Innovation trade and the size of exporting firms

Letizia Montinari, Massimo Riccaboni, Stefano Schiavo

Innovation, trade and the size of exporting firms

Letizia Montinari* Massimo Riccaboni† Stefano Schiavo‡

This draft: September 2013 – [PRELIMINARY and INCOMPLETE]

Abstract

This paper contributes to the literature explaining firm-level heterogenenity in the extensive margin of trade, defined as the number of products exported by each firm. We develop a model where firms must invest in R&D to maintain and increase their portfolio of goods: the process of product innovation by new and incumbent firms is such that the probability to capture new products is a function of the number of varieties already exported. This mechanism, together with the entry/exit dynamics that characterize the economy, gives rise to a Pareto distribution for the number of products exported by each firm. On the other hand, we model export sales as depending on exogenous preference shocks on the demand side, which leads to a lognormal distribution for the intensive margin of trade. Both predictions are consistent with a number of empirical findings recently emerged in the literature; this paper provides additional evidence based on a large dataset of French firms. Finally, a simple extension to the model allows us to derive some interesting insights on the behavior of multi-products firms: sales of different products across destinations are not uncorrelated, but show a rather strict hierarchy.

Keywords: international trade, extensive margin, innovation, preferential attachment, multi-product firms.

JEL classification: F14, F43, L11, O3

Financial support received through the project 'The international trade network: empirical analyses and theoretical models' (www.tradenetworks.it) funded by the Italian Ministry of Education, University and Research (PRIN 2009) is gratefully acknowledged.

*University of Trento and European Central Bank. e-mail: letizia.montinari@ecb.europa.eu

†IMT Institute for Advanced Studies and KU Leuven. e-mail: massimo.riccaboni@imtlucca.it

‡University of Trento and OFCE-DRIC. e-mail:stefano.schiavo@unitn.it

1 Introduction

Increasing availability of firm-level data has taught us that firm engagement in international markets differs widely. Empirical evidence suggests that this cross-sectional heterogeneity is primarily explained by the extensive margin, i.e. the difference in the number of products exported and/or destinations served by either countries or firms (Bernard et al, 2009).

In a recent contribution, Chaney (2011) discusses how existing trade models featuring heterogeneous firms (such as Melitz, 2003; Bernard et al, 2003; Chaney, 2008) are unable to make any prediction on the cross-sectional distribution of the extensive margin. Focusing on the number of destinations served by each exporter, he then proposes a model based on social network theory that accurately matches the empirical features of the data.

This paper contributes to the literature on the extensive margin of export (see Arkolakis and Muendler, 2010, for another example) by looking at the other component of the extensive margin of trade, i.e. the number of products exported by each firm. In addition, we provide a simple explanation for the very skewed distribution that characterizes also the intensive margin of trade and, by combining the two, we end up with a prediction for the size distribution of exporting firms. More specifically, we develop a model in the spirit of Simon (1955); Klette and Kortum (2004); Luttmer (2007) that describes the dynamics of both the intensive and the extensive margin of firm exports. The former is assumed to depend on a exogenous preference shock on the demand side, whereas firms must invest in R&D to maintain and increase their portfolio of goods, so that innovation governs the extensive margin of trade.

Our setup predicts a lognormal distribution for the intensive margin of trade and a Pareto distribution for the number of products exported by each firm. This second result is obtained from the dynamics of firm entry and exit and of the process of product innovation by new and incumbent firms, whereby the probability to capture new products is a function of the number of varieties already exported. Both these predictions are consistent with empirical findings recently emerged in the literature; we provide additional evidence based on a large dataset of French firms.

A simple extension to the baseline model, namely the inclusion of destination-specific fixed export costs, leads us to derive some interesting insights on the behavior of multi-product firms. In contrast with early work on multi-product firms (Bernard et al, 2011), our setup implies a rather strict hierarchy among products, so that best-selling varieties in one destination are more likely to be exported to many markets, and will sell a lot everywhere. This additional implication of the model is consistent with the evidence put forward by Arkolakis and Muendler (2010) and is further reflected in our data.

A large literature has documented the main empirical regularities that characterize international trade flows at various levels of aggregation. Moreover, many studies have

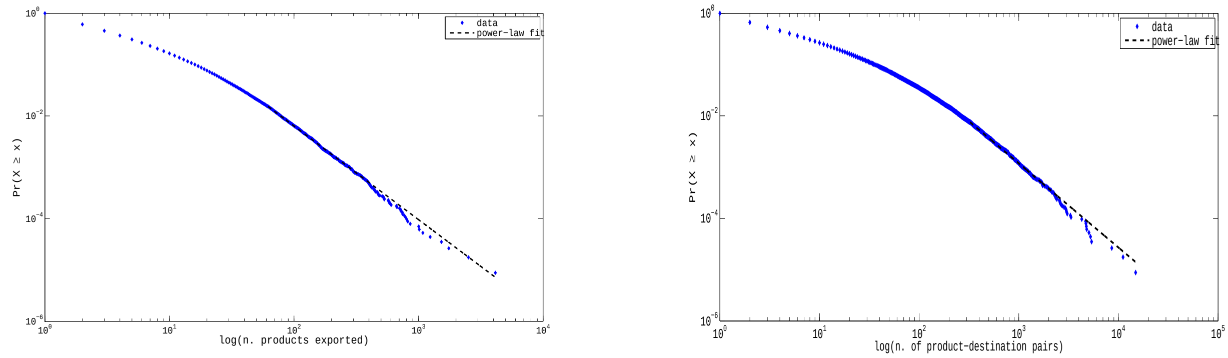

Figure 1: Top: Distribution of the number of product exported by firm, double logarithmic scale, with power-law fit. Bottom: Distribution of the number of product-destination pairs by firm, double logarithmic scale, with power-law fit. Data refers to year 2003.

further investigated the features of the different dimensions in which trade flows can be decomposed: the intensive and the extensive margin (see for instance Bernard et al, 2009). In a nutshell, what emerges from this body of work is that both margins are characterized by large and persistent heterogeneity, and this features hold across different countries and levels of aggregation. Hence, for instance, one learns that most exporting firms ship only a few products to a small number of foreign destinations (often just one), and that even at the country level, the most important products or destination markets account for the bulk of total export (Easterly et al, 2009, label these large flows "big hits" and show they are very rare).

In this paper we look at data on more than 100 000 French firms in the year 2003 exploiting data collected by the French Custom Service. Figure 1 shows the distribution of the extensive margin: irrespective of whether we look at the number of products exported (left panel) or at the combination of product-destination pairs (right panel), the distributions appear very skewed, and a power-law fit provides a good approximation of the data.[1]

Table 1: Number of products exported by each firm: summary statistics. Data refers to 2003 and are classified according to the 8-digit CN (Combined Nomenclature).

mean	7.31		
std. dev	26.57	% of firms selling:	
min	1.00	only 1 product	39.24
25th percentile	1.00	>10 products	14.96
median	2.00	>100 products	0.63
75th percentile	6.00	>100 products	0.007
max	4140.00		

Table 1 provides summary statistics that further characterize the very large heterogeneity in the data. While the number of products exported (according to the 8-digit

[1]The power-law fit is obtained using the methodology described in Clauset et al (2009).

Combined Nomenclature) ranges between 1 and 4140, 39.24% of firms export a single product, the median value is 2 and less than 15% of firms export more than 10 products.

Finally, Figure 2 shows the distribution of export sales (in logarithms) of products by different firms for 2003: the distribution is well approximated by a truncated normal.

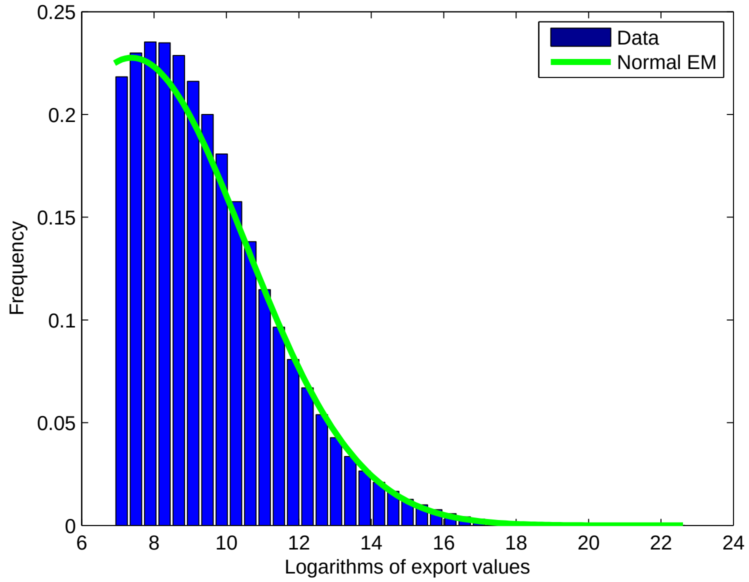

Figure 2: Distribution of log export sales by firm and product, with superimposed truncated normal density. Data for 2003, CN 8-digit classification.

The empirical regularities presented here appear robust to the specific year analyzed and to the level of aggregation chosen, both in terms of digits of the specific classification employed and the way export flows are identified (product Vs product-destination pair).

Having laid down the man empirical facts that motivate the paper, in the next Section we present the baseline version of the model, discuss its main implications, and perform some empirical analyses to validate them. Section 3 extends the model to consider fixed export costs and multi-products firms, and look at the relevant empirical evidence. Finally, Section 4 concludes.

2 Baseline Model

Firms are distributed over a finite set of C identical countries and engage in costly trade (iceberg type). At any time t, each country is populated by a continuum of identical consumers of measure $H_t = He^{\eta t}$, where $\eta \geq 0$ is the growth rate of the population. Time is continuum and denoted by t, with initial time $t = 0$. At each point in time, the representative consumer is endowed with one unit of labor and has the following utility function

$$U_t = E_t \left[\int_t^{\infty} ln(X_t) e^{-\rho t} dt \right] \tag{1}$$

where $\rho > 0$ is the discount factor and X_t is a composite good. The differentiated good X_t is produced with a continuum of varieties

$$X_t = \left(\int_{\omega \in \Omega_t} a_t(\omega) x_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where $x_t(\omega)$ is consumption of variety ω, $a_t(\omega)$ is a preference shock distributed lognormal which hits variety ω at time t, $\sigma > 1$ is the elasticity of substitution across varieties, and Ω_t is total mass of varieties at time t.[2] In each country, the composite good sector consists of a large group of monopolistically competitive producers. At any point in time, each variety ω is produced by one and only one firm.

Each variety is produced according to the following production technology

$$x_t = z_t l_t \tag{3}$$

where z is labor productivity which is common across varieties and firms. We assume that z evolves exogenously over time according to $z_t = ze^{\theta t}$.

2.1 Households

The representative household maximizes utility subject to a standard budget constraint. The corresponding first order conditions are

$$\frac{\dot{Y}}{Y} = r - \rho \tag{4}$$

$$x_t(\omega) = a_t(\omega) \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \frac{Y_t}{P_t} \tag{5}$$

where r is the interest rate and $P_t = \left(\int_{\omega \in \Omega_t} a_t(w) p_t(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}$ is the price index. Total household expenditure on the composite good X is

[2]Lognormally distributed preference shocks would result from a process of "impeded growth" à la Kalecki (1945) whereby a Gibrat diffusion process is stabilized by relaxing the original assumption that shocks are independent of size and assuming instead that there is a negative correlation between size and growth: $a_t(\omega) = a_{t-1}(\omega)^{(1-\beta)}$ $\forall t$. A possible interpretation for this process is a product life-cycle type of dynamic: young varieties are subject to higher shocks, i.e. are more volatile, while as products get older they reach maturity and their volatility decreases.

$$\int_{\omega \in \Omega_t} x_t(\omega) p_t(\omega) d\omega = Y_t$$

Equation (4) is the standard Euler equation and (5) is the demand for variety ω. We let total expenditure Y_t be the numeraire and set it to a constant over every period implying $r = \rho$ in the balanced growth path.

2.2 Production and Innovation

Trade costs are symmetric and of the standard iceberg type: $\tau > 1$ units shipped result in 1 unit arriving. Since there are no entry costs in the export market, active firms sell both to the domestic and to the foreign markets. At each point in time, the representative firm maximizes profits subject to the demand function (5) coming from the domestic and from the foreign markets. The pricing rule of the firm in the domestic market and in a foreign market are $p_{d,t}(\omega) = (\sigma/(\sigma-1))(w_t/z_t)$ and $p_{f,t}(\omega) = (\sigma/(\sigma-1))(w_t/z_t)\tau$ respectively. Revenues from sales of variety ω in the domestic market and in a foreign market are

$$r_{d,t}(\omega) = a_t(\omega) \left(\frac{p_{d,t}(\omega)}{P_{d,t}} \right)^{1-\sigma} Y_{d,t}$$

$$r_{f,t}(\omega) = a_t(\omega) \left(\frac{p_{f,t}(\omega)}{P_t^F} \right)^{1-\sigma} Y_{f,t}$$

As $r_{f,t}(\omega) = r_{d,t}(\omega)\tau^{1-\sigma}$, total revenues from domestic and foreign sales of variety ω are $r_t^{Tot}(\omega) = r_{d,t}(\omega) + (C-1)r_{d,t}(\omega)\tau^{1-\sigma}$. Profits from the domestic market a from a foreign market are

$$\pi_{d,t}(\omega) = a_t(\omega) \left(\frac{\sigma w_t}{(\sigma-1)P_{d,t}} \right)^{1-\sigma} \frac{Y_{d,t}}{\sigma} z_t^{\sigma-1}$$

$$\pi_{f,t}(\omega) = a_t(\omega) \left(\frac{\sigma w_t}{(\sigma-1)P_{f,t}} \right)^{1-\sigma} \frac{Y_{f,t}}{\sigma} z_t^{\sigma-1}$$

Total profits from the domestic market and the foreign markets are $\pi_t^{Tot}(\omega) = \pi_{d,t}(\omega) + (C-1)\pi_{d,t}(\omega)\tau^{1-\sigma}$.

2.2.1 Innovation by incumbents

To increase its portfolio of goods, a firm must invest in R&D activities. We assume that new innovations arrive following an exponential distributed waiting time with mean

$$\lambda_t = f(i_t) \tag{6}$$

where i_t is labor used in the replication process. We assume that f is increasing and exhibits strictly decreasing returns to scale. Each firm faces also the probability that some firm will innovate over a good it is currently producing. When this event occurs, the incumbent producer loses the good from its portfolio. An existing variety is lost with an exponentially distributed waiting time with mean

$$\mu_t = g(j_t) \tag{7}$$

where j_t is labor used to maintain existing products and g is (i) strictly decreasing and convex. We can rewrite the profits as $\frac{wl}{\sigma-1}n\phi$ where $\phi = [1 + (C-1)\tau^{1-\sigma}]$. The value of a firm with n product must satisfy the Bellman equation

$$\begin{aligned} r_t V_t(n) = \underset{\lambda,\mu}{max} E_t\{w_t n \left[\frac{l_t}{\sigma-1}\phi - (i+j)\right] + \lambda n \left[V_t(n+1) - V_t(n)\right] + \\ + \mu n \left[V_t(n-1) - V_t(n)\right]\}. \end{aligned} \tag{8}$$

The expectation operator is conditional on information at time t. As firms only learn about the preference shock over a new variety after they succeed in the innovation process, they expect to receive the same flow of profits from each innovation. Moreover, as each innovation guarantees the same expected flow of profits, all firms choose the same innovation rate λ_t and maintenance rate μ_t.

As in Klette and Kortum (2004), the unique solution to (8) must be proportional to the number of products $V(n) = vn$. The optimal levels of investment in new varieties and of maintenance of existing varieties are determined by

$$\lambda_t = f(i_t) \quad \mu_t = g(j_t) \quad v_t f'(i_t) = -v_t g'(j_t) = w_t \tag{9}$$

Intuitively, the innovation process described here suggests that firms base their investment policy only on the number of products they already export, independently of whether these products rank high or low in consumers' preferences (i.e. independently of the $a_t(\omega)$). This implies a form of compensation across different products exported by the same firm: profits generated by "best sellers" make up for the lower-than-average profits generated by products characterized by low $a_t(\omega)$. Such a result can easily be accommodated for firms exporting a large number of products (n_L), as total profits are likely to differ only slightly from those obtained from exporting n_L times the "average" product (i.e. a product yielding the average profit). On the other hand, it is possible that "unlucky" firms selling only a few products (n_S) characterized by low $a_t(\omega)$ may be unable to invest *as if* they were getting "average" profits. When this happens, then we could observe some departures from the equilibrium distribution of the number of products exported by firms (whose shape is discussed in Section 2.4 below) in the lower tail of the distribution.

2.2.2 Innovation by entrants

New varieties can also be produced from scratch by agents acting as entrepreneurs. At each point in time, agents are endowed with one unit of effort that can be allocated between two tasks: supplying labor or producing a new variety. Following Luttmer (2011), we assume that each agent has a skill vector (x, y), where x corresponds to the rate at which agents generate a new variety and y is the amount of labor supplied per unit of time. Agents with skill vectors that satisfy $v_t x > w_t y$ will become entrepreneurs, while agents with skills vectors that satisfy $v_t x < w_t y$ will supply labor to existing firms.

Let T be a time-invariant talent distribution defined over the set of all possible skill vectors with finite mean and density ψ. The resulting per capita supply of entrepreneurial effort is

$$E(v_t/w_t) = \int_{v_t x > w_t y} x dT(x, y) \tag{10}$$

for $\pi \in \Pi$. The per capita supply of labor is

$$L(v_t/w_t) = \int_{v_t x < w_t y} y dT(x, y) \tag{11}$$

Given a per capita stock of entrepreneurial activities $E(v_t, w_t)$ and a stock of varieties N_t, the rate ν_t at which new entrepreneurs add a new variety is determined by

$$\nu_t N_t = H_t E(v_t, w_t) \tag{12}$$

Labor market clearing requires

$$N_t(l_t + i_t + j_t) = H_t L(v_t, w_t) \tag{13}$$

2.3 Balanced growth

Along the balanced growth path, the measure of varieties grows at the rate η. The allocation of labor is constant (i, j, l). From the consumer's problem, aggregate variables w_t and C_t grow at a rate $k = \theta + \frac{\eta}{(\sigma-1)}$ with a rate that is larger when goods are less substitutable. The implied interest rate is $r = \rho + k$. The Bellman equation (8) implies that wages and the values of firms must satisfy

$$\frac{v}{w} = \frac{\frac{l}{\sigma-1}\phi - [i + j]}{r - k - [\lambda - \mu]} \tag{14}$$

where (i, j) and (λ, μ) satisfy (9).

Holding l fixed, these conditions imply that v/w is equal to the maximum subject to $[\lambda, \mu] = [f(i), g(j)]$ of the right-hand side of equation (14), as log as it is finite. As the total number of varieties grows at rate η, new entrepreneurs must contribute at the non-negative rate $\eta - [\lambda - \mu]$. If $E(v, w)$ is positive, from (12) we obtain the entrepreneurial steady-state supply of varieties

$$\frac{N}{H} = \frac{E(v, w)}{\eta - [\lambda - \mu].} \tag{15}$$

Alternatively, $E(v, w) = 0$ and $\eta = \lambda - \mu$. Along the balanced growth path, the market clearing condition becomes

$$\frac{N(l + i + j)}{H} = L(v, w) \tag{16}$$

Luttmer (2011) shows that if $\rho + k > k + \eta$ and $\eta > f(0) - g(0)$, for a positive $E(v, w)$, then equations (9), (14), (15) and (16) define the unique balanced growth path and $\eta > \lambda - \mu$. A balanced growth path can arise with $E(v, w) = 0$ if the talent distribution has bounded support. In this case, new varieties are only produced by existing firms.

2.4 The distribution of the number of products exported

In absence of fixed export costs, all active firms sell both to the domestic and to the foreign markets. A new variety can then be thought as a new trade link. Firms form new trade links by creating new commodities and lose trade links when some firms innovate over a good they are currently producing. It follows that one can identify the growth process of the number of products for an individual firm with the distribution of its links. Let us define $M_{n,t}$ the mass of a firm with n products at time t. The aggregate measure of products is

$$N_t = \sum_{n=1}^{\infty} n M_{n,t}. \tag{17}$$

The change in the number of firms with one commodity over time is

$$\dot{M}_{1,t} = \mu 2 M_{2,t} + v N_t - (\mu + \lambda) M_{1,t} \tag{18}$$

where λ, μ and $\nu = \eta - [\lambda - \mu]$ are constant along the balanced growth path. The number of firms with one commodity increases because firms with two commodities loses one or because of entry. The number decreases because firms with one commodity gain or lose one.

The number of firms with more than one commodity evolves according to

$$\dot{M}_{n,t} = \lambda(n-1) M_{n-1,t} + \mu(n+1) M_{n+1,t} - (\mu + \lambda) n M_{n,t} \tag{19}$$

for all $n - 1 \in N$. A stationary distribution for a firm exists if (18) and (19) have a solution for which $\frac{M_{n,t}}{N_t}$ is constant over time. Since along the balanced growth path N grows at rate η, $\dot{M}_t = \eta M_{n,t}$ for all $n \in N$. Given that N and M_n grow at the same rate

η, we can define

$$P_n = \frac{M_{n,t}}{\sum_{n=1}^{\infty} M_{n,t}} \tag{20}$$

for all $n \in N$. Equation (20) gives the fraction of firms with n commodities. We can also define the fraction of all commodities produced by firms of size n as

$$Q_n = \frac{nM_{n,t}}{\sum_{n=1}^{\infty} nM_{n,t}} \tag{21}$$

for all $n \in N$. Using these definitions we can rewrite (18) and (19) as

$$\eta Q_1 = \mu Q_2 + \upsilon - (\lambda + \mu) Q_1 \tag{22}$$

$$\frac{1}{n}\eta Q_n = \lambda Q_{n-1} + \mu Q_{n+1} - (\lambda + \mu) Q_n \tag{23}$$

Luttmer (2011) provides a solution for (22)-(23) in Appendix A of his paper. Luttmer (2011) shows that if λ, μ, η and $\nu = \eta - (\lambda - \mu)$ are positive, the sequence $\{\beta_n\}_{n=0}^{\infty}$ defined by the recursion $\beta_n = 1/(1 - (\lambda\beta_n/\mu) + (\eta + \lambda n)/\mu n)$ and the initial condition $\beta_0 = 0$ is monotone and converges to $min\{1, \mu/\lambda\}$. The only non-negative and summable solution to equations (22)–(23) is given by

$$Q_n = \frac{\nu}{\lambda} \sum_{k=0}^{\infty} \frac{1}{\beta_{n+k}} \left(\prod_{m=n}^{n+k} \beta_m \right) \prod_{m=n}^{n+k} \frac{\lambda\beta_m}{\mu} \tag{24}$$

For large n and $\lambda \neq \mu$ the distribution satisfies

$$Q_n \sim \frac{\nu}{\mid \lambda - \mu \mid} \prod_{m=1}^{n-1} \frac{\lambda\beta_m}{\mu} \tag{25}$$

If $\nu = 0$, the only non-negative and summable solution to equations (22)–(23) is identically zero, implying that there does not exist a stationary distribution in this case. If $\nu > 0$, equation (24) adds up to 1 by construction and defines a stationary distribution $\{P_n\}_{n=1}^{\infty}$ via $P_n \propto \frac{Q_n}{n}$. The mean number of links of a firm can be written as $1/\left(\sum_{n=1}^{\infty} Q_n/n\right)$ which is finite by construction.

If $\lambda < \mu$, Q_n is bounded above by a multiple of the geometrically declining sequence $(\lambda/\mu)^n$. When $\lambda > \mu$ then $(\lambda\beta_n/\mu) \uparrow 1$ and (25) declines at a rate that is slower than any given geometric rate.

Luttmer (2011) shows that under some parameter restrictions the connectivity distribution features a Pareto tail with a shape parameter greater than unity. If $\eta > 0$, $\lambda > \mu$ and $\eta > \lambda - \mu$, then the right tail probabilities $R_n = \sum_{k=n}^{\infty} P_k$ of the stationary connectivity distribution satisfy

$$\lim_{n\to\infty} n\left(1 - \frac{R_{n+1}}{R_n}\right) = \xi \quad (26)$$

where $\xi = \frac{\eta}{(\lambda-\mu)}$. That is, R_n is a regularly varying sequence with index $-\xi$ and $\xi > 1$.[3]

2.5 Implications of the model and empirical validation

Our model provides clear predictions for the evolution of both the intensive and the extensive margin of trade. The assumed evolution of preference shocks gives rise to a lognormal distribution of export sales for products (Kalecki, 1945). This is consistent with the empirical findings recently emerged in the empirical literature: in Section 1 we have shown that the lognormal distribution provides a good fit for the distribution of exports values by products (Figure 2).

Looking at distribution of export sales however is only one part of the story. In fact, this results depends on our assumption on preference shocks: assuming a negative correlation between size and growth (as in Kalecki, 1945) modifies the original Gibrat result in a way that stabilizes the resulting lognormal distribution, so that its variance no longer diverges. Kalecki's model of impeded growth can be validated by looking at the deviations of (the logs of) product sales from their mean value (see for instance Bottazzi et al, 2001). In particular, we run an OLS estimation of the following expression:

$$g_i(t) = \beta g_i(t-1) + \varepsilon_i(t) \quad (27)$$

where $g_i(t) = s_i(t) - \bar{s}(t)$, $s_i(t)$ is (the log of) firm size (defined in terms of export sales in the present context), and $\bar{s}_i(t)$ is (the log of) the average size of all firms. A pure Gibrat process would entail $\beta = 1$, whereas the impeded growth hypothesis implies $\beta < 1$.

Table 2: Test for "impeded growth"

size($t-1$)	0.846***
	(0.002)
constant	-0.634***
	(0.008)
Observations	430 494
Firms	67 661
R-squared	0.708
F-test: $\beta == 1$	
p-value	0.000

clustered standard errors (by firm) in parentheses
*** significant at 1%

Table 2 reports the results from OLS estimation of equation (27) where standard

[3] When the rate $\nu = \eta - (\lambda - \mu)$ goes to zero, the limiting tail index $\xi = 1$ associated with Zipf's law arises.

errors are clustered at the firm level to account for the possible correlation of observations pertaining to the same firm. The estimated coefficient is smaller than one ($\hat{\beta} = 0.846$), and an F-test formally rejects the hypothesis $\beta = 1$, thus lending credit to our assumption.

For what concerns the extensive margin, i.e. the number of products exported by each firm, the model generates a Pareto distribution. We look at the data on export by French firms in 2003, and run multiple statistical tests to detect the presence of a power-law in the data. In particular, we apply three different methodologies, namely the uniformly most powerful unbiased (UMPU) test developed by del Castillo and Puig (1999), the maximum entropy (ME) test by Bee et al (2011), and the test proposed by Gabaix and Ibragimov (2011).[4]

Table 3: Test for power-law upper-tail behavior in extensive margin (data for 2003)

	N. products by firm
UMPU	2100
	(1.83%)
	29.76%
ME	2300
	(2.01%)
	31.03%
GI	2680
	(2.34%)
	33.24%

First row: number of observations in power-law tail
Second row: percentage of observations in power-law tail
Third row: share of total num. of products in power-law tail

Table 3 displays the results of the different tests, which identify the number of (the largest) observations that belong to the power-law upper tail of the distribution. The three tests are in good agreement and find a power-law behavior limited to the top 2100–2680 observations. This is small in terms of number of observations, covering roughly 2% of the sample size, but nonetheless represents one third of all the products in the dataset.

The observation that the lower tail of the connectivity distribution does not follow a power-law can be rationalized by noting that, as already anticipated in Section 2.2.1 above, the innovation process described in the model assumes all firms invest *as if* they were getting "average" profits on all their products. This is not a constrains for firms exporting a large number of products, as a poor performance by some products can be compensated by other goods; on the contrary, it is possible that firms selling only a few "poor" products are constrained in their ability to invest in innovation since their revenues may be too low.[5] If this is the case, then the cumulative process that lies at the core of our

[4]For a quick overview of the different methodologies see Bee et al (2013).

[5]The model does not consider the financial system and the possibility that firms can access external resources. However, it is well known that financing innovation activities is particularly difficult given the intrinsic uncertainty of the process and of the associated returns.

model may not properly work, leading to some departures from the Pareto distribution in the lower tail, as we indeed observe in the data.

The absence of fixed export costs implies that each innovation leads to a new trade link. As a result, the evolution of the number of goods produced by a firm describes the dynamics of the number of trade links or the connectivity distribution of that firm. The growth in the number of varieties (due to population growth) and the dynamics of firm entry and exit ensures that the right tail of the connectivity distribution follows a Pareto with a tail index greater than unity (Luttmer, 2007). This prediction is consistent with the data, where we find of a shape parameter for the Pareto tail of the distribution of the extensive margin (i.e. number of products by firm) ranging between 1.73 and 1.83.

A further implication of the model has to do with the size of exporting firm, and results from the combination of the above two results. Growiec et al (2008) show that the size distribution of business firms depend on the shape of the distribution its components, namely the number of products sold (the extensive margin n) and the value of sales of each product (the intensive margin). In particular, when the former is Pareto and the latter is lognormal, the distribution of firm sales is a lognormal distribution multiplied by a stretching factor which increases with n: when n is small, the stretching factor is negligible and the distribution is close to a lognormal; on the contrary, for large n the size distribution shows a departure that leads to the emergence of a Pareto upper tail.

We look at this prediction of the model by running the three statistical tests used above (UMPU, ME and GI) on export sales at two different levels of aggregation: first we take export by firm-product and then total export by firm. Under the assumption of monopolistic competition each firm produces a different variety, so that export sales by firm-product represent the empirical counterpart of "products" as defined in the model.

Table 4: Test for power-law upper-tail behavior in export values. Data for 2003 at different levels of aggregation (by firm-product and by firm)

	Total export by	
	firm-product	firm
UMPU	650	400
	(0.08%)	(0.35%)
	43.17%	55.49%
ME	850	800
	(0.10%)	(0.70%)
	46.70%	65.45%
GI	1230	590
	(0.15%)	(0.51%)
	51.71%	61.05%

First row: number of observations in power-law tail
Second row: percentage of observations in power-law tail
Third row: share of total trade in power-law tail

Table 4 shows that indeed the length of the power-law tail in the data increases upon aggregation, with the three tests in good agreement. The share of observations belonging to the Pareto upper tail spans a larger share of the sample: depending on the specific methodology adopted, the Power-law is 3 to 7 times larger (as a share of the sample) than the one we observe in disaggregated data. Even if it is still limited to the very top of the distribution of export sales, the power-law behavior in the upper tail accounts for 55–65% of total exports (up from 43–51% when looking at product-level data).

3 Extension: Fixed Export Costs

An important extension of the model entails the inclusion of fixed entry costs in foreign markets. For the sake of clarity, we move one step at a time and first look at what happens when fixed costs are symmetric across countries. In a second step we relax this assumption and extend the model to the more interesting case of country-specific fixed costs.

3.1 Symmetric case

If the entry costs associated with export are equal across destinations, we have that (i) firms that have captured only (few) low-selling products are prevented from exporting; (ii) there is hierarchy among products whereby only the best-selling items (those hit by higher preference shocks) are exported.

As productivity is the same across firms and varieties and fixed export costs are equal across destinations, whether a product is exported or not only depends on the preference shock. Let us define a^*as the *preference* cut-off level which makes profits from selling variety ω in a foreign market equal to zero $a_t^* = sup\{a : \pi_t^F(a) = 0\}$

As before, profits from selling variety ω in the domestic market are

$$\pi_t^D(\omega) = a_t(\omega) \left(\frac{\sigma w_t}{(\sigma - 1)P_t^D} \right)^{1-\sigma} \frac{Y_t^D}{\sigma} z_t^{\sigma-1} \tag{28}$$

whereas profits from sales into the foreign markets become

$$\pi_t^{F_c} = a_t(\omega) \left(\frac{\sigma w_t \tau}{(\sigma - 1)P_t^{F_c}} \right)^{1-\sigma} \frac{Y_t^F}{\sigma} z_t^{\sigma-1} - f_e \tag{29}$$

where f_e represents the (symmetric) entry costs into a foreign market.

Thus, under the assumptions that country as symmetrical so that $Y_t^D = Y_t^F$, total profits from a non-exported variety are $\pi_t^D(\omega)$, and profits from an exported variety are $\pi_t^{Tot} = \pi_t^D \phi - f_e(C-1)$, with $\phi = [1 + (C-1)\tau^{1-\sigma}]$.

As before, firms only learn about the preference shock on a new variety once they succeed in the innovation process. Therefore, they expect to receive the same flows of profits from each new variety and thus choose the same innovation rate λ and maintenance rate μ.

The Bellman equation becomes

$$r_t V_t(n) = \underset{\lambda,\mu}{max} E_t\{w_t n_t \left[\alpha \frac{l_t}{\sigma-1} + \beta\left(\frac{l_t}{\sigma-1}\phi - f_e(C-1)\right) - (i+j)\right] + \\ + \lambda n \left[v V_t(n+1) - V_t(n)\right] + \mu n \left[V_t(n-1) - V_t(n)\right]\} \tag{30}$$

where α and β are the fractions of products sold in the domestic and in the foreign markets respectively and $\phi = [1 + (C-1)\tau^{1-\sigma}]$.

Finally, equation (14) becomes

$$\frac{v}{w} = \frac{\alpha \frac{l_t}{\sigma-1} + \beta\left(\frac{l_t}{\sigma-1}\phi - f_e(C-1)\right) - [i+j]}{r - k - [\lambda - \mu]}. \tag{31}$$

Equations (22) and (23) still describe the evolution of the fraction of commodities produced by a firm, but they do not describe the evolution of the number of its trade links as not all new varieties will be now exported. However, since the number of products sold by each firm and the preference shocks that determines whether each product is exported or not are independent, the number of products exported by each firms is simply a random sample from the overall population of products sold domestically.[6] Since a random sample taken from a power-law follows the same distribution, we can still conclude that the distribution of the number of exported varieties is power-law. In this respect then, the introduction of fixed entry costs does not entail any significant change to the implications of the model, but simply provide us a mechanism to explain the existence of non-exporting firms.

3.2 Contry-specific fixed costs

A more interesting case is the one issuing from relaxing the assumption of symmetric countries. More specifically, suppose there are heterogeneous fixed entry cost into foreign markets (F_j, with $j = 1, \ldots, C$ indexing destinations) and that we can order rank countries on the basis of these fixed entry cost such that $F_1 < F_2 < \cdots < F_C$. This leads us to derive some further interesting (and more realistic) implications relative to the behavior of multi-product firms. In particular, there will be a is hierarchy among products whereby top-selling items are more likely to be shipped to many destinations (see Arkolakis and Muendler, 2010, for evidence along these lines).

[6] In fact, as it is common in this class of models, we also have that the range of exported varieties is a subset of products available domestically. In other words, there are no products that are exported but not consumed at home.

This is in contrast with early work on multi-product firms (Bernard et al, 2011), that assume sales across markets are uncorrelated, but appears to be consistent with recent empirical evidence

[...]

3.3 Empirical evidence on multi-product firms

Arkolakis and Muendler (2010) present empirical evidence based on a large sample of Brazilian firms to show that there is a strict cross-country hierarchy in export sales by multi-product firms. This is to say that successful products in one market are also best-sellers in other destinations. In particular, they look at the US and Argentina (Brazil's two main export destinations) as reference markets, and compare export behavior there with export in other markets. They find that, within a firm, the best-selling products in the reference market have higher sales than other products in all other markets. Indeed, the rank-correlation between sales in the US (Argentina) and sales in the rest of the world is as high as 0.837 (0.860). Furthermore, lower ranked products tend to be shipped to fewer destinations.

[...]

4 Conclusion

We develop a dynamic model of innovation by new and incumbent firms in the spirit of Klette and Kortum (2004) and Luttmer (2007) where firms must invest in order to capture new products and maintain their existing portfolio. This process gives rise to a cumulative dynamic whereby large firms tend to invest more and therefore grow even larger. The power-law distribution of the number of products exported by each firm well approximates the data, so that the model provides a novel explanation for the large heterogeneity in the extensive margin among exporting firms.

Moreover, the model predicts a lognormal distribution of export sales of each product and the emergence of a power-law tail in the size distribution of exporting firms. Both these predictions are well matched by data on a large sample of French exporting firms.

[...]

References

Arkolakis C, Muendler MA (2010) The extensive margin of exporting products: A firm-level analysis. Working Papers 16641, NBER

Bee M, Riccaboni M, Schiavo S (2011) Pareto versus lognormal: A maximum entropy test. Physical Review E 84:026,104, DOI 10.1103/PhysRevE.84.026104

Bee M, Riccaboni M, Schiavo S (2013) The size distribution of US cities: Not Pareto, even in the tail. Economics Letters 120(2):232–237

Bernard A, Eaton J, Jensen JB, Kortum S (2003) Plants and productivity in international trade. American Economic Review 93(4):1268–1290

Bernard A, Jensen JB, Redding S, Schott P (2009) The margins of us trade. American Economic Review 99(2):487–93

Bernard AB, Redding SJ, Schott PK (2011) Multiproduct firms and trade liberalization. The Quarterly Journal of Economics 126(3):1271–1318

Bottazzi G, Dosi G, Lippi M, Pammolli F, Riccaboni M (2001) Innovation and corporate growth in the evolution of the drug industry. International Journal of Industrial Organization 19(7):1161–1187

del Castillo J, Puig P (1999) The best test of exponentiality against singly truncated normal alternatives. Journal of the American Statistical Association 94:529–532

Chaney T (2008) Distorted gravity: The intensive and extensive margins of international trade. American Economic Review 98(4):1707–21

Chaney T (2011) The network structure of international trade. Working Papers 16753, NBER

Clauset A, Shalizi C, Newman M (2009) Power-law distributions in empirical data. SIAM Review 51:661–703

Easterly W, Reshef A, Schwenkenberg J (2009) The power of exports. Policy Research Working Paper Series 5081, World Bank

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

Growiec J, Pammolli F, Riccaboni M, Stanley HE (2008) On the size distribution of business firms. Economics Letters 98(2):207–212

Kalecki M (1945) On the gibrat distribution. Econometrica 13(02):161–170

Klette TJ, Kortum S (2004) Innovating firms and aggregate innovation. Journal of Political Economy 112(05):986–1018

Luttmer EGJ (2007) Selection, growth, and the size distribution of firms. The Quartely Journal of Economis 122(03):1103–1144

Luttmer EGJ (2011) On the mechanics of firm growth. The review of Economics Studies 78(03):1042–1068

Melitz MJ (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71(6):1695–1725

Simon HA (1955) On a class of skew distribution functions. Biometrika 42(3/4):425–440

Appendices

A Household's problem

The representative household maximizes the following intertemporal utility function subject to an intertemporal budget constraint

$$U_t = E_t \left[\int_t^{\infty} ln(X_t) e^{-\rho t} dt \right]$$

$$\dot{A}_t = r_t A_t + w_t - X_t P_t$$

where X_t is a composite good, P_t is the price of the composite good, w_t is the wage rate, and A_t is the value of the household's asset holdings. At any period t, the representative consumer is endowed with one unit of labor. Total spending in final good at t is $Y_t = P_t X_t$.

The consumer's problem is solved in two steps.

A.1 First step: dynamic consumption problem

The current value Hamiltonian is

$$H(C_t, A_t, v_t) = ln(C_t) + v_t[r_t A_t + w_t - X_t P_t]$$

The first order conditions are

$$X_t : \; v_t = \frac{1}{X_t P_t} = \frac{1}{Y_t} \tag{32}$$

$$A_t : \; \frac{\dot{v_t}}{v_t} = \rho - r_t \tag{33}$$

Taking the time derivative of (28), we get

$$\frac{\dot{v_t}}{v_t} = -\frac{\dot{Y_t}}{Y_t}$$

using (29), we get the standard Euler equation

$$\frac{\dot{Y_t}}{Y_t} = r_t - \rho$$

We set Y_t to a constant over every period implying $r = \rho$ in the balanced growth path.

A.2 Second step: static choice across varieties

The representative consumer chooses the optimal bundle of varieties to consume given its budget constraint

$$X_t = \left(\int_{\omega \in \Omega_t} a_t(\omega) x_t(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

$$\int_{\omega \in \Omega_t} x_t(\omega) p_t(\omega) d\omega = Y_t$$

The corresponding first order condition gives the demand function for variety ω

$$x_t(\omega) = a_t(\omega) \left(\frac{p_t(\omega)}{P_t} \right)^{-\sigma} \frac{Y_t}{P_t}$$

B The problem of the firm

Let $p_t^{DD}$and x_t^{DD} be price and quantity for the domestic market and $p_t^{DF_c}$,$x_t^{DF_c}$ prices and quantities for the foreign market.

The representative firm maximizes the following profit function:

$$\pi_t = \left(p_t^{DD} x_t^{DD} - \frac{w_t x_t}{z_t} \right) + \sum_{c=1}^{C-1} \left(p_t^{DF_c} x_t^{DF_c} - \frac{w_t x_t}{z_t} \tau \right)$$

subject to

$$x_t^{DD} = a_t \left(\frac{p_t^{DD}}{P_t^D} \right)^{-\sigma} \frac{Y_t^D}{P_t^D}$$

$$\sum_{c=1}^{C-1} x_t^{DF_c} = a_t \left(\frac{p_t^{DF_c}}{P_t^{F_c}} \right)^{-\sigma} \frac{Y_t^{F_c}}{P_t^{F_c}}$$

The resulting first order conditions give the price for the domestic and the foreign market respectively

$$p_t^{DD} = \frac{\sigma}{\sigma - 1} \frac{w_t}{z_t}$$

$$p_t^{DF_c} = \frac{\sigma}{\sigma - 1} \frac{w_t}{z_t} \tau$$