Journal of International Economics 85 (2011) 42–52

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

Power laws in firm size and openness to trade: Measurement and implications ☆

Julian di Giovanni [a], Andrei A. Levchenko [b,c,*], Romain Rancière [d,e,f]

[a] Research Department, International Monetary Fund, 700 19th Street NW, Washington, DC, 20431, USA
[b] Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109, USA
[c] National Bureau of Economic Research, Cambridge, MA, USA
[d] Paris School of Economics, 48 Boulevard Jourdan, F-75014 Paris, France
[e] Research Department, International Monetary Fund, 700 19th Street NW, Washington, DC, 20431, USA
[f] Center for Economic Policy and Research, London, UK

ARTICLE INFO

Article history:
Received 9 April 2010
Received in revised form 6 May 2011
Accepted 9 May 2011
Available online 19 May 2011

JEL classification:
F12
F15

Keywords:
Firm size distribution
International trade
Power laws

ABSTRACT

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, both for the economy as a whole and at the industry level. Furthermore, the differences between power law coefficients for exporters and non-exporters are larger in sectors that are more open to trade. 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.

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 et al., 2009).

The literature has emphasized the importance of the precise value of the power law exponent, ζ. For instance, for the distribution of firm size in the U.S., 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. 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, 2011; di Giovanni and Levchenko, 2010a), regulation of entry (di Giovanni and Levchenko, 2010b), 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

☆ We are grateful to Xavier Gabaix, the editor (Jonathan Eaton), and an anonymous referee 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.

* Corresponding author at: Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109, USA. Tel.: +1 734 764 3296; fax: +1 734 764 2769.

E-mail addresses: JdiGiovanni@imf.org (J. di Giovanni), alev@umich.edu (A.A. Levchenko), ranciere@pse.ens.fr (R. Rancière).

URLs: http://julian.digiovanni.ca (J. di Giovanni), http://alevchenko.com (A.A. Levchenko), http://romainranciere.com (R. Rancière).

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

0022-1996/$ – see front matter
doi:10.1016/j.jinteco.2011.05.003

into account international trade could be misleading regarding the deep parameters of the economy.[2]

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 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 is highly statistically significant. In addition, we show that, as predicted by theory, 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.

Results based on the economy as a whole are not conclusive evidence that exporting behavior *per se* generates the difference in power law coefficients between exporters and non-exporters. For instance, it is well known that exporters tend to be larger than non-exporters, and thus the findings could be driven by a departure from an exact power law for larger firms. We address this concern in two ways. First, to focus specifically on the role of international trade, we estimate power law coefficients for exporting and non-exporting firms by sector. At the disaggregated industry level, the estimates exhibit the same pattern: power law coefficients for exporters are systematically lower than for non-exporters. More tellingly, the differences between power law coefficients are larger in industries that are more open to trade, a striking regularity consistent with the theoretical intuition developed in the paper. Second, we re-estimate the power laws on sub-samples of exporters and non-exporters of the same mean size, confirming the main result. 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 are used to pin down crucial parameters in calibrated heterogeneous firm models (see, among many others, Helpman et al., 2004; Chaney, 2008; Buera and Shin, 2010; di Giovanni and Levchenko, 2010a). 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 (2010b) 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 firm 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 et al. (forthcoming) 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.[3] 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. 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

This section describes the firm size distribution in the canonical heterogeneous firms model, and in particular how it is affected by international trade. Throughout this section, sales are the measure of firm size, though all the empirical results are reported below for both sales and employment as an alternative size variable. Firm sales, S_i, in the economy are said to follow a power law if their distribution is described by[4]:

$$Pr(S_i > s) = Cs^{-\zeta}. \tag{1}$$

The canonical monopolistic competition model with CES demand and heterogeneous firms implies that domestic sales of firm i in market n are given by

$$D_i = M_n \times B_i,$$

where M_n 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. In this expression, a_i is the marginal cost of firm i, and ε is the elasticity of substitution between CES varieties.[5]

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 the canonical heterogeneous firms model this is equivalent to assuming that firm productivity is Pareto, but with a different exponent. To see this, suppose that firm productivity has the Pareto cdf: $\Pr(1/a < y) = 1 - \left(\frac{b}{y}\right)^{\theta}$. In the autarkic economy, where $S_i = D_i$, the power law follows:

$$Pr(S_i > s) = Pr(M_n B_i > s) = Pr\left(a_i^{1-\varepsilon} > \frac{s}{M_n}\right) = \left(b^{\varepsilon-1} M_n\right)^{\frac{\theta}{\varepsilon-1}} s^{-\frac{\theta}{\varepsilon-1}}, \tag{2}$$

satisfying Eq. (1) 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 Fig. 1. In addition, this calculation shows that $S_i \sim \text{Pareto}\left(b^{\varepsilon-1} M_n, \frac{\theta}{\varepsilon-1}\right)$.

The result that the power law exponent is 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 international trade.

To start exporting from country n to country m, firm i must pay the fixed cost κ_{mni} that potentially varies by firm, and an iceberg per-unit cost of $\tau_{mn} > 1$. It is easy to verify that export sales by firm i to market m can be expressed as $M_m^* B_i$, where M_m^* is m's market size *from the*

[2] 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.

[3] See Arkolakis (2009, 2010) for related theoretical treatments.

[4] Unless otherwise noted, in the discussion below all parameter values are nonnegative.

[5] $M_n = \frac{Y_n}{P_n^{1-\varepsilon}}\left(\frac{\varepsilon}{\varepsilon-1}\omega_n\right)^{1-\varepsilon}$, where Y_n is total expenditure, P_n is the price level, and ω_n is the cost of the input bundle. For details of this type of model, see, e.g., di Giovanni and Levchenko (2010b).

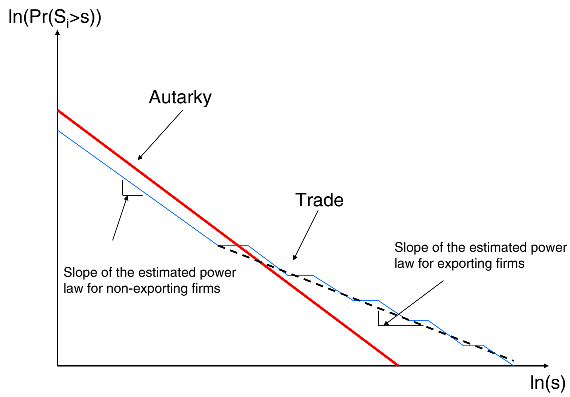

Fig. 1. The analytical power law in the Melitz–Pareto model: multiple export markets.

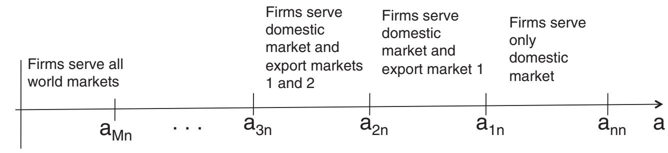

Fig. 2. Partition of firms.

perspective of the firm that exports there from n, and B_i is defined above. As is well known, variable profits from exporting to m are equal to $M_m^* B_i / \varepsilon$, and firm i exports to that market only if

$$\frac{M_m^* B_i}{\varepsilon} \geq \kappa_{mni}. \tag{3}$$

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. Eq. (3) defines a partition of firms according to how many markets m the firm serves. A greater range of firms will serve markets with higher M_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 Fig. 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 Fig. 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, as in Eaton et al. (forthcoming). In this case, there will not be a single "exporting cutoff," above which all firms export, and below which none do. Instead, there will be both exporting and non-exporting firms with the same exact productivity, or equivalently, domestic sales. 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]

[6] Hanson and Xiang (2011) use U.S. motion picture exports to provide empirical evidence that fixed exporting costs are global rather market-specific.

Defining $\phi \equiv \frac{M_m^*}{M_n}$ to be the ratio of the foreign market size relative to the domestic one, the exporting decision condition (3) 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 $X_i = M_n^* B_i = \phi D_i$, and its total sales, $S_i = D_i + X_i = (1 + \phi) D_i$. The distributions of domestic sales, export sales, and total sales are described in the following proposition. The proof is presented in Appendix A.

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

$$p_D(x) = kx^{-\zeta - 1} 1_{x > \underline{D}} \tag{4}$$

$$p_X(x) = Kx^{-\zeta - 1} H\left(\frac{x}{\phi}\right) 1_{x > \phi \underline{D}} \tag{5}$$

$$p_S(x) = kx^{-\zeta - 1}\left[1 - H(x) + H\left(\frac{x}{1 + \phi}\right)\ (1 + \phi)^{\zeta}\right]\ 1_{x > (1 + \phi)\underline{D}} + kx^{-\zeta - 1} 1_{\underline{D} < x < (1 + \phi)\underline{D}}, \tag{6}$$

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 Eq. (6) that describes the distribution of total sales, fitting instead the simple power law relationship (1) 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 $k'' > 0$ and $\alpha > 0$, and $H(x/\phi) = k''(x^*)^{\alpha}$ for $x > x^*$. Then, Eq. (5) implies that the distribution of export sales is given by:

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

and thus the power law exponent of X is $\zeta - \alpha$ for $x < x^*$, and ζ for $x \geq x^*$. When H has a high slope, the Pareto exponent of X is lower than that of domestic sales: there are fewer small exporters, 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

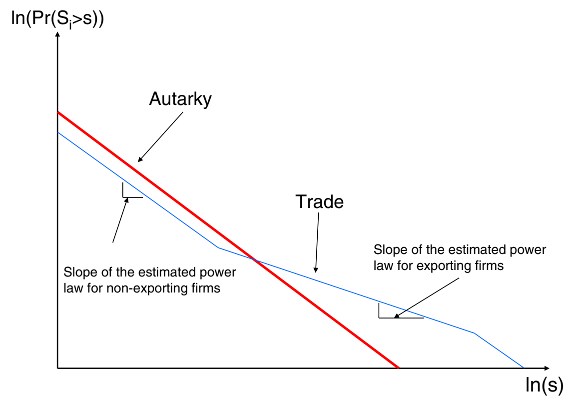

Fig. 3. The analytical power law in the Melitz–Pareto model: stochastic export entry costs.

of exports converges to the exponent of domestic sales. Such a possibility is depicted in Fig. 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 Figs. 1 and 3, is to use only non-exporting firms in the estimation of ζ. The second, suggested by Eq. (2) 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, 2009). The minimum sales cutoff we select is Euro 750,000 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 Euro 100,000.

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.

Appendix 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 million vs. Euro 6.4 million). We also consider the number of employees as an alternative measure of firm size. Appendix Table A1 reports the summary statistics for the number of employees in each panel (all firms, non-exporting, and exporting firms). When measured by employees, the average exporter is approximately three times larger than the average non-exporter (69.8 vs. 25.4 workers).

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 (1), which in natural logs becomes:

$$ln(Pr(S_i > s)) = ln(C) - \zeta ln(s). \tag{7}$$

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 a 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 Eq. (1) has a probability density function

$$f(s) = C\zeta s^{-(\zeta+1)}. \tag{8}$$

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

[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.

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:

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

where i indexes firms, and S_i are sales. This is the estimator suggested by Gabaix and Ibragimov (2011). 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. Panel I uses total sales as the measure of firm size. We can see that the fit of the data for sales 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 (2001) finds for the U.S. Fig. 4 presents the results graphically for sales. 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. Panel II reports estimates on the number of employees instead. The fit is again quite good and the point estimates are similar to those for sales.

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. Once again, Panel I uses sales as the measure of firm size, while Panel II uses employees. 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

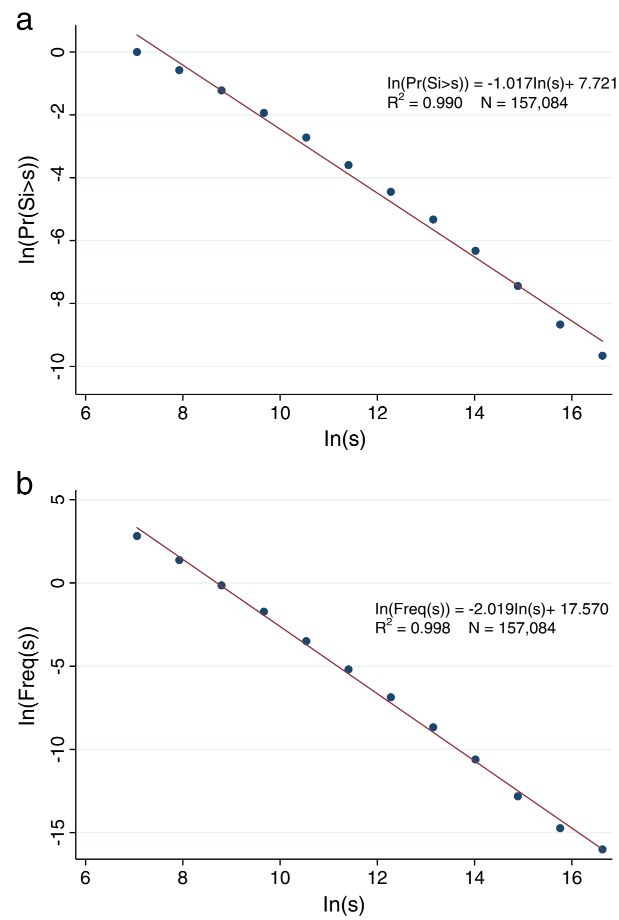

Fig. 4. Power laws in the distribution of firm size, all firms. Notes: This figure reports the estimated power laws in firm size based on total sales and all firms. The power laws are estimated with two different methods, the cdf (panel a) and the pdf (panel b).

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.

Fig. 5 presents the sales 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 large firms that export than those that do not. The pdf plot conveys the same message.

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

Table 1
Power law in firm size, all firms: sales and employees.

	(1) CDF	(2) PDF	(3) ln(Rank −0.5)
I. Sales			
ζ	1.017	1.019	0.825
	(0.032)	(0.031)	(0.004)
R^2	0.990	0.998	0.991
No. of firms	157,084	157,084	157,084
II. Employees			
ζ	1.078	1.093	0.790
	(0.072)	(0.083)	(0.003)
R^2	0.958	0.985	0.906
No. of firms	152,429	152,429	152,429

Notes: This table reports the estimates of power laws in firm size (total sales or employees), using the three methods described in the text. Column (1) estimates the CDF of the power law specified in Eq. (7). Column (2) estimates the PDF of the power law specified in Eq. (8). Column 3 regresses ln(Rank −0.5) of the firm in the size distribution on the natural logarithm of its size (Gabaix and Ibragimov, 2011).

[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 coefficients in Columns (1) compared to (2) and (3) compared to (4).

Table 2
Power laws in firm size, non-exporting and exporting firms: sales and employees.

	CDF		PDF		ln(Rank −0.5)		
	(1) Exporters	(2) Non-exporters	(3) Exporters	(4) Non-exporters	(5) Exporters	(6) Non-exporters	(7) t-stat
I. Sales							
ζ	0.964	1.055	0.967	1.095	0.738	1.029	37.03**
	(0.042)	(0.011)	(0.041)	(0.044)	(0.006)	(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	
II. Employees							
ζ	0.967	1.251	0.949	1.111	0.724	0.891	28.54**
	(0.078)	(0.074)	(0.100)	(0.041)	(0.004)	(0.004)	
R^2	0.939	0.967	0.974	0.991	0.906	0.870	
No. of firms	66,040	86,389	66,040	86,389	66,040	86,389	

Notes: This table reports the estimates of power laws in firm size (total sales or employees) 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 Eq. (7). Columns (3)–(4) estimate the PDF of the power law specified in Eq. (8). Columns (5)–(6) regress ln (Rank −0.5) of the firm in the size distribution on the natural logarithm of its size (Gabaix and Ibragimov, 2011). 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.

is still quite good. The range of the R^2's is 0.972 to 0.999 for sales, and 0.870 to 0.991 for employees.

It could be that the results are driven by the fact that exporters are on average larger than non-exporters. To check robustness of the results, we therefore compare the exporters' group to a non-exporters' group while dropping smallest non-exporting firms so that the mean sales of the two groups is equal. This way, the power laws for exporters and non-exporters are estimated on the samples with identical mean firm size. The results, reported in Appendix Table A3, are robust to matching up the mean size of non-exporters and exporters.

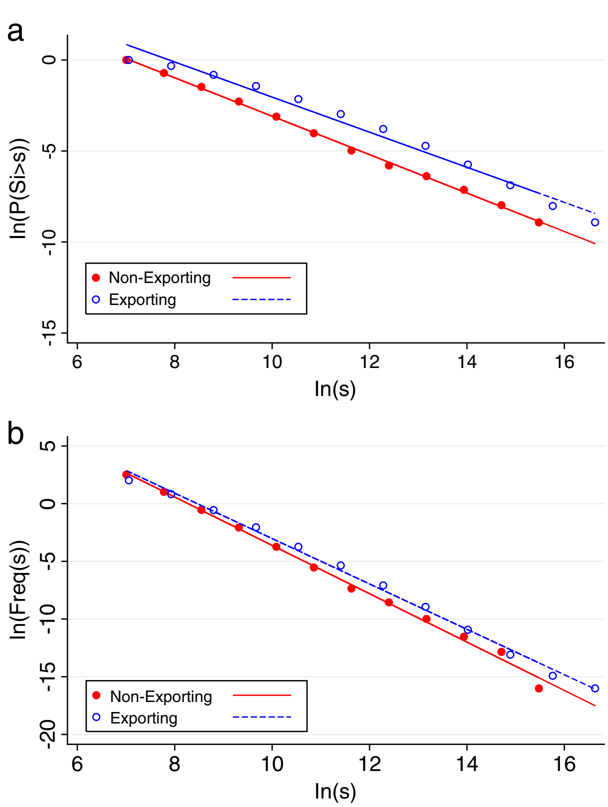

Fig. 5. Power laws in the distribution of firm size, exporting and non-exporting firms. Notes: This figure reports the estimated power laws in firm size for exporting and non-exporting firms separately. The power laws are estimated with two different methods, the CDF (panel a) and the pdf (panel b).

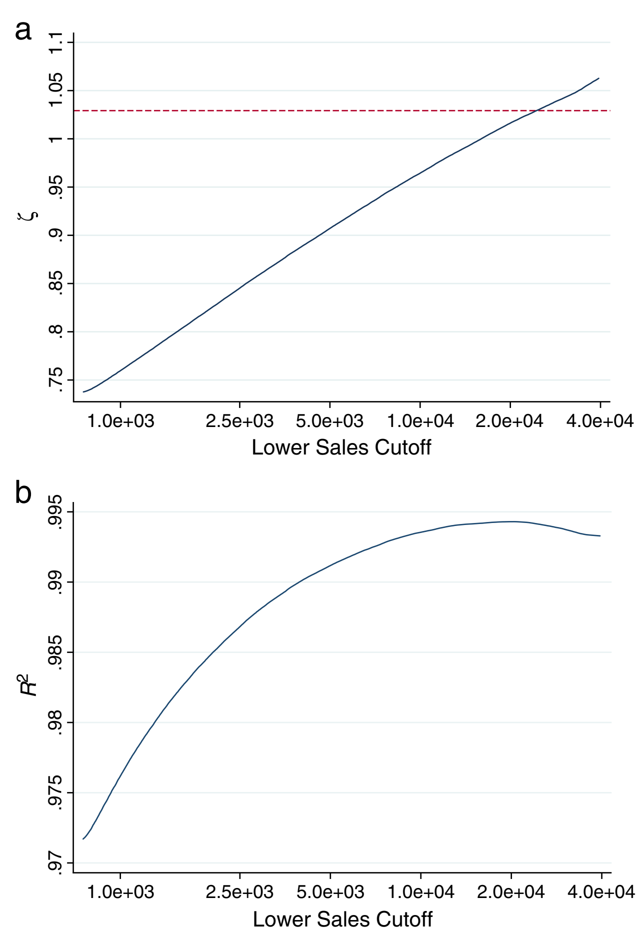

Fig. 6. Power law coefficient for exporting firms. Notes: The top panel of this Figure depicts the power law coefficient estimated on exporting firms on the vertical axis, plotted against the minimum sales cutoff on the horizontal axis. The dashed horizontal line in the top panel is the power law coefficient for domestic firms. The bottom panel depicts the concomitant evolution in the R^2's of the estimates.

Table 3
Power laws in firm size by sector, non-exporting and exporting firms: sales.

	Exporting firms				Non-exporting firms				
Sector	(1) ζ	(2) Std. error	(3) R^2	(4) No. of firms	(5) ζ	(6) Std. error	(7) R^2	(8) No. of firms	(9) t-stat
Agriculture, Forestry, and Fishing	1.010	0.046	0.985	982	1.418	0.037	0.997	2967	6.98**
Food Products	0.609	0.016	0.908	2876	0.937	0.021	0.984	4155	12.57**
Apparel and Leather Products	0.818	0.034	0.958	1135	1.287	0.092	0.990	394	4.79**
Printing and Publishing	0.808	0.025	0.971	2136	1.127	0.035	0.997	2056	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	1540	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	4166	1.338	0.028	0.984	4556	16.81**
Electrical Machinery	0.710	0.027	0.979	1394	1.437	0.059	0.991	1179	11.18**
Mineral Products	0.656	0.031	0.948	919	0.979	0.031	0.948	2062	7.49**
Textiles	0.844	0.038	0.919	1008	1.194	0.091	0.992	346	3.56**
Wood and Paper Products	0.765	0.026	0.958	1695	1.234	0.045	0.982	1511	9.01**
Chemicals, Plastic, and Rubber	0.662	0.018	0.935	2613	0.888	0.039	0.976	1046	5.27**
Metals	0.793	0.017	0.976	4574	1.241	0.031	0.993	3244	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	2.90**
Automotive Sales and Repair	0.737	0.016	0.947	4516	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	13.70**
Transport	0.856	0.017	0.970	5339	1.014	0.016	0.995	8293	6.91**
Professional Services	0.814	0.012	0.987	8687	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	1304	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.

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 Fig. 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. Fig. 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, at sector level we should expect to see the same patterns as found above for the aggregate. Examining the predictions of the heterogeneous firm model at the industry level is important for at least two reasons. First, estimates for the economy as a whole are not conclusive evidence that exporting behavior per se generates the differences in power law coefficients. For instance, it is well known that exporters tend to be larger than non-exporters, and thus the findings could be driven by a departure from an exact power law for larger firms. Some of this concern can be assuaged by constraining the sample of exporters and non-exporters to have the same mean size, as we do above. A more convincing approach is to exploit differences in trade openness by sector to provide further evidence for the role of exporting in generating these results. In addition, looking at sector-level data can also reassure us that the results are not driven by compositional effects.

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.[10] 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

[10] Appendix Table A4 presents the sector-level estimates for employees rather than total sales for size. The results are the same.

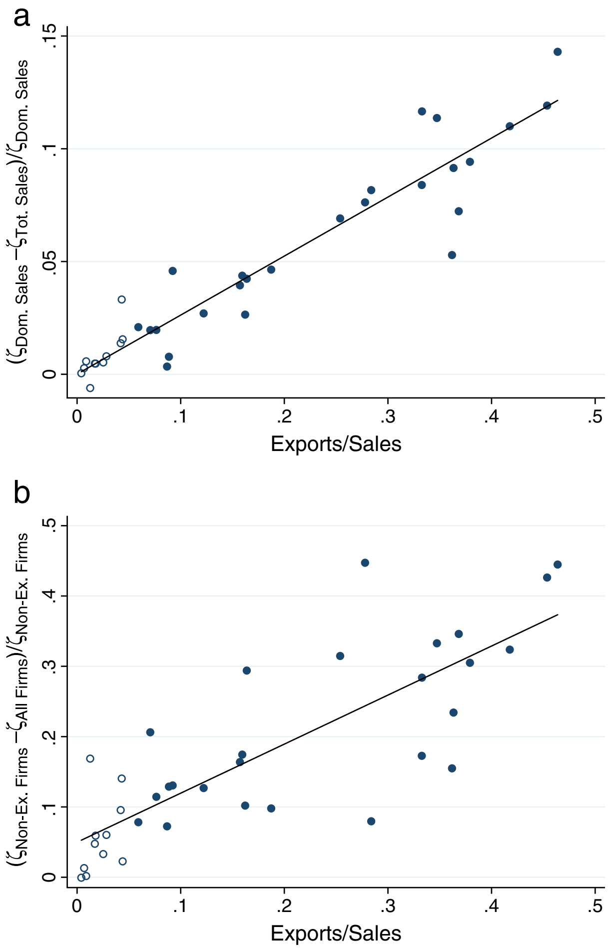

Fig. 7. Deviations in power law estimates and openness at sector level. Notes: The Figure plots the differences between the power law coefficients at sector level against trade openness. In both panels, sector-level exports relative to total sales are on the horizontal axis. In the top panel, on the 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.

sectors that are more open to trade. Fig. 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, Fig. 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 non-exporting 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, 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: 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

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)	(0.004)
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 Eq. (7). Column (2) estimates the PDF of the power law specified in Eq. (8). Column 3 regresses ln(Rank −0.5) of the firm in the size distribution on the natural logarithm of its size (Gabaix and Ibragimov, 2011).

results. Of course, this correction cannot be implemented using employment instead of sales, since employment data cannot be broken down into workers producing for domestic market and those producing for exports.

5. Conclusion

It has been known since at least the 1940s and the 1950s 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. The firm's export sales are given by:

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

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

$$S_i = \begin{cases} D_i & \text{if } \frac{\phi D_i}{\varepsilon} < \kappa_i; \text{ probability } 1-H(D_i) \\ (1+\phi)\ D_i & \text{if } \frac{\phi D_i}{\varepsilon} \geq \kappa_i; \text{ probability } H(D_i). \end{cases} \qquad \text{A.2}$$

From Eq. (A.1), the probability of exports, conditioning on domestic size is:

$$P(X_i > 0 | D_i) = H(D_i).$$

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:

$$p_D(x) = kx^{-\zeta-1} 1_{x > \underline{D}}, \qquad (A.3)$$

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

We next calculate the distribution of exports. To do that, we consider an arbitrary "test function" g (continuous and non-zero over on a compact set), and calculate the expected value of a test function $g(X)$. First, given (A.1),

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

Therefore,

$$\begin{aligned} E[g(X_i)] &= E[E[g(X_i)|D_i]] \\ &= E[(1-H(D_i))\ g(0) + H(D_i)\ g(\phi D_i)] \\ &= \int_D (1-H(D))\ g(0) p_D(D) + H(D)\ g(\phi D)\ p_D(D)\ dD \\ E[g(X_i)] &= \left(\int_D (1-H(D))\ p_D(D)\right) g(0) + \int_{x>0} \left(H\left(\frac{x}{\phi}\right) p_D\left(\frac{x}{\phi}\right) \frac{1}{\phi}\right) g(x)\ dx. \end{aligned} \qquad (A.4)$$

Eq. (A.4) implies that the probability measure associated with x has a point mass $\int_D (1-H(D))p_D(D)$ 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

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

for a constant k' such that $\int_{x>0} p_X(x)dx = 1$. With the Pareto density of baseline sizes (A.3),

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

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

We can calculate the distribution of S_i using a similar approach. From Eq. (A.2), a reasoning analogous to the one used for exports shows:

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

Thus, with the Pareto specification for D:

$$p_S(x) = kx^{-\zeta-1}\left[1-H(x) + H\left(\frac{x}{1+\phi}\right)(1+\phi)^{\zeta}\right] 1_{x > (1+\phi)\underline{D}} + kx^{-\zeta-1} 1_{\underline{D} < x < (1+\phi)\underline{D}}. \qquad (A.5)$$

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

Table A1
Summary statistics.

	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	2894	89,232	0	1.95E+07
Employees	152,429	44.64	626	1	168,386
Correlation (Sales, Employees) = 0.67					
	Non-exporting firms				
	No. of firms	Mean	St. dev.	Min	Max
Total Sales	90,006	6434	86,009	751	1.67E+07
Employees	86,389	25.42	583	1	168,386
Correlation (Sales, Employees) = 0.77					
	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	6777	136,456	1	1.95E+07
Employees	66,040	69.79	678	1	96,856
Correlation (Sales, Employees) = 0.79					

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	3949	0.092
Food Products	−0.661	0.011	0.962	−0.691	0.012	0.964	7031	0.164
Apparel and Leather Products	−0.842	0.030	0.970	−0.908	0.033	0.968	1529	0.368
Printing and Publishing	−0.895	0.020	0.988	−0.913	0.020	0.989	4192	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	2511	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	8722	0.347
Electrical Machinery	−0.798	0.022	0.995	−0.931	0.026	0.996	2573	0.464
Mineral Products	−0.809	0.021	0.972	−0.845	0.022	0.971	2981	0.159
Textiles	−0.855	0.033	0.938	−0.968	0.037	0.949	1354	0.333
Wood and Paper Products	−0.845	0.021	0.981	−0.908	0.023	0.979	3206	0.254
Chemicals, Plastic, and Rubber	−0.680	0.016	0.949	−0.748	0.017	0.955	3659	0.363
Metals	−0.863	0.014	0.989	−0.952	0.015	0.988	7818	0.379
Electrical and Electronic Components	−0.678	0.025	0.976	−0.770	0.029	0.978	1450	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	4632	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	3888	0.028
Hotels and Restaurants	−1.473	0.064	0.989	−1.503	0.066	0.993	1051	0.025
Recreational, Cultural, and Athletic Activities	−1.279	0.046	0.996	−1.306	0.047	0.996	1542	0.042
Health and Social Action	−1.111	0.021	0.978	−1.114	0.021	0.978	5853	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.

Table A3
Power laws in firm size, non-exporting and exporting firms: robustness to size distribution.

	CDF		PDF		ln(Rank −0.5)		
	(1) Exporters	(2) Non-exporters	(3) Exporters	(4) Non-exporters	(5) Exporters	(6) Non-exporters	(7) t-stat
I. Sales							
ζ	0.964	1.112	0.967	1.116	0.738	1.106	30.88**
	(0.042)	(0.025)	(0.041)	(0.040)	(0.004)	(0.011)	
R^2	0.981	0.995	0.996	0.997	0.972	0.998	
No. of firms	67,078	19,447	67,078	19,447	67,078	19,447	

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. The non-exporting sample has been trimmed of small firms, such that the mean of sales for non-exporting firms equals the mean of sales for exporting firms. Columns (1)–(2) estimate the CDF of the power law specified in Eq. (7). Columns (3)–(4) estimate the PDF of the power law specified in Eq. (8). Columns (5)–(6) regress ln(Rank −0.5) of the firm in the size distribution on the natural logarithm of its size (Gabaix and Ibragimov, 2011). 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 A4
Power laws in firm size by sector, non-exporting and exporting firms: employees.

	Exporting firms				Non-exporting firms				
Sector	(1) ζ	(2) Std. error	(3) R^2	(4) No. of firms	(5) ζ	(6) Std. error	(7) R^2	(8) No. of firms	(9) t-stat
Agriculture, Forestry, and Fishing	0.850	0.039	0.869	952	0.975	0.026	0.806	2868	2.66**
Food Products	0.698	0.018	0.889	2850	0.994	0.022	0.878	4082	10.28**
Apparel and Leather Products	0.778	0.033	0.900	1128	0.819	0.059	0.791	390	0.62
Printing and Publishing	0.808	0.025	0.905	2105	0.986	0.031	0.894	2013	4.48**
Pharmaceuticals, Perfumes, and Beauty Products	0.555	0.032	0.841	597	0.626	0.076	0.921	137	0.86
Furniture, Household Goods	0.789	0.028	0.923	1536	1.205	0.055	0.899	964	6.72**
Automotive	0.598	0.034	0.924	602	1.045	0.081	0.905	336	5.10**
Transport Equipment	0.553	0.040	0.940	389	0.730	0.071	0.925	212	2.18*
Non-electrical Machinery	0.809	0.018	0.918	4154	1.158	0.024	0.872	4534	11.59**
Electrical Machinery	0.756	0.029	0.947	1391	1.191	0.049	0.891	1174	7.65**
Mineral Products	0.727	0.034	0.927	917	0.960	0.031	0.855	1928	5.08**
Textiles	0.850	0.038	0.858	1004	0.995	0.076	0.861	344	1.71+
Wood and Paper Products	0.835	0.029	0.911	1693	1.131	0.041	0.880	1504	5.90**
Chemicals, Plastic, and Rubber	0.733	0.020	0.893	2601	0.903	0.040	0.889	1019	3.80**
Metals	0.875	0.018	0.929	4566	1.129	0.028	0.872	3213	7.55**
Electrical and Electronic Components	0.667	0.030	0.920	972	0.975	0.063	0.877	472	4.38**
Fuels	0.507	0.105	0.884	47	0.473	0.087	0.888	59	−0.25
Water, Gas, Electricity	0.382	0.091	0.970	35	0.523	0.044	0.892	286	1.40
Automotive Sales and Repair	0.795	0.017	0.882	4436	0.953	0.011	0.859	14,204	7.78**
Wholesale Trade, Intermediaries	0.793	0.008	0.907	19,651	0.894	0.009	0.883	19,154	8.28**
Transport	0.815	0.016	0.906	5312	0.854	0.014	0.898	7965	1.87+
Professional Services	0.760	0.012	0.925	8508	0.914	0.010	0.876	17,230	10.11**
Research and Development	0.738	0.072	0.900	210	0.648	0.076	0.888	147	−0.86
Personal and Domestic Services	0.920	0.105	0.953	153	1.142	0.054	0.877	887	1.88+
Education	0.889	0.083	0.873	231	1.041	0.041	0.875	1267	1.64+

Notes: This table reports the estimates of power laws in firm size (employees) 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.
+ Significant at the 10% level.

References

Arkolakis, C., May 2009. A unified theory of firm selection and growth, mimeo, Yale University.

Arkolakis, C., 2010. Market penetration costs and the new consumers margin in international trade. Journal of Political Economy 118 (6), 1151–1199 December.

Axtell, R.L., 2001. Zipf distribution of U.S. firm sizes. Science 293 (5536), 1818–1820 September.

Buera, F., Shin, Y., 2010. Financial Frictions and the Persistence of History: A Quantitative Exploration, NBER Working Paper No. 16400. September.

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

Chaney, T., 2008. Distorted gravity: the intensive and extensive margins of international trade. American Economic Review 98 (4), 1707–1721 September.

di Giovanni, J., Levchenko, A.A., 2010a. Country Size, International Trade and Aggregate Fluctuations in Granular Economies, RSIE Discussion Paper 585. November.

di Giovanni, J., Levchenko, A.A., 2010b. Firm Entry, Trade, and Welfare in Zipf's World, NBER Working Paper No. 16313. August.

Easterly, W., Reshef, A., Schwenkenberg, J., 2009. The Power of Exports, World Bank Policy Research Working Paper 5081. October.

Eaton, J., Kortum, S., Kramarz, F., December forthcoming. An anatomy of international trade: evidence from French firms, Econometrica.

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

Gabaix, X., 2009. Power laws in economics and finance. Annual Reviews of Economics 1, 255–293.

Gabaix, X., 2011. The granular origins of aggregate fluctuations. Econometrica 79 (3), 733–772 May.

Gabaix, X., Ibragimov, R., 2011. Rank-1/2: a simple way to improve the OLS estimation of tail exponents. Journal of Business and Economic Statistics 29 (1), 24–39 January.

Gabaix, X., Landier, A., 2007. Why has CEO pay increased so much? Quarterly Journal of Economics 123, 49–100.

Hanson, G.H., Xiang, C., 2011. Trade barriers and trade flows with product heterogeneity: an application to U.S. motion picture exports. Journal of International Economics 83 (1), 14–26 January.

Helpman, E., Melitz, M.J., Yeaple, S.R., 2004. Export versus FDI with heterogeneous firms. American Economic Review 94 (1), 300–316 March.

Hinloopen, J., van Marrewijk, C., 2006. Comparative Advantage, the Rank–Size Rule, and Zipf's Law, Tinbergen Institute Discussion Paper 06-100/1.

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

Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725 November.

Rossi-Hansberg, E., Wright, M., 2007. Urban structure and growth. Review of Economic Studies 74, 597–624.

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