Working paper

2015-06

INNOVATION, TRADE AND THE SIZE OF EXPORTING FIRMS

Letizia Montinari
European Commission, Joint Research Centre, Institute for Prospective Technological Studies

Massimo Riccaboni
IMT Institute for Advanced Studies, Lucca
KU Leuven

Stefano Schiavo
OFCE – SciencesPo Paris and University of Trento

March 2015

Innovation, trade and the size of exporting firms

Letizia Montinari* Massimo Riccaboni† Stefano Schiavo‡

This draft: March, 2015

Abstract

This paper contributes to the literature explaining firm-level heterogeneity in the extensive margin of trade, defined as the number of products exported by each firm. We develop a dynamic model where firms must invest in R&D to maintain and increase their portfolio of goods: the process of product innovation by incumbent firms is such that the probability to capture new products is a function of the number of varieties already exported. Varieties can also be produced from scratch by new entrepreneurs. The entry/exit dynamics of varieties, together with population growth that characterize the economy, gives rise to a distribution for the number of products exported by each firm with a heavy right tail, which is consistent with the data. This markedly heterogeneous behavior in export markets occur even if we do not assume any heterogeneity in productivity to start with. On the other hand, we assume that differences in export sales across products originate from the demand-side of the model, in the form of a product-specific preference attribute. Finally, a simple extension of 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. We thank Michele Bernini, Mauro Caselli, Francesco Di Comite, Andrea Fracasso, Michele Imbruno and Simone Salotti for helpful comments on an earlier draft of the paper. All the usual disclaimers apply.

*Corresponding author: European Commission, Joint Research Centre (JRC), Institute for Prospective Technological Studies (IPTS), C/Inca Garcilaso, 3, E-41092 Seville, Spain. Email: Letizia.MONTINARI@ec.europa.eu. The views expressed in this article are those of the author and do not necessarily reflect the official policy or position of the European Commission.

†IMT Institute for Advanced Studies, Lucca (Italy) and KU Leuven (Belgium). E-mail: massimo.riccaboni@imtlucca.it

‡University of Trento (Italy) and OFCE-DRIC (France). 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 (2014) 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 by looking at the other component of the extensive margin of trade, i.e. the number of products exported by each firm (Arkolakis and Muendler, 2010, have similar aims).

We achieve this goal by means of a dynamic model in the spirit of Simon (1955); Klette and Kortum (2004) and Luttmer (2011) in which incumbent firms must invest in R&D to enlarge the portfolio of goods they produce and export. In particular, the more products a firm has, the more resources it can devote to R&D and the more likely it will capture new opportunities. This cumulative growth process governs the evolution of the extensive margin of trade and gives rise to a skewed distribution of the number of products exported by each firm with a heavy right tail. Such a distribution closely matches the empirical evidence, as we show by looking at comprehensive data on French manufacturing firms.

One of the remarkable features of the paper, that we regard as one of its main contributions, is its ability to generate large heterogeneity in firm behavior (most notably in the number of product exported) despite firms are *ex ante* identical and homogeneous in all their characteristics. In fact, we do not need to assume a skewed productivity distribution to obtain the result of firms producing and selling a (very) different number of products. Rather, as in Klette and Kortum (2004) and Luttmer (2011), it is the cumulative and stochastic nature of the process of innovation that drives the heterogeneity in the extensive margin.

A simple extension to the baseline model, namely the inclusion of destination-specific fixed export costs, allows us to derive some interesting insights on the behavior of multi-product firms as well. 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 in every market. This additional implication of the model is consistent with the recent evidence put forward by Arkolakis and Muendler (2010) and Mayer et al (2014).

Relative to the existing literature, we build on the theoretical framework developed by Luttmer (2011), and extend it to the open economy. In this setup, which enjoys a long history in industrial economics at least since Simon (1955), the dynamics of firm entry and exit and the process of product innovation by new and incumbent firms generate an advantage for large firms, which are more successful in securing new business opportunities (i.e. new products). We find that this approach lends itself very well to explain the thick-tailed distribution of the number of product sold by each firm (the extensive margin of trade).

A second stream of the trade literature on which we draw is comprised by recent studies investigating the behavior of multi-product firms (e.g. Arkolakis and Muendler, 2010; Bernard et al, 2011; Mayer et al, 2014). In fact, we borrow from Bernard et al (2011) the idea that each product variety is characterized by an "attribute" that represents consumers' taste for that product, generates heterogeneity in export sales across varieties, and determines the behavior of the *intensive* margin of trade. Hence, we introduce heterogeneity from the demand rather than the supply side of the model. The importance of demand-side influences in determining dispersion in firm size and revenue-based productivity has been emphasized by Syverson (2004, p. 549), and exploited by Di Comite et al (2014) to explain differentiation in export markets.

Last, we also touch upon the size distribution of business firms, as the model is consistent with the results presented by Growiec et al (2008), who find that the very skewed distribution of firm size can emerge through the interaction between the intensive and extensive margin of a firm's portfolio of products. Our model's prediction on the size distribution of firms is also in line with the empirical evidence based French firms that we present in this paper.

The paper is organized as follows: the next section provides a quick glance at the data that motivates the paper and we seek to explain, Section 3 presents the model, while Section 4 goes back to the empirical findings and discusses them in light of the theory. Section 5 provides additional evidence on multi-product firms and, by including fixed export costs, extends the model to match those empirical regularities. Finally, Section 6 concludes.

2 A Glance at the Empirical Facts

A large literature has documented a series of empirical regularities that characterize international trade flows at various levels of aggregation. Moreover, many studies have 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, most exporting firms ship only a few products to a small number of foreign destinations (often just one), and the most important products or destination markets account for the bulk of total export, both at the firm and the country level (Easterly et al, 2009, label these large flows "big hits" and show they are very rare). This Section presents evidence based on data covering more than 30 000 French (manufacturing) exporting firms in the year 2003. The data are collected by the French Customs Service and are similar to those used in a number of other studies (e.g. Eaton et al, 2011; Mayer et al, 2014).[1] It is worth noting that the evidence presented here below is robust to the specific year analyzed, and to the level of aggregation chosen, both in terms of digits of the specific classification and the way export flows are identified (product Vs product-destination pair).[2]

Figure 1 shows the distribution of the extensive margin: irrespective of whether we look at the number of products exported or at the combination of product-destination pairs (meaning that we treat the same product shipped to different destinations as different varieties), the distributions appear very skewed, and a power-law fit provides a good approximation of the data.[3]

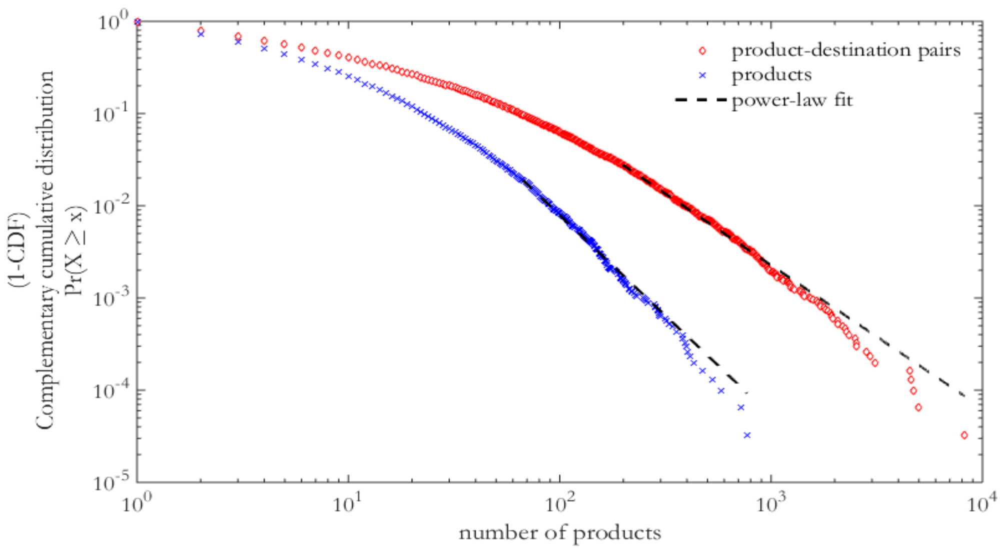

Figure 1: Complementary cumulative distribution of the number of products (blue) and product-destination pairs (red) exported by firm in double logarithmic scale, with power-law fit. Data refers to the year 2003.

The claim that the two distributions are characterized by heavy tails can be further

[1]The model developed in the paper refers to all goods *produced* by the firm, not only those *exported*. Empirically, however, we do not have information on the varieties that are not exported and thus take the number of products exported as a proxy for those actually produced. It is worth noting that the model remains valid even if some of the varieties produced are not traded due to fixed export costs.

[2]Chaney (2014) analyzes cross-sectional data covering the years 1986–1992 and finds results that are qualitatively similar to ours in terms of the heterogeneity in the extensive margin.

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

substantiated by calculating their *obesity index*, as proposed by Cooke et al (2014).[4] The values taken by the obesity index are 0.821 (number of products) and 0.872 (number of product-destination pairs): these are closer to the theoretical value implied by a Zipf's law ($\pi^2 - 9 \approx 0.87$) than to the values associated with, say, an exponential distribution (0.75).

Table 1 provides summary statistics that further characterize the very large heterogeneity in the extensive margin (here defined as the number of products exported) of French firms. While the number of products exported by each firm (according to the 8-digit Combined Nomenclature) ranges between 1 and 770, 26.91% of firms export a single product, the median value is 4 and less then a quarter of firms export more than than 10 products.

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	9.67		
std. dev	20.96	% of firms selling:	
min	1.00	only 1 product	26.91
25th percentile	1.00	1–5 products	40.43
median	4.00	>10 products	23.13
75th percentile	10.00	>100 products	0.81
max	770.00	>500 products	0.01

Export sales are also characterized by a high degree of heterogeneity, both within and across firms. Most papers featuring heterogeneous firms postulate a Pareto or power-law distribution of productivity which, under the common assumption of a CES demand structure, translates into a power-law distribution of export sales (and firm size).[5]

While the power-law assumption is very convenient from a modeling point of view because of the peculiar analytical properties of this distribution, recent evidence suggests that a lognormal distribution provides a better fit to the data (Head et al, 2014). Figure 2 confirms this hypothesis: the main panel depicts an histogram of (log) export sales by firm and product, and a (truncated) normal fit. The inset presents a Q-Q plot that compares the empirical quantiles of the data against the theoretical quantiles derived from a truncated lognormal distribution whose parameters are estimated from the data. The plot confirms that the (truncated) lognormal provides a very good approximation of the data.[6]

[4] The obesity index is based on the heuristic that, in the case of heavy-tailed distributions, larger observations lie further apart than smaller observations, and is defined as $Ob(X) = prob(X_1 + X_4 > X_2 + X_3 | X_1 \leq X_2 \leq X_3 \leq X_4)$, where X_i are values randomly sampled from the data. In our application, the index is computed by bootstrapping based on $1,000$ random samples of 4 observations.

[5] Most if not all the papers that postulate a power-law productivity distribution refer to Axtell (2001) to justify empirically their choice.

[6] The parameters of the (truncated) lognormal distribution against which we compare the actual trade values are estimated using the EM algorithm discussed in Bee (2006).

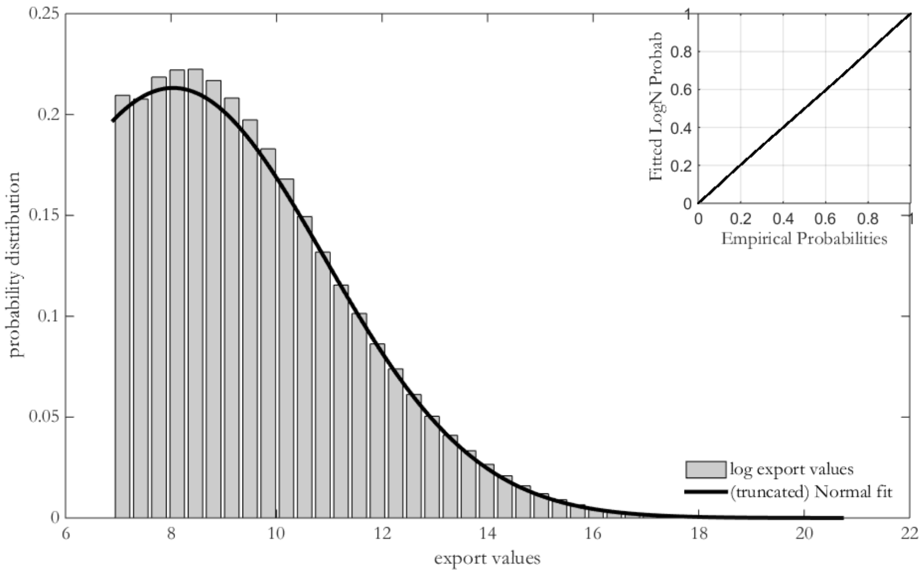

Figure 2: Distribution of log export sales by firm and product, with superimposed truncated normal density. Inset: quantile-quantile plot of empirical versus theoretical quantiles. Data for 2003, CN 8-digit classification.

In any case, both the lognormal and the power-law distribution imply a high degree of heterogeneity in export sales, with a small fraction of very large products and a large majority of small sales. Bee et al (2011) discuss the difficulty associated with discriminating the tail behavior of the two distributions in actual samples, while (as noted also by Head et al, 2014) the main benefit of the lognormal is to provide a good fit of the data also for the body of the distribution, not only for its upper tail.

While the sum of power-law distributed random variables still follows a power-law, no closed form solution exist for the sum of lognornal distributions. This implies that, if product sales follow a lognormal distribution, firm size (defined as total sales of all products) cannot be characterized precisely. The debate on the exact shape of the distribution of firm size is a very old one in economics and its discussion goes beyond the scope of this Section. Here it suffices to say that our data confirm the most recent evidence pointing toward significant departures from a power-law distribution (Rossi-Hansberg and Wright, 2007; Bernard et al, 2011, see for instance) and closer resemblance to lognormality. Using data on French firms Bee et al (2014) confirm this latter finding and add that lognormality appears particularly suitable to describe the size distribution of single-product firms, whereas a power-law upper tail emerges for the class of multi-product companies.

Comparing the distribution of export sales at the product level with aggregate exports by firms, we also find that the second distribution displays heavier tails and features an (upper) tail behavior consistent with a power-law. We look at this feature of the data first by looking at the obesity index: the index increases from 0.918 (product-destination level) to 0.924 (summing over destinations for each distinct product within firms) , to 0.981

(total exports by firm) indicating that aggregation gives rise to a heavier tail. Then, we run three statistical tests aimed at discriminating between power-law and lognormal tail behavior, 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 procedure proposed by Gabaix and Ibragimov (2011, GI henceforth).[7] Table 2 reports the number of (largest) observations for which the hypothesis of power-law behavior cannot be rejected at the 5% confidence level together with the relative percentile of the distribution and the share of total export sales that belong to this power-law tail. The first column refers to export sales for each product, while the second pertains to total export by firm. In either case, power-law behavior is restricted to a very small number of observations: the different tests gives slightly different results, but they all agree in finding a very limited power-law. This is consistent with the view that the lognormal offers a much better fit to the data. However, Table 2 does show that the length of the power-law tail increases upon aggregation, and again the three tests are in good agreement. Depending on the specific methodology adopted, the power-law is 6 to 10 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 68–73% of total exports (up from 48–53% when looking at product-level data).

Table 2: 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	550	500
	(0.19%)	(1.97%)
	48.12%	68.11%
ME	750	600
	(0.25%)	(1.64%)
	52.65%	71.03%
GI	800	700
	(0.27%)	(2.30%)
	53.61%	73.46%

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

The next Section outlines a dynamic open economy model in which firms invest in innovation in order to enlarge their portfolio of goods. Such a setup manages to replicate many of the empirical regularities just described, and that, for clarity, we summarize below:

- **FACT 1.** Within firms, the distribution of export sales by product (intensive

[7]For details on the properties of the test and their theoretical underpinning, see Bee et al (2011)

margin) follows a normal distribution.

- **FACT 2.** The distribution of the number of products by firm (extensive margin) is characterized by a heavy upper tail which is well approximated by a power-law fit.
- **FACT 3.** The distribution of firms size, defined as total exports sales for all products, follows a normal distribution with a power-law behaviour in the upper tail explained by the presence of multi product firms.

3 The Model

Firms are distributed over a finite set of C identical countries. 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 $t = 0$.

3.1 Households

The intertemporal utility of the representative consumer is

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

where $\rho > 0$ is the discount factor and X_t denotes aggregate consumption. X_t is a CES composite of differentiated goods

$$X_t = \left[\sum_{j=1}^{N_t} a_j^{\frac{1}{\sigma}} \left(x_{j,t} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{2}$$

where $x_{j,t}$ is consumption of variety j, $\sigma > 1$ is the elasticity of substitution across varieties, and N_t is total mass of varieties at time t, and a_j is an exogenous product-specific attribute that captures consumer tastes, in a way similar to Bernard et al (2011).[8] The preference attribute a_j comes from a time-invariant distribution $\Gamma(a)$ which is continuous in the domain $0 < a < \infty$ with mean $\bar{a}$.[9] We assume that $\Gamma(a)$ is common to all firms in

[8] A similar assumption is made by Eaton et al (2011) and Easterly et al (2009) who assume that an exogenous demand shock specific to good i in a given market n affects consumer demand for that variety. As the preference shock on variety j affects symmetrically consumer tastes for that particular variety, a_j can be interpret as the α_s parameter in Di Comite et al (2014), that is an index of quality of variety j (vertical differentiation). Our choice to model firms heterogeneity as coming from consumer tastes is consistent with the findings by Syverson (2004) who shows that demand-side conditions play an important role in explaining persistent firm-level dispersion.

[9] By assuming $\Gamma(a)$ to be a lognormal distribution we obtain that export sales are also lognormally distributed, a feature that closely match our data. A similar assumption is done in Eaton et al (2011), who assume that the logarithm of demand shocks follow a normal distribution with zero mean and finite variance, and show that this specific distributional choice provides estimates matching data on sales by

each country. Furthermore, in each country, the number of varieties that firms produce may vary as the result of a stochastic innovation process.

The representative household maximizes utility subject to the standard budget constraint. The resulting demand for a variety j is

$$x_{j,t} = a_j \left(\frac{p_{j,t}}{P_t}\right)^{-\sigma} \frac{Y_t}{P_t} \tag{3}$$

where $Y_t = \sum_{j=1}^{N_t} p_{j,t} x_{j,t}$ is total household expenditure on the composite good X and P_t is the price index defined as

$$P_t = \left[\sum_{j=1}^{N_t} a_j \left(p_{j,t}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}.$$

Here, as in the original formulation by Dixit and Stiglitz (1977), we are assuming that N_t is large enough so that each variety has no effect on the aggregate price index. Aggregate consumption growth and the interest rate are related via the standard Euler equation

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

with r being the interest rate and ρ the discount factor.

3.2 Firms

Labor is the sole factor of production, and each variety j is produced according to the following technology:

$$x_{j,t} = z_{j,t} l_{j,t} \tag{5}$$

where $l_{j,t}$ is the amount of labor used in the production process and $z_{j,t}$ is labor productivity.

Firms sell their products in both the domestic and in foreign markets: in this baseline setting, there are no entry costs in export markets, implying that all varieties produced are sold in the domestic market and in all foreign markets. To ship goods abroad, firms incur symmetric trade costs of the standard iceberg type: $\tau > 1$ units shipped result in 1 unit arriving the destination market. Firms set the price of each variety j in order to maximize (static) profits, given the wage rate w_t, yielding the standard result that the optimal price is given by a mark-up over marginal costs:

$$p_{j,t} = \frac{\sigma}{(\sigma - 1)} \frac{w_t}{z_{j,t}} \tau. \tag{6}$$

French manufacturing firms.

Indicating with $p_{j,t}^D$ and $p_{j,t}^F$ the domestic and foreign price respectively, and given that $\tau = 1$ for domestic sales, we can write $p_{j,t}^F = p_{j,t}^D \tau = \frac{\sigma}{(\sigma-1)} \frac{w_t}{z_{j,t}} \tau$ (see Appendix B for details on the derivations).

Revenues from sales of variety j in the domestic and in a foreign market are given by:

$$r_{j,t}^D = a_j \left(\frac{p_{j,t}^D}{P_t^D} \right)^{1-\sigma} Y_t^D \tag{7}$$

$$r_{j,t}^F = a_j \left(\frac{p_{j,t}^F}{P_t^F} \right)^{1-\sigma} Y_t^F \tag{8}$$

If all countries are identical, so that $P_t^D = P_t^F$ and $Y_t^D = Y_t^F = Y_t$, then $r_{j,t}^F = r_{j,t}^D \tau^{1-\sigma}$. Total revenues from sales of variety j in the domestic and all the (C-1) foreign countries are thus equal to $r_{j,t} = r_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right]$.

Profits from sales of variety j in the domestic and in a foreign market are given respectively by:

$$\pi_{j,t}^D = a_j \frac{Y_t}{\sigma} \left(\frac{\sigma}{\sigma - 1} \frac{w_t}{z_{jt} P_t} \right)^{1-\sigma} = \frac{r_{j,t}^D}{\sigma} \tag{9}$$

$$\pi_{j,t}^F = a_j \frac{Y_t}{\sigma} \left(\frac{\sigma}{\sigma - 1} \frac{w_t \tau}{z_{jt} P_t} \right)^{1-\sigma} = \frac{r_{j,t}^F}{\sigma} \tag{10}$$

As $\pi_{j,t}^F = \pi_{j,t}^D \tau^{1-\sigma}$, total profits from sales of variety j in all markets can be expressed as $\pi_{j,t} = \pi_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right]$.

At each point in time, for a firm i with $n_{i,t}$ products, total revenues and profits equal

$$r_{i,t}^{Tot}(n_{i,t}) = \sum_{j=1}^{n_{i,t}} r_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right] \tag{11}$$

$$\pi_{i,t}^{Tot}(n_{i,t}) = \sum_{j=1}^{n_{i,t}} \pi_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right]. \tag{12}$$

We assume that all firms are equal in terms of productivity and, moreover, that all varieties yield the same revenues and profits. This is necessary to derive the balanced growth path of the economy in a tractable way, as shown by Luttmer (2011).

In what follows we assume that productivity is common across all varieties and firms, i.e. $z_{j,t} = z_t$, so that varieties are all sold at the same price ($p_{j,t} = p_t$).[10] The common productivity z_t evolves exogenously over time according to the following law of motion: $z_t = z e^{\theta t}$.

[10] In Appendix D we show an alternative version of our model where heterogeneity on the demand side is obtained using the more common assumption of productivity heterogeneity across firms. As shown there, our main results are not affected by the chosen approach to model heterogeneity in the intensive margin because CES preferences and monopolistic competition firm productivity and consumer tastes have similar effects on revenues (Bernard et al, 2011).

To be able to characterize the balanced-growth path of the economy, it is convenient to express average revenues and profits for a firm with $n_{i,t}$ products as a function of the average preference attribute hitting its varieties $\tilde{a} = \sum_{j=1}^{n_t} \frac{1}{n_t} a_j$:

$$r_{i,t}(\tilde{a}, n_{i,t}) = \sum_{j=1}^{n_{i,t}} \frac{1}{n_t} r_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right] \tag{13}$$

$$\pi_{i,t}(\tilde{a}, n_{i,t}) = \sum_{j=1}^{n_{,it}} \frac{1}{n_t} \pi_{j,t}^D \left[1 + (C-1)\tau^{1-\sigma}\right]. \tag{14}$$

As the portfolio of goods $n_{i,t}$ that a firm produces increases, by the law of large numbers we expect $\tilde{a} \to \bar{a}$, so that $r_{i,t}(\tilde{a}) \simeq r_{i,t}(\bar{a})$ and $\pi_{i,t}(\tilde{a}) \simeq \pi_{i,t}(\bar{a})$, where $\bar{a}$ is the population average of the distribution $\Gamma(a)$. As a consequence, as long as the number of products is large enough, the average revenues and profits per variety will be approximately the same for all firms. In other words, this means that the intensive and extensive margins are independent of each other (as, for instance, in Bernard et al, 2011) and firm revenues and profits only depend on the number of products sold.[11] Intuitively, we are assuming that firms devote the same amount of resources to all products: intra-firms adjustments may well occur as long as they do not impact on the total amount of resources used by each firm (conditional on the number of products sold $n_{i,t}$). Best-selling products, which require higher levels of production will absorb more labor, compensating the existence (within the firm) of "below-average" products that use fewer resources.[12]

This assumption is crucial to allow us to characterize the balanced growth path and to derive a closed-form solution for the distribution of the number of products sold by each firm (see Section 3.5 below). In fact, in order to have a single distribution we need that firms' investment choices only depends on the number of products in their portfolio (as in Klette and Kortum, 2004 or Luttmer, 2011), disregarding the possibility that each product may generate a different level of revenues and profits. If, on the other hand, firms exporting many products were more likely to produce a "best selling" variety, so that average profits per product were positively associated with the number of products sold, the cumulative process of preferential attachment that underlies the model's dynamics would be magnified and our conclusions would be reinforced.

[11]From an empirical point of view, we find that the number of product exported by a firm barely correlates with average export per product: considering firms selling at least two products, the correlation coefficient is 0.03 (an analogous lack of substantial correlation between the intensive and extensive margins of trade is reported by Bernard et al, 2007 for US firms and by Bernard et al, 2014 for Belgian ones).

[12]This may not be true for firms with a small number of products for which the average preference attribute may be different than $\bar{a}$. The existence of those firms has important implications for the shape of the distribution of the number of products by firm that we will discuss in Section 4.

3.3 New varieties

We follow Luttmer (2011) to model the entry/exit dynamics of varieties in our economy. New varieties can be produced by incumbent firms as a result of innovation activities. New varieties can also be produced from scratch by agents acting as entrepreneurs. The respective rates at which these two events occur in equilibrium are λ and ν. Existing varieties can also be lost when firms innovate over existing products. The rate at which this occurs in equilibrium is μ. Finally, firms with only one product exit the market when some other firm innovates over the good they are currently producing. Thus, the number of varieties evolve according to

$$\dot{N} = (\nu_t + \lambda_t - \mu_t)N_t \tag{15}$$

An initial condition determines N_0, but is irrelevant for the equilibrium outcome.

3.3.1 Innovation by incumbents

To increase and maintain its portfolio of goods, a firm must invest in R&D activities. We assume that new innovations arrive following a Poisson process with exponentially distributed waiting time of the form

$$\lambda_t = f(i_t) \tag{16}$$

where i_t represents the resources (labor) invested in the innovation 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 variety it is currently producing. We assume that a firm innovating over an existing variety can produce a better type of that variety without changing the product specific attribute that captures consumer tastes over that . When this event occurs, the incumbent producer loses that variety from its portfolio. An existing variety is lost with an exponentially distributed waiting time with mean

$$\mu_t = g(h_t) \tag{17}$$

where h_t is labor used to "maintain" existing products and g is strictly decreasing and convex.[13] Given the constant markup $\sigma/(\sigma-1)$ of price over the marginal cost, firms' average profit per variety $\pi_t(\tilde{a})$ can be rewritten as $wl/(\sigma-1)\phi$ where $\phi = [1 + (C-1)\tau^{1-\sigma}]$ (see Appendix C for derivations). The average value of a variety $v_t(\tilde{a})$ must satisfy the

[13]This is one of the main departures from Klette and Kortum (2004), who assume a constant and homogeneous degree of "creative destruction" μ.

Bellman equation

$$r_t v_t(\tilde{a}) = \max_{\lambda,\mu} \left[w_t \left(\frac{l_t}{\sigma - 1}\phi - (i+h) \right) + v_t(\tilde{a})(\lambda - \mu) + \dot{v}_t(\tilde{a}) \right]. \tag{18}$$

From an economic point of view, the first term inside the square brackets corresponds to profits coming from variety j at time t minus the costs associated to innovation (i) and imitation (h) activities for that particular variety. The second term captures the expected gain from innovating over variety j and the expected loss from having another firm innovating over variety j.

The optimal levels of investment in new varieties and of maintenance of existing varieties are determined by

$$\lambda_t = f(i_t) \quad \mu_t = g(h_t) \quad v_t(\tilde{a})f'(i_t) = -v_t(\tilde{a})g'(h_t) = w_t \tag{19}$$

Since the distribution of the product-specific preference attribute $\Gamma(a)$ is constant and common to all firms, the law of large numbers implies that, as the number of products per firm grows large, the average value per variety $\tilde{a}$ tends to the population average $\bar{a}$, that is common to all firms. Hence, abstracting for the moment from deviations pertaining to firms with low $n_{i,t}$, along the balanced growth path, there will be a unique innovation rate λ and maintenance rate μ common to all firms.

3.3.2 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 $v_t(\tilde{a})x > w_t y$ will become entrepreneurs, while agents with skills vectors that satisfy $v_t(\tilde{a})x < w_t y$ 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

$$E(v_t(\tilde{a})/w_t) = \int\limits_{v_t(\tilde{a})x > w_t y} x dT(x, y) \tag{20}$$

for $\pi \in \Pi$. Per capita labor supply is

$$L(v_t(\tilde{a})/w_t) = \int\limits_{v_t(\tilde{a})x < w_t y} y dT(x, y). \tag{21}$$

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

$$\nu_t N_t = H_t E(v_t(\tilde{a})/w_t). \tag{22}$$

Labor market clearing requires

$$N_t(l_t + i_t + h_t) = H_t L(v_t(\tilde{a})/w_t). \tag{23}$$

3.4 Balanced growth

As the average value per variety is the same across firms, along the balanced growth path, firms will allocate (on average) the same fraction of labor (i, h, l) per variety. This implies that the measure of varieties will grow at the same rate of population η. From the consumer's problem, wages w_t and per capita consumption $c_t = \frac{X_t}{H_t}$ 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 (18) implies that wages and the average values of a variety must satisfy the present-value condition:

$$\frac{v(\tilde{a})}{w} = \frac{\frac{l}{\sigma-1}\phi - [i+h]}{r - k - [\lambda - \mu]} \tag{24}$$

where (i, h) and (λ, μ) satisfy (19).

As the total number of varieties grows at rate η, new entrepreneurs must contribute at the non-negative rate $\eta - [\lambda - \mu]$. If $E(v(\tilde{a})/w)$ is positive, from (22) we obtain the entrepreneurial steady-state supply of varieties

$$\frac{N}{H} = \frac{E(v(\tilde{a})/w)}{\eta - [\lambda - \mu]}. \tag{25}$$

Alternatively, $E(v(\tilde{a})/w) = 0$ and $\eta = \lambda - \mu$. Along the balanced growth path, the market clearing condition will be

$$\frac{N(l + i + h)}{H} = L(v, w). \tag{26}$$

Luttmer (2011) shows that if $\rho > \eta$ and $\eta > f(0) - g(0)$, for a positive $E(v, w)$, then equations (19), (24), (25) and (26) 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.

3.5 The distribution of the number of products exported

With no fixed export costs, each variety is sold in the domestic and in all foreign markets: hence, each new product represents $C-1$ new trade links. Firms form new trade links by creating new commodities and lose trade links when some firm innovates 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 trade links (i.e. connectivity distribution in network jargon). Let us define $M_{n,t}$ the mass of firms with n products at time t. The aggregate measure of products is

$$N_t = \sum_{n=1}^{\infty} nM_{n,t}. \tag{27}$$

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

$$\dot{M}_{1,t} = \mu 2M_{2,t} + vN_t - (\mu+\lambda)M_{1,t} \tag{28}$$

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 lose 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

$$\dot{M}_{n,t} = \lambda(n-1)M_{n-1,t} + \mu(n+1)M_{n+1,t} - (\mu+\lambda)nM_{n,t} \tag{29}$$

for all $n-1 \in N$. A stationary distribution for a firm exists if (28) and (29) 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

$$P_n = \frac{M_{n,t}}{\sum\limits_{n=1}^{\infty} M_{n,t}} \tag{30}$$

for all $n \in N$. Equation (30) gives the fraction of firms with n commodities. We can also define the fraction of all commodities produced by firms of size n as

$$Q_n = \frac{nM_{n,t}}{\sum\limits_{n=1}^{\infty} nM_{n,t}} \tag{31}$$

for all $n \in N$.

Using these definitions we can rewrite (28) and (29) as

$$\eta Q_1 = \mu Q_2 + v - (\lambda+\mu)Q_1 \tag{32}$$

$$\frac{1}{n}\eta Q_n = \lambda Q_{n-1} + \mu Q_{n+1} - (\lambda+\mu)Q_n. \tag{33}$$

Luttmer (2011) provides a solution for (32)-(33). He also shows that under some parameter restrictions a stationary distribution of the number of products by firm exists and features a Pareto tail with a shape parameter greater than unity. In particular, $\nu > 0$ assures that a stationary distribution exists. Then, 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

$$\lim_{n\to\infty} n\left(1 - \frac{R_{n+1}}{R_n}\right) = \xi \tag{34}$$

where $\xi = \frac{\eta}{(\lambda-\mu)}$ and R_n is a regularly varying sequence with index $-\xi$ and $\xi > 1$.[14]

We shall now return to the stylized facts presented in Section 2 above, and discuss how the theoretical framework presented here can accommodate them.

4 Discussion

Our model exploits a dynamic process of innovation similar to those presented by Klette and Kortum (2004) and Luttmer (2011) to derive an equilibrium where firms produce a different number of varieties. In our setting, firms also sell different quantities of each variety due to the preference attributes associated with each of them. In the baseline model, as there are no entry costs in export markets, all varieties will be sold both in the domestic and in all foreign markets. This implies that the number of varieties produced by firms coincides with the the number of varieties they export, as happens in the empirical analysis where we only have information on products shipped abroad.

A crucial assumption of the model is that the product-specific attributes representing consumers' taste come from the same time-invariant distribution. This implies that as the portfolio of goods of a firm increases, its average preference attribute $\tilde{a}$ will better approximate the mean ($\bar{a}$) of the distribution $\Gamma(a)$. As a consequence, average revenue and profit per variety will approximately be the same across firms, allowing us to characterize the balanced growth path and to derive a unique distribution of the number of products by firm.

As in (Luttmer, 2011), the thick right tail of the distribution of the number of products exported by each firm is the result of the stochastic dynamics of entry and exit of varieties and of population growth. In particular, population must grow at a positive rate $\eta > 0$. In addition, along the balanced growth path, the probability of capturing a new variety must be higher than the probability of losing one of them: $\lambda > \mu$. These conditions,

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

together with $\eta > (\lambda - \mu)$, give rise to a distribution for the number of products exported by each firm featuring a Pareto upper tail with a tail index greater than unity. This is consistent with our findings that suggest the presence of heavy tails in the distribution of the extensive margin.[15]

In addition, the existence of firms with a small number of products can explain some departures from the equilibrium distribution of the number of products. In particular, 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 3.3.1 above, the innovation process described in the model assumes all firms invest *as if* they were getting "average" profits on all their products. While this is not a constraint for firms exporting a large number of products —as poor performance by some varieties can be compensated by other items featuring above-average sales— it is possible that firms selling only a few "unpopular" products are constrained in their ability to invest in innovation since their revenues may be too low.[16] If this is the case, then the cumulative process that lies at the core of our model may not properly work for firms in the lower tail of the distribution, leading to some departures from the power-law distribution, as it appears in Figure 1 above.

In Section 2 we have shown that, upon aggregation, a power-law upper tail emerges in the distribution of export sales (by firm), which is otherwise well approximated by a lognormal distribution at the level of single product. This behavior results from the interaction of the extensive and intensive margin of export in a context where firm size can be measured by total export sales (if, as in our baseline model, there are no fixed export costs). In fact, Growiec et al (2008) show that when the sales of each product (the intensive margin here) follow a lognormal distribution and the number of products sold by each firm (n, the extensive margin) follows a power-law, the distribution of total 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. In the next Section, we present an extension of the model with country-specific fixed export costs, which delivers some interesting features consistent with the behavior of multi-product firms.

[15]Estimates of the shape parameter of a possible power-law fit to our data range between 1.5 and 2.3 depending on the methodology adopted and the focus on product Vs product-destination pairs.

[16]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.

5 Extension: Country-Specific Fixed Export Costs

An interesting extension of the model entails the inclusion of country-specific fixed entry costs in foreign markets. Indeed, this allows the model to replicate a number of empirical regularities concerning the behavior of multi-product firms in different markets. We present a quick overview of the relevant empirical evidence and then illustrate how a simple modification of the model can account for these additional facts as well.

5.1 Evidence on multi-product firms

A number of recent papers present evidence pointing toward a rather strict hierarchy among the products exported by each firm in different markets. In particular, Mayer et al (2014) compute the global ranking of each product sold by each firm (based on total export sales) and compare it with the local ranking, i.e. the ranking within each destination market. They report an average correlation of 0.68, which appears not driven by firms exporting 1 product to 1 destination only, but rather reflect a broader phenomenon. Indeed, the correlation is still 0.59 for firms exporting more than 50 products to more than 50 countries. Arkolakis and Muendler (2010) take a slightly different route to look at the same phenomenon, namely the persistence of product ranking across destination markets. Using a large sample of Brazilian firms, they focus on two reference markets, the US and Argentina (Brazil's two main export destinations), and compare export behavior there with export in all other markets. They find that, within a firm, the best-selling products in the reference market have higher sales in all other markets as well. 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.

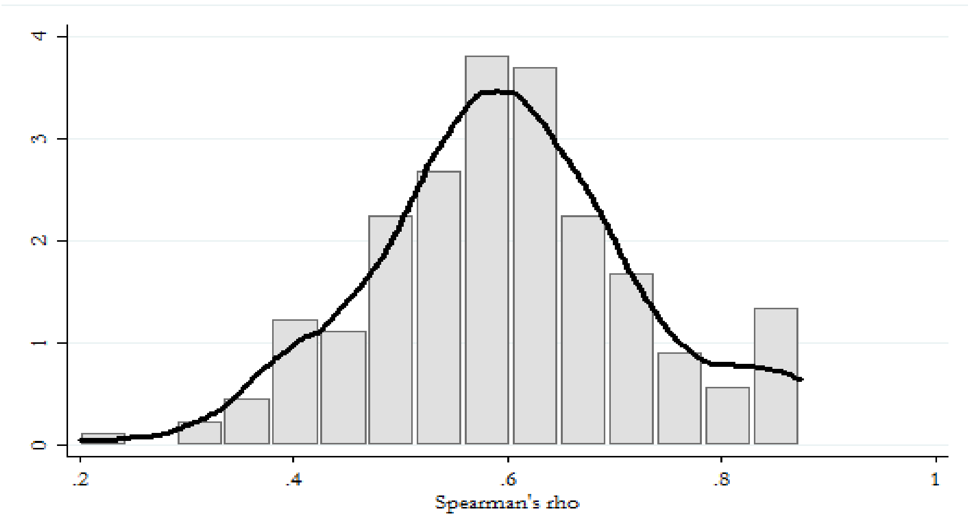

Figure 3: Distribution of correlation coefficients between global and local ranks

Here we add some further evidence based on our own data. First of all, we replicate the approach by Mayer et al (2014) and find results that are comparable to theirs, as reported in Table 3 and illustrated in Figure 3. The average rank correlation across the 199 destinations for which we have at least 20 observations is 0.60, and the values range between a minimum of 0.20 and a maximum of 0.87. Interestingly, the largest values refer to France main trading partners (in 2003), namely Germany, Spain and the UK. This suggests that in large (and closer) markets the ranking of products is more strongly correlated, whereas marginal destinations (where fewer products are exported) tend to drive down the value of the correlation. The distributions of Spearman's correlation coefficients between the global and each local ranking appears well behaved, with higher frequency clustered near the mean value and a symmetric shape (see Figure 3).

Table 3: Correlation between global and local ranking of products

destination	Spearman's rho	observations
Germany	0.874	69188
Spain	0.872	59868
UK	0.872	53666
overall	mean	0.599
	st. dev.	0.126
	min	0.200
	2nd quarter	0.520
	median	0.596
	3rd quarter	0.676
	max	0.874

Table 4 looks more systematically to the issue by using France's three main trading partners as reference markets. In so doing, we are able to replicate some of the analyses described in Arkolakis and Muendler (2010) and add further interesting information. Overall, results are very similar to those reported for Brazilian firms. For each reference market, column 1 of Table 4 reports the average number of destinations served by varieties with a given rank in the reference market. So, for instance, we see that the top selling variety in Germany ships on average to 12.46 markets, whereas a firm's 32nd most popular product in Spain is exported to roughly 10 countries. Column 2 displays the average number of foreign markets to which a firm with at least as many products as the corresponding rank ships. Clearly, firms with more products tend to serve more markets. However, among these destinations, higher ranking products cover a higher proportion (see column 3): *overlap* represents the share of destinations served by the firm, which are covered by products of a given ranking. This percentage goes down rather quickly: while the top product is shipped to 67% of all destinations covered by firms serving at least one destination other than Germany, the share goes down to 48% for the second product and 34% for the fourth. A very similar patters is found for the other two reference markets, namely Spain and the UK.

Table 4: Overlap between reference country and rest of the world by product rank

	Reference market: Germany				Reference market: Spain				Reference market: UK			
rank*	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
1	12.46	18.68	67%	11097	13.04	19.62	66%	10246	13.79	20.72	67%	9437
2	11.39	23.85	48%	6275	12.31	25.51	48%	5662	12.83	26.60	48%	5260
4	10.04	29.76	34%	3340	10.96	31.76	34%	2973	11.38	33.03	34%	2747
8	9.43	36.96	26%	1572	9.72	38.75	25%	1411	10.74	40.53	26%	1280
16	9.19	44.34	21%	679	9.65	45.38	21%	634	9.78	47.70	20%	582
32	9.11	53.83	17%	259	10.02	53.54	19%	240	9.09	52.71	17%	233
64	9.00	61.83	15%	92	10.45	60.56	17%	89	9.55	59.91	16%	82
128	9.14	71.09	13%	22	7.55	78.18	10%	22	10.13	60.26	17%	23

* *rank* indicates the product rank for the firm in the specific reference market.
The analysis is restricted to firms-products shipping to the reference market and at least one other destination.

Moreover, column 1 of Table 4 suggests a positive correlation between a product's rank in the main export market and the number of destinations served. Although the relationship is not monotone, we do observe lower ranking products being shipped to fewer markets. To look more precisely to this issue, we compute the rank correlation between a product's sales in a reference market and the number of destinations covered. Our reference markets are again France's three main export partners (assumed to be the relevant reference market for each single product) as before, to which we add the main export destination for each firm. Results (see Table 5) show a positive and significant correlation ranging between 0.34 and 0.42, which is robust to using all firms rather than focusing on enterprises exporting more than one product and/or serving more than one foreign market.

Table 5: Correlation between export sales in top destination and number of destinations served.

	destination			
	GER	SPA	GBR	Top by firm
2-5 All firms	0.339	0.358	0.384	0.407
	(65105)	(56971)	(50752)	(294172)
Firms serving more than 1 dest.	0.351	0.373	0.389	0.423
	(64000)	(56269)	(50518)	(276322)
Firms exporting more than 1 prod.	0.347	0.369	0.391	0.409
	(64141)	(56121)	(50172)	(285986)
Firms exporting more than 1 prod. to more than 1 dest.	0.354	0.378	0.393	0.420
	(63307)	(55652)	(50040)	(274435)

Spearman's rank correlation coefficient; (number of observations in parentheses).

How does the theoretical framework presented above relate to the behavior of multi-product firms illustrated here? A simple extension to the model, namely the inclusion of country-specific fixed export costs, allows us to generate the same hierarchy among products exported to different markets that one encounters in the data (**FACT 4**). The extension of the model is presented in the next sub-Section.

5.2 Model extension

Suppose there are heterogeneous fixed entry cost into foreign markets (f_c, with $c = 1, \ldots, C-1$ indexing destinations) and that we can rank foreign countries on the basis of these fixed entry costs such that $f_1 < f_2 < \cdots < f_{C-1}$. This leads us to derive some further implications which are consistent with recent empirical evidence on the behavior of multi-product firms. In particular, there will be a hierarchy among products whereby top-selling varieties are more likely to be shipped to many destinations (see Arkolakis and Muendler, 2010, for evidence along these lines). 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.

As productivity is the same across varieties, whether a product is exported at least in the destination with the lowest entry cost f_1 only depends on its preference attribute a. Let us define a^* as the *preference* cut-off level which makes profit from selling variety j in the foreign market with the lowest entry cost f_1 equal to zero: $a_t^* = sup\{a \, : \pi_t^{F(f_1)}(a) = 0\}$ where $F(f_1)$ is the foreign market with the lowest entry cost f_1. As before, profit coming from selling variety j in the domestic market are

$$\pi_{t,j}^D = a_j \left(\frac{\sigma w_t}{(\sigma-1)P_t^D} \right)^{1-\sigma} \frac{Y_t^D}{\sigma} z_t^{\sigma-1} \tag{35}$$

whereas profit from selling variety j into a foreign market with entry cost f_c are

$$\pi_{t,j}^{F_c} = a_j \left(\frac{\sigma w_t \tau}{(\sigma-1)P_t^{F_c}} \right)^{1-\sigma} \frac{Y_t^{F_c}}{\sigma} z_t^{\sigma-1} - f_c \tag{36}$$

As before, profit from a non-exported variety are simply given by (35) whereas total profit from an exported variety become $\pi_{t,j} = \pi_{t,j}^D + \sum_{c=1}^{C-1} \pi_{t,j}^{F_c}$. Due to asymmetric entry costs, the number of varieties available to consumers will now differ across countries, which implies $Y_t^D \neq Y_t^{F_c}$ and $P_t^D \neq P_t^{F_c}$. As a consequence, variety j will be associated with different revenues and profits in each foreign destination. However, as all firms in a given country will face the same fixed entry costs in foreign markets, the average value of a variety will still be approximately the same across firms. Thus, firms will still choose the same innovation rate λ and maintenance rate μ. The Bellman equation becomes

$$r_t v_t(\tilde{a}) = \max_{\lambda,\mu} \left[w_t \left(\frac{l_t}{\sigma-1} + \sum_{c=1}^{C-1} \beta_c \left(\frac{l_t^c}{\sigma-1} \tau^{1-\sigma} - f_c \right) - (i+h) \right) + v_t(\tilde{a})(\lambda - \mu) + \dot{v}_t(\tilde{z}) \right] \tag{37}$$

where $\beta_c \in [0,1]$ is the share of varieties sold in country c with entry cost f_c.[17] Finally,

[17] An extreme case is when entry costs associated with export are equal across destinations $f_c = f$. Under this assumption total average profit of a firm become $[\frac{l_t}{\sigma-1} + \beta(C-1)\left(\frac{l_t}{\sigma-1}\tau^{1-\sigma} - f\right)]$.

equation (24) becomes

$$\frac{v}{w} = \frac{\frac{l_t}{\sigma-1} + \sum_{c=1}^{C-1} \beta_c \left(\frac{l_t^c}{\sigma-1} \tau^{1-\sigma} - f_c \right) - [i+h]}{r - k - [\lambda - \mu]}. \tag{38}$$

Equations (32) and (33) still describe the evolution of the fraction of commodities produced by a firm of size n. However, the distribution of the number of products of a firm of size n does not coincide with the distribution of its trade links (i.e. its connectivity distribution) as in presence of entry costs in foreign markets only the “best" varieties (those featuring higher levels of consumers' preference) will be exported. However, since the number of products sold by each firm and the preference attribute which determines whether a 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.[18] 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 country-specific fixed entry costs does not entail any significant change to the implications of the model concerning the distribution of the number of products exported by firms. However, the model with country specific entry cost predicts that firms reach foreign markets with their "best" varieties first. This is in line with recent evidence on multi-product firms provided among the others by Arkolakis and Muendler (2010).[19]

This extension of the model, which allows for country-specific fixed entry costs, generates a strict hierarchy among the products exported by each firm. In particular, varieties characterized by strong preferences by consumers (i.e. by high a_j) will be shipped to more countries. Furthermore, since we assume the product-specific preference attribute to be common across destinations, we explain why top-selling varieties tend to be common across destinations.

6 Conclusion

We develop a dynamic model of innovation by new and incumbent firms where firms must invest and innovate 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. Despite the assumption of firm homogeneity in productivity, the model yields a power-law distribution for the number of products exported by each firm, which provides a good approximation of the data. Hence, the

[18]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.

[19]On sequential exporting see also Albornoz et al (2012).

model provides a novel explanation for the large heterogeneity in the extensive margin of trade that characterizes exporting firms, one that does not assume firm heterogeneity to start with but rather derive it as a result.

Second, introducing asymmetric fixed entry costs into foreign markets into the model, we can explain the behavior of multi-product firms: in particular, the stability of product ranking across destinations and the fact that best-selling products are shipped to more destinations.

The paper offers multiple contributions to the existing literature: first, we extend the model by Luttmer (2011) to the open economy and show that this setup can go a long way explaining the dynamics of the extensive margin of trade. Second, we highlight the interrelation between a series of empirical regularities that are commonly thought as independent, namely the number of products exported by each firm, the size distribution of business enterprises, the concentration of export flows at the firm level, and the number of foreign markets covered by each product.

The model can also be used as a benchmark against which to evaluate the performance of firms, helping to investigate the determinants of better- or worse-than-expected performances in foreign markets. We claim this is a promising avenue for further research with potentially interesting policy implications.

References

Albornoz F, Calvo Pardo HF, Corcos G, Ornelas E (2012) Sequential exporting. Journal of International Economics 88(1):17–31

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

Axtell RL (2001) Zipf distribution of U.S. firm sizes. Science 293(5536):1818–1820

Bee M (2006) Estimating the parameters in the loss distribution approach: how can we deal with truncated data? In: Davis E (ed) The advanced measurement approach to operational risk, Riskbooks, London, pp 123–144

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 (2014) Where Gibrat meets Zipf: Scale and Scope of French Firms. DEM Discussion Papers 2014/03, Department of Economics and Management, University of Trento

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 B, Redding S, Schott P (2007) Firms in international trade. Journal of Economic Perspectives 21(3):105–130

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

Bernard AB, Van Beveren I, Vandenbussche H (2014) Multi-product exporters and the margins of trade. The Japanese Economic Review 65(2):142–157

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 (2014) The network structure of international trade. American Economic Review 104(11):3600–3634

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

Cooke RM, Nieboer D, Misiewicz J (2014) Fat-Tailed Distributions: Data, Diagnostics and Dependence. Wiley-ISTE

Di Comite F, Thisse JF, Vandenbussche H (2014) Verti-zontal differentiation in export markets. Journal of International Economics 93(1):50–66

Dixit AK, Stiglitz JE (1977) Monopolistic Competition and Optimum Product Diversity. American Economic Review 67(3):297–308

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

Eaton J, Kortum S, Kramarz F (2011) An anatomy of international trade: Evidence from french firms. Econometrica 79(5):1453–1498

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

Head K, Mayer T, Thoenig M (2014) Welfare and trade without pareto. American Economic Review 104(5):310–316

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

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

Mayer T, Melitz M, Ottaviano G (2014) Market size, competition, and the product mix of exporters. American Economic Review 104(2):495–536

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

Rossi-Hansberg E, Wright ML (2007) Establishment size dynamics in the aggregate economy. American Economic Review 97(5):1639–1666

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

Syverson C (2004) Product Substitutability and Productivity Dispersion. The Review of Economics and Statistics 86(2):534–550

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

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

The first order conditions are

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

$$A_t: \; \frac{\dot{v_t}}{v_t} = \rho - r_t \tag{A-2}$$

Taking the time derivative of (A-1), we get

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

using (A-2), we get the standard Euler equation

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

A.2 Second step: static choice across varieties

The representative household chooses the optimal bundle of varieties to consume (X_t) given its budget constraint

$$\max_{x_{j,t}} \quad X_t = \left[\sum_{j=1}^{N_t} a_j^{\frac{1}{\sigma}} \left(x_{j,t} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

$$\text{subject to} \quad \sum_{j=1}^{N_t} x_{j,t} p_{j,t} = Y_t.$$

The first order condition (FOC) for any variety j is given by:

$$\frac{\sigma}{\sigma-1} \left[\sum_{j=1}^{N_t} a_j^{\frac{1}{\sigma}} \left(x_{j,t} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} a_j^{\frac{1}{\sigma}} \left(\frac{\sigma-1}{\sigma} \right) x_j^{-\frac{1}{\sigma}} = \iota p_j$$

where ι is the relevant Lagrangian multiplier. Taking the ratio between the FOCs for any two varieties i,j we obtain

$$x_i = \frac{Y_t a_i p_{i,t}^{-\sigma}}{\sum_{j=1}^{N_t} a_j p_{j,t}^{1-\sigma}}.$$

The price index P_t can be defined as the level of income necessary to buy 1 unit of the bundle X_t. Setting $X_t = 1$ and solving for the associated expenditure level one gets

$$P_t = \left[\sum_{j=1}^{N_t} a_j p_{j,t}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}.$$

Hence, demand for variety j can be written as

$$x_{j,t} = a_j \left(\frac{p_{j,t}}{P_t} \right)^{-\sigma} \frac{Y_t}{P_t}.$$

B Firm's optimal pricing rule

Let $p_t^D$and x_t^D be price and quantity of variety j for the domestic market and p_t^F,x_t^F prices and quantities of variety j for any of the $(C-1)$ foreign markets (we omit j to simplify notation).

Firms maximize the following profit function to find the optimal price of variety j in the domestic and in the foreign market:

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

subject to

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

$$\sum_{F=1}^{C-1} x_t^F = a\left(\frac{p_t^F}{P_t^F}\right)^{-\sigma}\frac{Y_t^F}{P_t^F}.$$

The resulting first order conditions are

$$p_t^D = \frac{\sigma}{(\sigma-1)}\frac{w_t}{z_t}$$

$$p_t^F = \frac{\sigma}{(\sigma-1)}\frac{w_t}{z_t}\tau = p_t^D\tau.$$

C Profit per variety: the Bellman equation

Abstracting from the time index t for the sake of simplicity, profits for each variety j are given by $\pi_j = p_j x_j - \tau l_j w$; using $x_j = z_j l_j$ and $p_j = \frac{\sigma}{(\sigma-1)}\frac{w}{z_j}\tau$, we can rewrite profits as

$$\pi_j = x_j(p_j - \tau l_j) = x_j\left(\frac{\sigma}{(\sigma-1)}\frac{\tau w}{z} - \tau\frac{w}{z_j}\right) = \frac{x_j w\tau}{z_j(\sigma-1)} = \frac{w\tau}{(\sigma-1)}l_j.$$

D Heterogeneous productivity levels

In this section, we abandon the assumption that varieties are hit by exogenous preference shock and assume that each variety is produced with a different level of productivity from a time-invariant distribution common across varieties and firms. Everything else being equal to our baseline model, the main mechanisms are still at work and the main predictions still valid. As before, the intertemporal utility of the representative consumer is

$$U_t = E_t\left[\int_t^{\infty} ln(X_t)e^{-\rho t}dt\right] \tag{A-3}$$

where $X_t = [\sum_{j=1}^{N_t} x_{j,t}^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}$. Varieties are distinguished only by their productivity levels indexed by $z > 0$ and are produced according to the following production technology

$$x_{j,t} = z_j l_t \tag{A-4}$$

Productivity levels z come from a distribution $\Gamma(z)$ which is continuous in the domain

$0 < z < \infty$ with mean $\bar{z}$. The resulting demand for a variety j is now given by

$$x_{j,t} = \left(\frac{p_{j,t}}{P_t}\right)^{-\sigma} \frac{Y_t}{P_t} \tag{A-5}$$

where $P_t = [\sum_{j=1}^{N_t} (p_{j,t})^{1-\sigma}]^{\frac{1}{1-\sigma}}$. is the price index. As before, the static profit maximization problem for a variety yields prices

$$p^D_{j,t} = \frac{\sigma}{\sigma-1} \frac{w_t \tau}{z_j} \tag{A-6}$$

$$p^F_{j,t} = \frac{\sigma}{\sigma-1} \frac{w_t \tau}{z_j} \tag{A-7}$$

for the domestic market a foreign market respectively. Revenues from sales of variety j in the domestic market and in a foreign market are

$$r^D_{j,t} = \left(\frac{p^D_{j,t}}{P_t}\right)^{1-\sigma} Y_t \tag{A-8}$$

$$r^F_{j,t} = \left(\frac{p^F_{j,t}}{P_t}\right)^{1-\sigma} Y_t \tag{A-9}$$

Total revenues from variety j are $r_{j,t} = r^D_{j,t}(1 + (C-1)\tau^{1-\sigma})$. Profits from variety j in the domestic market and in a foreign market are

$$\pi^D_{j,t} = \left(\frac{\sigma w_t}{(\sigma-1)P_t}\right)^{1-\sigma} \frac{Y_t}{\sigma} z_j^{\sigma-1} \tag{A-10}$$

$$\pi^F_{j,t} = \left(\frac{\sigma w_t}{(\sigma-1)P_t}\right)^{1-\sigma} \frac{Y_t}{\sigma} z_j^{\sigma-1} \tau \tag{A-11}$$

As $\pi^F_{j,t} = \pi^D_{j,t}\tau^{1-\sigma}$, total profits from varity j is $\pi_{j,t} = \pi^D_{j,t}[1 + (C-1)\tau^{1-\sigma}]$. At each point in time, for a firm with $n_{i,t}$ products aggregate revenues and profits are

$$r^{Agg}_{j,t} = \sum_{j=1}^{n_t} r^D_{j,t}(1 + (C-1)\tau^{1-\sigma}) \tag{A-12}$$

$$\pi^{Agg}_{j,t} = \sum_{j=1}^{n_t} \pi^D_{j,t}(1 + (C-1)\tau^{1-\sigma}) \tag{A-13}$$

As before, we can express average revenues and profits of a firm with $n_{i,t}$ products as a function of a summary measure that is its average productivity $\tilde{z} = \sum_{j=1}^{n_t} \frac{1}{n_t} z_j$

$$r_t(\tilde{z}, n_{i,t}) = \sum_{j=1}^{n_t} \frac{1}{n_t} r^D_{j,t}(1 + (C-1)\tau^{1-\sigma}) \tag{A-14}$$

$$\pi_t(\tilde{z}, n_{i,t}) = \sum_{j=1}^{n_t} \frac{1}{n_t} \pi^D_{j,t}(1 + (C-1)\tau^{1-\sigma}) \tag{A-15}$$

As the portfolio of goods $n_{i,t}$ that a firm produces increases, by the law of large number we expect $\tilde{z} \to \bar{z}$, so that $r_{i,t}(\tilde{z}) \simeq r_{i,t}(\bar{z})$ and $\pi_{i,t}(\tilde{z}) \simeq \pi_{i,t}(\bar{z})$, where $\bar{z}$ is the average productivity of the distribution $\Gamma(z)$. Thus, for a large enough n, the average revenues and profits per variety will be approximately the same for all firms. As in the model with exogenous preference shocks, this assumption guarantees the existence of a balance growth path and allows us to derive a unique distribution of the number of products produced (and exported) by firms (i.e. connectivity distribution) as shown in Sections 3.3–3.5. Moreover, as productivity is now constant over time, in the balanced growth wages w_t and per capita consumption c_t grow at rate $k = \eta/(\sigma - 1)$.

E Solution of the system (32)–(33)

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 (32)–(33) is given by (the proof of the solution to the system is provided by Luttmer, 2011 in Appendix A of his paper)

$$Q_n = \frac{\nu}{\lambda} \sum_{k=0}^{\infty} \frac{1}{\beta_{n+k}} \left(\prod_{m=n}^{n+k} \beta_m \right) \prod_{m=n}^{n+k} \frac{\lambda \beta_m}{\mu} \tag{A-16}$$

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

$$Q_n \sim \frac{\nu}{\mid \lambda - \mu \mid} \prod_{m=1}^{n-1} \frac{\lambda \beta_m}{\mu}. \tag{A-17}$$

If $\nu = 0$, the only non-negative and summable solution to equations (32)–(33) is identically zero, implying that there does not exist a stationary distribution in this case. If $\nu > 0$, equation (A-16) 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 (A-17) 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

$$\lim_{n \to \infty} n \left(1 - \frac{R_{n+1}}{R_n}\right) = \xi \tag{A-18}$$

where $\xi = \frac{\eta}{(\lambda - \mu)}$. That is, R_n is a regularly varying sequence with index $-\xi$ and $\xi > 1$.Finally, when the rate $\nu = \eta - (\lambda - \mu)$ goes to zero, the limiting tail index $\xi = 1$ associated with Zipf's law arises.