefficients across sectors thus provides remarkable supporting evidence that international trade changes the distribution of firm size in systematic and predictable ways.

4.3 Corrections and Quantitative Implications

What is the economic significance of these differences? In section 2.1, we show analytically that the power law exponent in firm size, ζ, is related to $\frac{\theta}{\varepsilon-1}$, where θ is the parameter in the distribution of productivities, and ε is the elasticity of substitution between goods. As such, estimates of ζ have been used in empirical and quantitative analyses to pin down this combination of parameters (see, e.g. Helpman et al. 2004, Chaney 2008). Above, we showed that estimating power laws without regard for exporting behavior implies that the estimated ζ does not actually equal $\frac{\theta}{\varepsilon-1}$.

However, it is still possible to recover a reliable estimate of $\frac{\theta}{\varepsilon-1}$ in two simple ways suggested by theory. The first relies on the observation that in the sample of non-exporting firms only, ζ does in fact correspond to $\frac{\theta}{\varepsilon-1}$. Thus, in any dataset that contains explicit information on whether or not a firm exports (without necessarily reporting the value of exports), $\frac{\theta}{\varepsilon-1}$ can be estimated by restricting attention to non-exporting firms.

The non-exporters estimates of Table 2 do precisely that. Comparing the conventional estimates of ζ that are based on all firms in the economy (Table 1) with the estimates based on the non-exporting sample (columns 2, 4, and 6 of Table 2) thus allows us to get a sense of how far off are the conventional estimates of $\frac{\theta}{\varepsilon-1}$. In practice, for the CDF and PDF estimators, this does not turn out to be a large difference: the Table 1 coefficients for all firms are only about 3.6% off from the "true" estimates of $\frac{\theta}{\varepsilon-1}$, that are based on the non-exporting firms only. For the log-rank-log-size estimator, this difference is larger, 20%.

The second correction suggested by theory is that for all firms, the domestic sales will obey the power law with exponent $\frac{\theta}{\varepsilon-1}$. Thus, another conceptually correct way of estimating this combination of parameters is to fit a power law on domestic sales for all firms, non-exporting and exporting. Table 4 reports the results. Several things are worth noting. First, as predicted by theory, the power law estimates in Table 4 are higher in absolute value than in Table 1 for total sales. Once again, we see that ignoring exporting behavior leads to power law estimates that are too low. Second, the coefficients in Table 4 are quite similar to the non-exporter coefficients of Table 2, suggesting that the two different corrections we propose deliver mutually consistent results.

5 Conclusion

It has been known since at least the 1940's and the 1950's that the probability distributions for many economic variables can be described by power laws. Fifty years on, renewed interest in the causes and consequences of these phenomena coincides with the advancement and application of theoretical frameworks that model heterogeneous firms. This paper argues that theories of heterogeneous firms can fruitfully inform the empirics of estimating power laws in firm size. We set up a simple version of such a model, and show that international trade affects the distribution of firm size systematically: the exponent of the power law among exporting firms is lower than among the non-exporting firms. We then use a comprehensive dataset of French firms to demonstrate that this prediction holds very strongly in the data.

Two corrections can be implemented to obtain power law estimates consistent with theory. The main advantage of the methods proposed here is simplicity: all they require is an appropriate modification of either the estimation sample, or the variable to be used. In practice, it turns out that at the economy-wide level, the estimated corrected coefficients do not differ much from the traditional, unadjusted ones. One possible conclusion from this exercise is that if one is interested in a ballpark estimate of the extent of firm size heterogeneity, systematic biases introduced by exporting behavior are not that large. However, we would caution against generalizing this conclusion to other countries and settings. For instance, the impact of exporting behavior could be much greater in smaller countries, or in developing ones, or in individual industries. Thus, a more general approach to obtaining reliable estimates would be to implement the corrections suggested in this paper.

Appendix A Proof of Proposition 1

Proof:

From (5), the probability of exports, conditioning on domestic size is:

$$P\left(X_i > 0 \mid D_i\right) = H\left(D_i\right).$$

Call p_Y the density of a generic variable Y. We start from the postulate (e.g., coming from random growth) that the distribution of baseline sizes is:

$$p_D\left(x\right) = kx^{-\zeta-1}1_{x>\underline{D}}, \tag{A.1}$$

where k is an integration constant, $k = \zeta \underline{D}^{\zeta}$.

We next calculate the distribution of exports. To do that, we consider an arbitary "test function" g (continuous and non-zero over on a compact set), and calculate the expected value of a test function $g\left(X\right)$. First, given (5),

$$E\left[g\left(X_i\right) \mid D_i\right] = \left(1 - H\left(D_i\right)\right) g\left(0\right) + H\left(D_i\right) g\left(\phi D_i\right).$$

Therefore,

$$\begin{aligned} E\left[g\left(X_i\right)\right] &= E\left[E\left[g\left(X_i\right) \mid D_i\right]\right] \\ &= E\left[\left(1 - H\left(D_i\right)\right) g\left(0\right) + H\left(D_i\right) g\left(\phi D_i\right)\right] \\ &= \int_D \left(1 - H\left(D\right)\right) g\left(0\right) p_D\left(D\right) + H\left(D\right) g\left(\phi D\right) p_D\left(D\right) dD \\ E\left[g\left(X_i\right)\right] &= \left(\int_D \left(1 - H\left(D\right)\right) p_D\left(D\right)\right) g\left(0\right) + \int_{x>0} \left(H\left(\frac{x}{\phi}\right) p_D\left(\frac{x}{\phi}\right) \frac{1}{\phi}\right) g\left(x\right) dx. \end{aligned} \tag{A.2}$$

Equation (A.2) implies that the probability measure associated with x has a point mass $\int_D \left(1 - H\left(D\right)\right) p_D\left(D\right)$ on $X = 0$, and a density $H\left(\frac{x}{\phi}\right) p_D\left(\frac{x}{\phi}\right) \frac{1}{\phi}$ for $x > 0$. Hence, the density associated with the restriction of the exports to $X > 0$ is

$$p_X\left(x\right) = k' p_D\left(\frac{x}{\phi}\right) H\left(\frac{x}{\phi}\right) \frac{1}{\phi}$$

for a constant k' such that $\int_{x>0} p_X\left(x\right) dx = 1$. With the Pareto density of baseline sizes (A.1),

$$p_X\left(x\right) = K x^{-\zeta-1} H\left(\frac{x}{\phi}\right) 1_{x>\phi\underline{D}}$$

for a constant $K = k' \phi^{\zeta} k$.

We can calculate the distribution of S_i using a similar approach. From (6), a reasoning analogous to the one used for exports shows:

$$p_S\left(x\right) = p_D\left(x\right)\left(1 - H\left(x\right)\right) + p_D\left(\frac{x}{1+\phi}\right) H\left(\frac{x}{1+\phi}\right) \frac{1}{1+\phi}.$$

Thus, with the Pareto specification for D:

$$p_S(x) = kx^{-\zeta-1}\left[1 - H(x) + H\left(\frac{x}{1+\phi}\right)(1+\phi)^{\zeta}\right]1_{x>(1+\phi)\underline{D}} + kx^{-\zeta-1}1_{\underline{D}<x<(1+\phi)\underline{D}}. \quad \text{(A.3)}$$

We see a Pareto shape in the tails, but modulated by the export probability function H. ■

References

Arkolakis, Costas, "Market Penetration Costs and the New Consumers Margin in International Trade," 2008. Mimeo, Yale University.

———, "A Unified Theory of Firm Selection and Growth," May 2009. mimeo, Yale University.

Axtell, Robert L., "Zipf Distribution of U.S. Firm Sizes," *Science*, September 2001, *293* (5536), 1818–1820.

Buera, Francisco and Yongseok Shin, "Financial Frictions and the Persistence of History: A Quantitative Exploration," April 2008. mimeo, UCLA and Washington University in St. Louis.

Champernowne, D., "A model of income distribution," *Economic Journal*, 1953, *83*, 318–351.

Chaney, Thomas, "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, September 2008, *98* (4), 1707–1721.

di Giovanni, Julian and Andrei A. Levchenko, "International trade and Aggregate Fluctuations in Granular Economies," February 2009a. RSIE Discussion Paper 585.

——— and ———, "Firm Entry, Trade, and Welfare in Zipf's World," September 2009b. RSIE Discussion Paper 591.

Easterly, William, Ariell Reshef, and Julia Schwenkenberg, "The Power of Exports," October 2009. World Bank Policy Research Working Paper 5081.

Eaton, Jonathan and Samuel Kortum, "Technology in the Global Economy: A Framework for Quantitative Analysis," 2005. Book manuscript, New York University and University of Chicago.

———, ———, and Francis Kramarz, "An Anatomy of International Trade: Evidence From French Firms," December 2008. NBER Working Paper No. 14610.

Gabaix, Xavier, "Zipf's law for cities: an explanation," *Quarterly Journal of Economics*, 1999, *114*, 739–767.

———, "Power Laws in Economics and Finance," *Annual Reviews of Economics*, 2009a, *1*, 255–93.

———, "The Granular Origins of Aggregate Fluctuations," August 2009b. NBER Working Paper No. 15286.

——— and Augustin Landier, "Why has CEO pay increased so much?," *Quarterly Journal of Economics*, 2007, *123*, 49–100.

——— and Rustam Ibragimov, "Rank-1/2: A Simple Way to Improve the OLS Estimation of Tail Exponents," 2008. Forthcoming, *Journal of Business and Economic Statistics*.

Hanson, Gordon H. and Chong Xiang, "Testing the Melitz Model of Trade: An Application to U.S. Motion Picture Exports," October 2008. NBER Working Paper 14461.

Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple, "Export versus FDI with Heterogeneous Firms," *American Economic Review*, March 2004, *94* (1), 300–316.

Hinloopen, Jeroen and Charles van Marrewijk, "Comparative advantage, the rank-size rule, and Zipf's law," 2006. Tinbergen Institute Discussion Paper 06-100/1.

Luttmer, Erzo, "Selection, growth, and the size distribution of firms," *Quarterly Journal of Economics*, 2007, *122*, 1103–1144.

Melitz, Marc J., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, November 2003, *71* (6), 1695–1725.

Rossi-Hansberg, Esteban and Mark Wright, “Urban structure and growth,” *Review of Economic Studies*, 2007, *74*, 597–624.

Zipf, G., *Human Behavior and the Principle of Least Effort*, Cambridge, Mass.: Addison-Wesley, 1949.

Table 1. Power Law in Firm Size, All Firms

	(1) CDF	(2) PDF	(3) ln(Rank-0.5)
ζ	1.017	1.019	0.825
	(0.032)	(0.031)	(0.003)
R^2	0.990	0.998	0.991
No. of firms	157,084	157,084	157,084

Notes: This table reports the estimates of power laws in firm size (total sales), using the three methods described in the text. Column (1) estimates the CDF of the power law specified in equation (9). Column (2) estimates the PDF of the power law specified in equation (10). Column 3 regresses log(Rank-0.5) of the firm in the size distribution on log of its size (Gabaix and Ibragimov 2008).

Table 2. Power Laws in Firm Size, Non-Exporting and Exporting Firms

	CDF		**PDF**		**ln(Rank-0.5)**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Exporters	*Non-Exporters*	*Exporters*	*Non-Exporters*	*Exporters*	*Non-Exporters*	*t-stat*
ζ	0.964	1.055	0.967	1.095	0.738	1.029	46.24**
	(0.042)	(0.011)	(0.041)	(0.044)	(0.004)	(0.005)	
R^2	0.981	0.999	0.996	0.996	0.972	0.998	
No. of firms	67,078	90,006	67,078	90,006	67,078	90,006	

Notes: This table reports the estimates of power laws in firm size (total sales) for non-exporting and exporting firms separately, using the three methods described in the text. Columns (1)-(2) estimate the CDF of the power law specified in equation (9). Columns (3)-(4) estimate the PDF of the power law specified in equation (10). Columns (5)-(6) regress log(Rank-0.5) of the firm in the size distribution on log of its size (Gabaix and Ibragimov 2008). The last column reports the $t-$statistic for the test of the difference between the coefficients in columns (5) and (6). **: significant at the 1% level.

Table 3. Power Laws in Firm Size By Sector, Non-Exporting and Exporting Firms

	Exporting Firms				**Non-Exporting Firms**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sector	ζ	Std. Error	R^2	No. of firms	ζ	Std. Error	R^2	No. of firms	*t-stat*
Agriculture, Forestry, and Fishing	1.010	0.046	0.985	982	1.418	0.037	0.997	2,967	6.98**
Food Products	0.609	0.016	0.908	2,876	0.937	0.021	0.984	4,155	12.57**
Apparel and Leather Products	0.818	0.034	0.958	1,135	1.287	0.092	0.990	394	4.79**
Printing and Publishing	0.808	0.025	0.971	2,136	1.127	0.035	0.997	2,056	7.44**
Pharmaceuticals, Perfumes, and Beauty Products	0.512	0.029	0.903	605	0.604	0.071	0.975	145	1.19
Furniture, Household Goods	0.755	0.027	0.969	1,540	1.490	0.068	0.993	971	10.08**
Automotive	0.531	0.030	0.958	608	0.651	0.049	0.903	347	2.07**
Transport Equipment	0.554	0.040	0.975	393	0.877	0.084	0.991	218	3.48**
Non-electrical Machinery	0.785	0.017	0.967	4,166	1.338	0.028	0.984	4,556	16.81**
Electrical Machinery	0.710	0.027	0.979	1,394	1.437	0.059	0.991	1,179	11.18**
Mineral Products	0.656	0.031	0.948	919	0.979	0.031	0.948	2,062	7.49**
Textiles	0.844	0.038	0.919	1,008	1.194	0.091	0.992	346	3.56**
Wood and Paper Products	0.765	0.026	0.958	1,695	1.234	0.045	0.982	1,511	9.01**
Chemicals, Plastic, and Rubber	0.662	0.018	0.935	2,613	0.888	0.039	0.976	1,046	5.27**
Metals	0.793	0.017	0.976	4,574	1.241	0.031	0.993	3,244	12.79**
Electrical and Electronic Components	0.648	0.029	0.958	977	1.181	0.077	0.992	473	6.48**
Fuels	0.378	0.076	0.955	49	0.470	0.077	0.924	75	0.85
Water,Gas, Electricity	0.362	0.081	0.944	40	0.622	0.038	0.967	529	2.9**
Automotive Sales and Repair	0.737	0.016	0.947	4,516	1.029	0.012	0.981	14,648	14.84**
Wholesale Trade, Intermediaries	0.760	0.008	0.967	20,216	0.923	0.009	0.994	20,265	13.7**
Transport	0.856	0.017	0.970	5,339	1.014	0.016	0.995	8,293	6.91**
Professional Services	0.814	0.012	0.987	8,687	1.155	0.012	1.000	18,165	19.72**
Research and Development	0.751	0.072	0.983	219	0.832	0.093	0.976	159	0.69
Personal and Domestic Services	1.011	0.116	0.967	153	1.663	0.078	0.997	898	4.66**
Education	0.989	0.091	0.971	238	1.387	0.054	0.995	1,304	3.77**

Notes: This table reports the estimates of power laws in firm size (total sales) for non-exporting and exporting firms separately, for each individual sector, estimated using the log-rank-log-size estimator. The last column reports the $t-$statistic for the test of the difference between the coefficients in columns (1) and (5). **: significant at the 1% level; *: significant at the 5% level.

Table 4. Power Law in Firm Size, All Firms, Domestic Sales Only

	(1) **CDF**	(2) **PDF**	(3) **ln(Rank-0.5)**
ζ	1.048	1.055	0.869
	(0.030)	(0.027)	(0.003)
R^2	0.992	0.998	0.992
No. of firms	157,084	157,084	157,084

Notes: This table reports the estimates of power laws in firm size (domestic sales), using the three methods described in the text. Column (1) estimates the CDF of the power law specified in equation (9). Column (2) estimates the PDF of the power law specified in equation (10). Column 3 regresses log(Rank-0.5) of the firm in the size distribution on log of its size (Gabaix and Ibragimov 2008).

Figure 1. The Analytical Power Law in the Melitz-Pareto Model: Multiple Export Markets

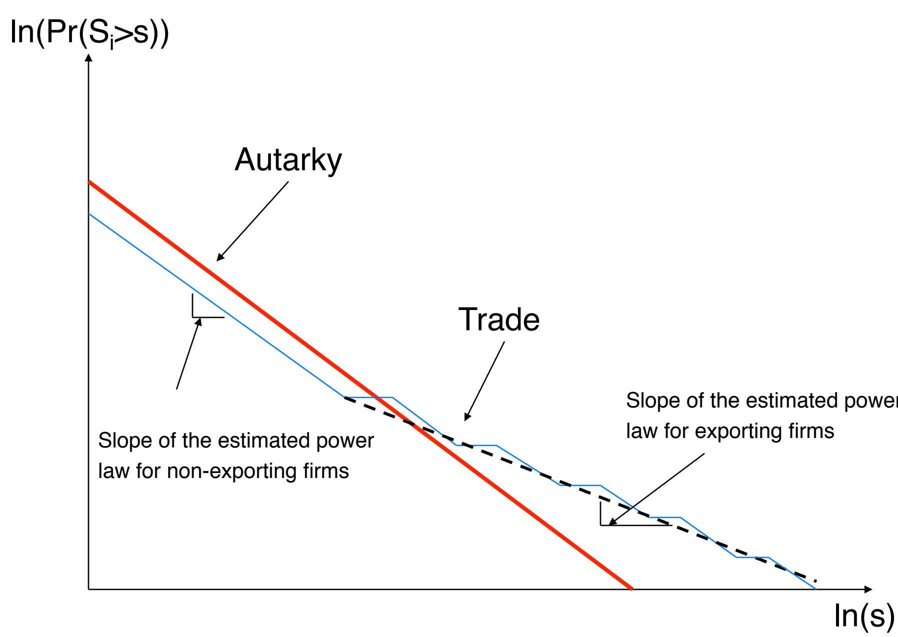


Figure 2. Partition of Firms

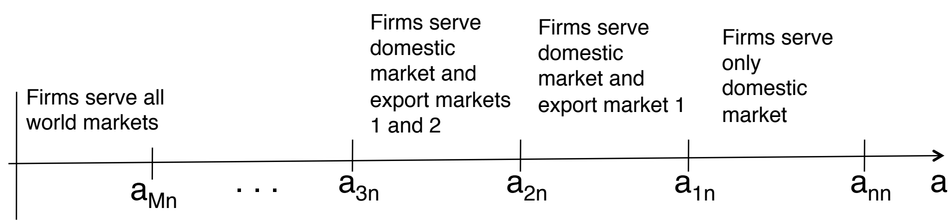


Figure 3. The Analytical Power Law in the Melitz-Pareto Model: Stochastic Export Entry Costs

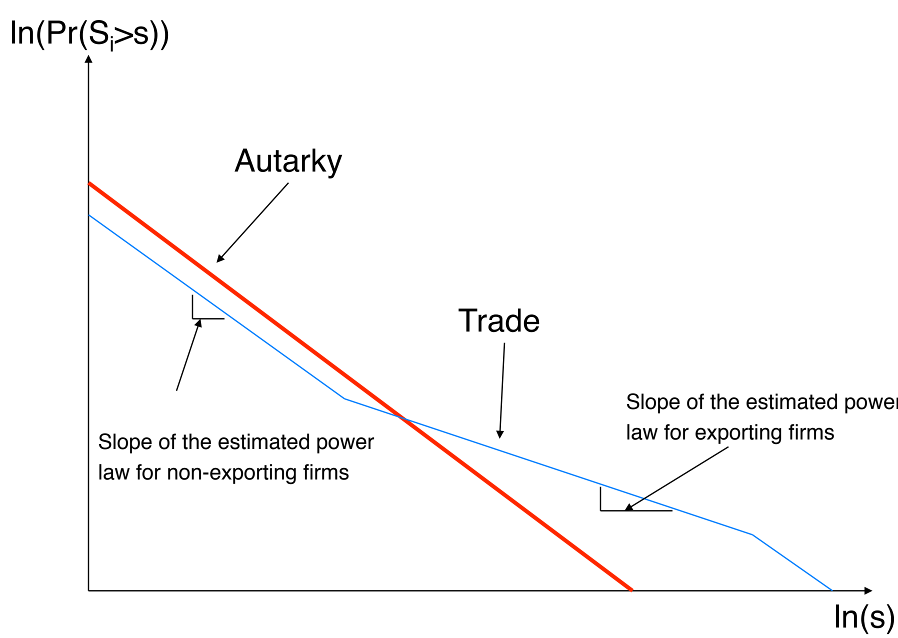
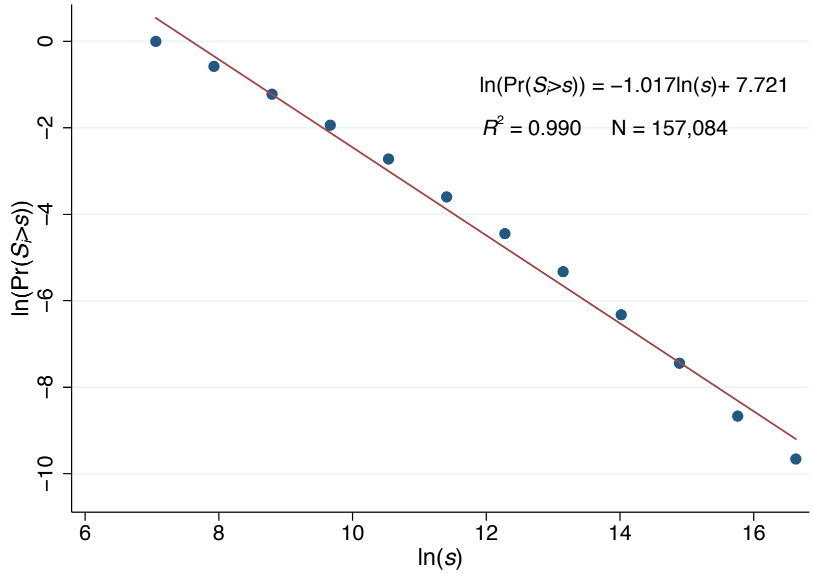
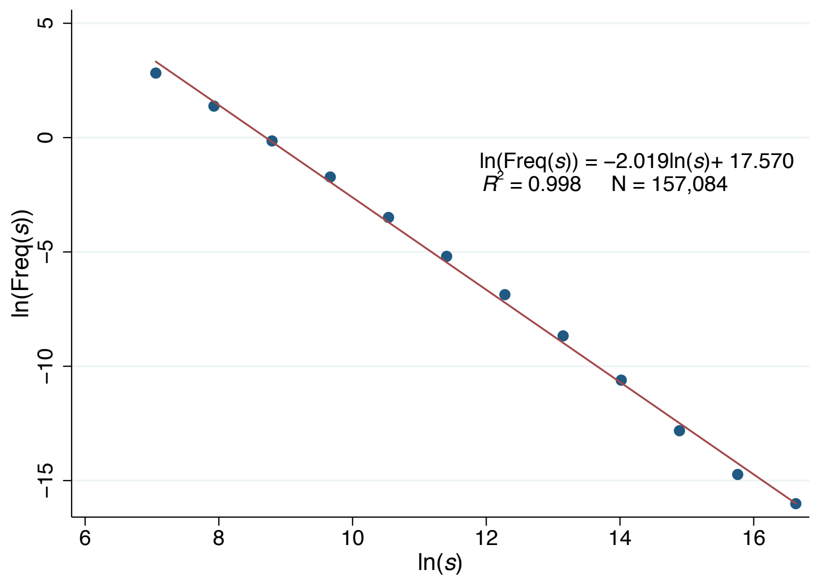


Figure 4. Power Laws in the Distribution of Firm Size, All Firms



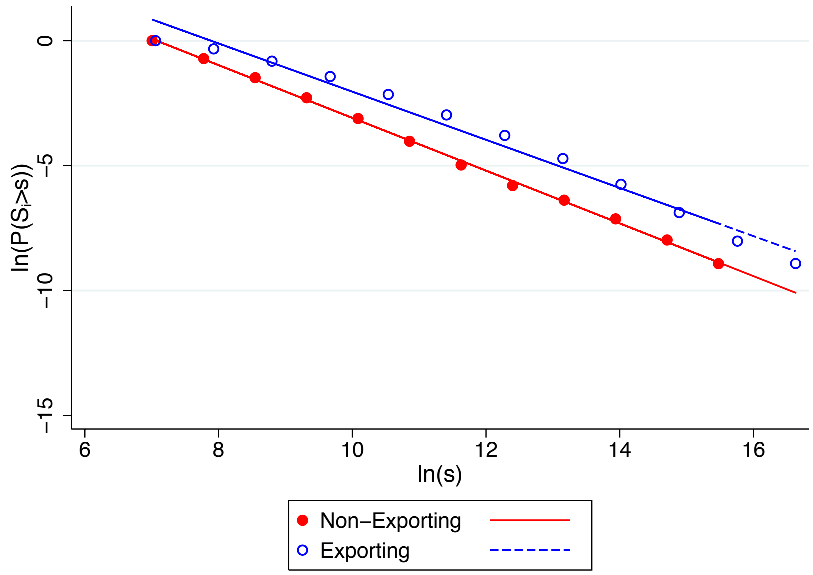
(a)



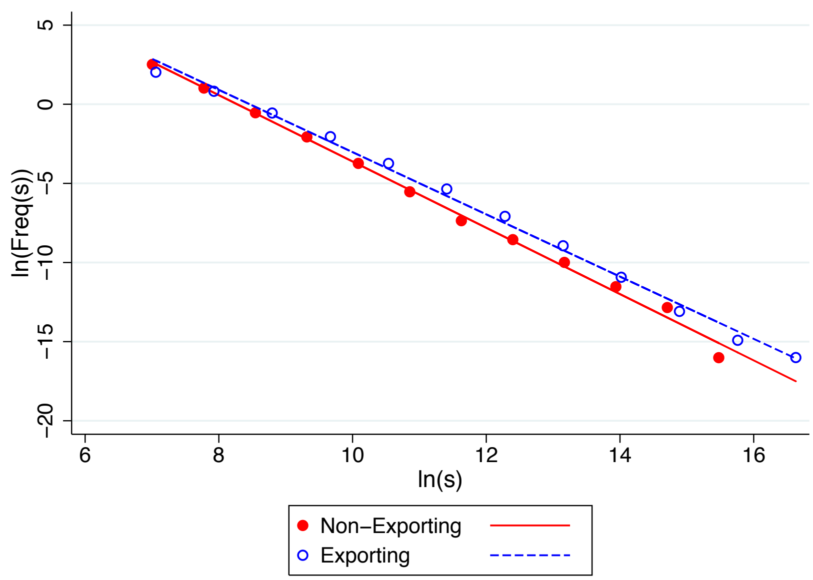
(b)

Notes: This figure reports the estimated power laws in firm size based on total sales and all firms. The power laws are estimated with two different methods, the cdf (panel a) and the pdf (panel b).

Figure 5. Power Laws in the Distribution of Firm Size, Exporting and Non-Exporting Firms



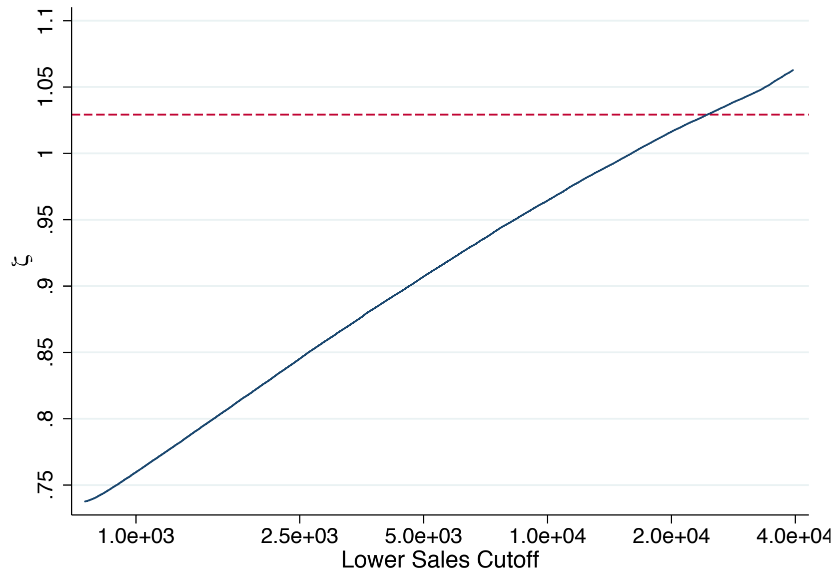
(a)



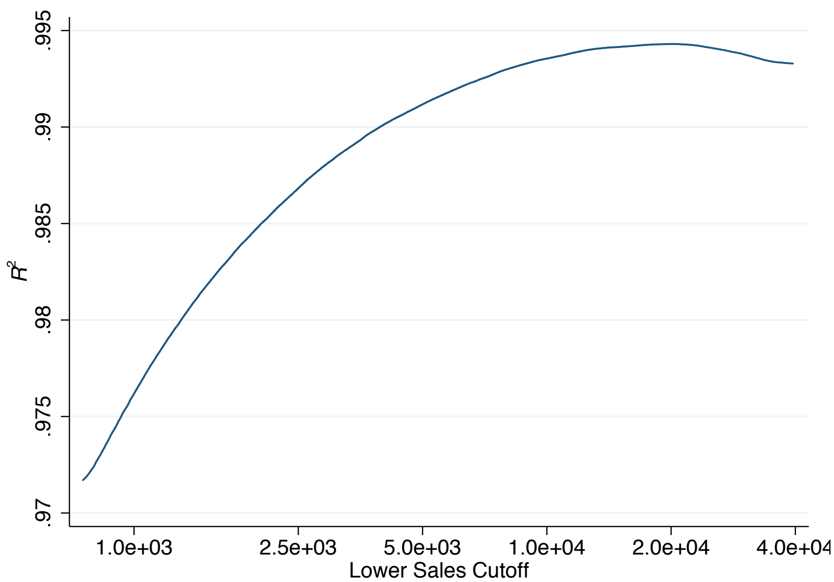
(b)

Notes: This figure reports the estimated power laws in firm size for exporting and non-exporting firms separately. The power laws are estimated with two different methods, the CDF (panel a) and the pdf (panel b).

Figure 6. Power Law Coefficient for Exporting Firms



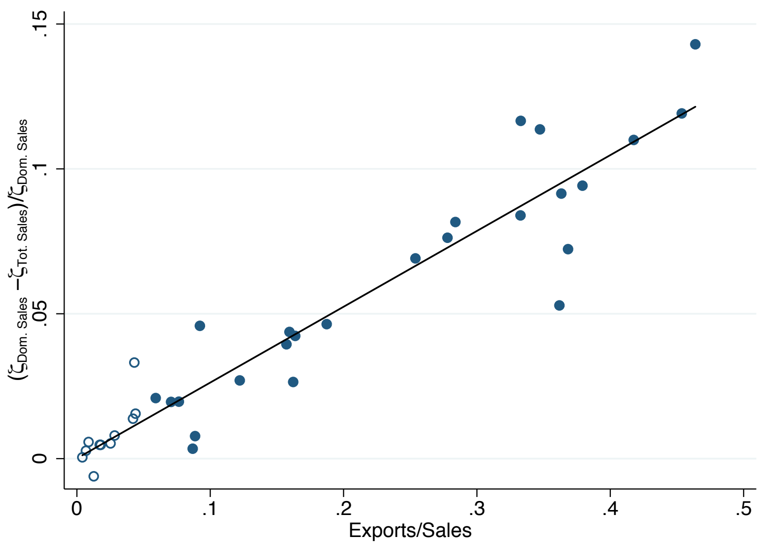
(a)



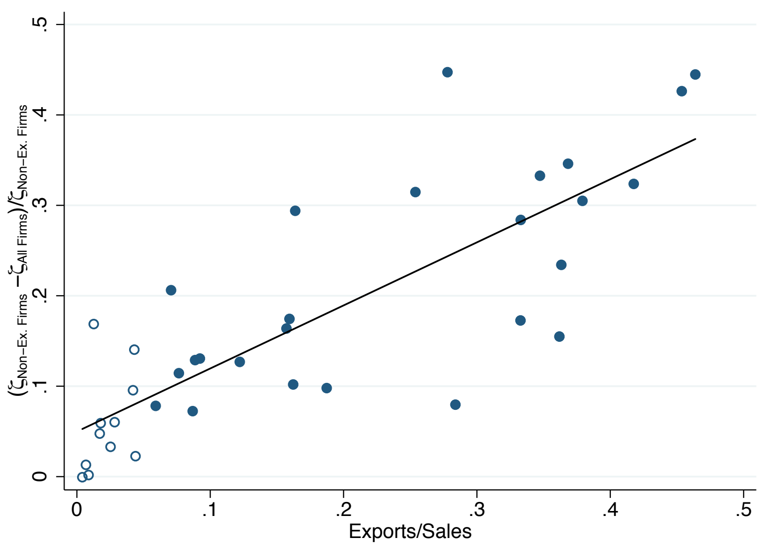
(b)

Notes: The top panel of this Figure depicts the power law coefficient estimated on exporting firms on the vertical axis, plotted against the minimum sales cutoff on the horizontal axis. The dashed horizontal line in the top panel is the power law coefficient for domestic firms. The bottom panel depicts the concomitant evolution in the R^2's of the estimates.

Figure 7. Deviations in Power Law Estimates and Openness at Sector Level



(a)



(b)

Notes: The Figure plots the differences between the power law coefficients at sector level against trade openness. In both panels, sector-level exports relative to total sales are on the horizontal axis. In the top panel, on the vertical axis is the percentage difference between the power law exponent estimated on domestic sales only and the power law exponent estimated on total sales: $\frac{\zeta_{dom} - \zeta_{total}}{\zeta_{dom}}$. In the bottom panel, on the vertical axis is the percentage difference between the power law exponent estimated on sales of non-exporting firms only and the power law exponent estimated on total sales of all firms: $\frac{\zeta_{nonex} - \zeta_{total}}{\zeta_{nonex}}$. The non-tradeable sectors are denoted by hollow dots, and the tradeable sectors by solid dots. The straight line is the OLS fit through the data.

Table A1. Summary Statistics

	(1)	(2)	(3)	(4)	(5)
	All Firms				
	No. of firms	Mean	St. Dev.	Min	Max
Total Sales	157,084	14,024	254,450	751	6.16E+07
Export Sales	157,084	2,894	89,232	0	1.95E+07
	Non-Exporting Firms				
	No. of firms	Mean	St. Dev.	Min	Max
Total Sales	90,006	6,434	86,009	751	1.67E+07
	Exporting Firms				
	No. of firms	Mean	St. Dev.	Min	Max
Total Sales	67,078	24,208	376,185	752	6.16E+07
Export Sales	67,078	6,777	136,456	1	1.95E+07

Notes: This table presents the summary statistics for the variables used in the estimation. Sales figures are in thousands of Euros.

Table A2. Power Laws in Firm Size By Sector, All Sales and Domestic Sales

	All Sales			Domestic Sales				
	(1) ζ	(2) Std. Error	(3) R^2	(4) ζ	(5) Std. Error	(6) R^2	(7) No. of firms	(8) Exports/Sales
Tradeable Sectors								
Agriculture, Forestry, and Fishing	-1.233	0.028	0.997	-1.292	0.029	0.998	3,949	0.092
Food Products	-0.661	0.011	0.962	-0.691	0.012	0.964	7,031	0.164
Apparel and Leather Products	-0.842	0.030	0.970	-0.908	0.033	0.968	1,529	0.368
Printing and Publishing	-0.895	0.020	0.988	-0.913	0.020	0.989	4,192	0.071
Pharmaceuticals, Perfumes, and Beauty Products	-0.510	0.026	0.912	-0.539	0.028	0.927	750	0.362
Furniture, Household Goods	-0.824	0.023	0.989	-0.892	0.025	0.992	2,511	0.278
Automotive	-0.538	0.025	0.987	-0.588	0.027	0.993	955	0.333
Transport Equipment	-0.593	0.034	0.987	-0.667	0.038	0.994	611	0.418
Non-electrical Machinery	-0.893	0.014	0.989	-1.007	0.015	0.986	8,722	0.347
Electrical Machinery	-0.798	0.022	0.995	-0.931	0.026	0.996	2,573	0.464
Mineral Products	-0.809	0.021	0.972	-0.845	0.022	0.971	2,981	0.159
Textiles	-0.855	0.033	0.938	-0.968	0.037	0.949	1,354	0.333
Wood and Paper Products	-0.845	0.021	0.981	-0.908	0.023	0.979	3,206	0.254
Chemicals, Plastic, and Rubber	-0.680	0.016	0.949	-0.748	0.017	0.955	3,659	0.363
Metals	-0.863	0.014	0.989	-0.952	0.015	0.988	7,818	0.379
Electrical and Electronic Components	-0.678	0.025	0.976	-0.770	0.029	0.978	1,450	0.454
Fuels	-0.422	0.054	0.957	-0.434	0.055	0.957	124	0.162
Water,Gas, Electricity	-0.577	0.034	0.979	-0.579	0.034	0.978	569	0.087
Automotive Sales and Repair	-1.146	0.008	0.996	-1.152	0.008	0.996	38,100	0.089
Wholesale Trade, Intermediaries	-0.896	0.009	0.980	-0.903	0.009	0.979	19,164	0.122
Transport	-0.806	0.006	0.982	-0.828	0.006	0.985	40,481	0.187
Professional Services	-1.107	0.007	0.996	-1.112	0.007	0.996	52,780	0.157
Research and Development	-0.915	0.011	0.987	-0.959	0.012	0.988	13,632	0.284
Personal and Domestic Services	-0.545	0.011	0.984	-0.548	0.011	0.984	4,632	0.077
Education	-0.959	0.011	0.986	-0.959	0.011	0.986	15,329	0.059
Non-Tradeable Sectors								
Construction	-0.548	0.037	0.994	-0.567	0.038	0.995	439	0.017
Retail Trade	-0.966	0.008	0.998	-1.006	0.009	0.999	26,852	0.018
Real Estate	-0.892	0.012	0.990	-0.899	0.012	0.991	11,183	0.004
Financial Services	-0.766	0.056	0.983	-0.834	0.061	0.988	378	0.009
Post and Telecommunications	-1.367	0.018	0.991	-1.374	0.018	0.991	12,152	0.043
Business Services	-0.928	0.021	0.994	-0.941	0.021	0.994	3,888	0.028
Hotels and Restaurants	-1.473	0.064	0.989	-1.503	0.066	0.993	1,051	0.025
Recreational, Cultural, and Athletic Activities	-1.279	0.046	0.996	-1.306	0.047	0.996	1,542	0.042
Health and Social Action	-1.111	0.021	0.978	-1.114	0.021	0.978	5,853	0.007
Public Administration	-0.901	0.157	0.990	-0.896	0.156	0.989	66	0.013
Unions and Extra-Territorial Activities	-1.012	0.055	0.993	-1.027	0.055	0.993	687	0.044

Notes: This table reports the estimates of power laws in firm size, for total sales (Columns 1-3), and for domestic sales (Columns 4-6), for each individual sector, estimated using the log-rank-log-size estimator. The last column reports the exports/total sales ratio in each sector.