An Anatomy of International Trade:

Evidence from French Firms

Jonathan Eaton,*Samuel Kortum,†and Francis Kramarz‡

October 2005

Abstract

We develop an equilibrium model of worldwide competition across a range of goods. Our model encompasses Ricardian and monopolistic competition as special cases. We parameterize the model to gauge its ability to capture the export behavior of French manufacturing firms.

Key words: International trade, exporting, market penetration

*Department of Economics, New York University, and National Bureau of Economic Research (jonathan.eaton@nyu.edu).

†Department of Economics, University of Minnesota, Federal Reserve Bank of Minneapolis, and National Bureau of Economic Research (kortum@econ.umn.edu).

‡CREST-INSEE and CEPR (kramarz@ensae.fr).

1   Introduction

A literature that emerged over the last 10 years has exploited microlevel data to measure various features of the export behavior of individual producers. Bernard and Jensen (1995, 1999), Roberts and Tybout (1997), Clerides, Lach, and Tybout (1998, henceforth CLT), Aw, Chung, and Roberts (1998), and Hallward-Driemeier, Iarossi, and Sokoloff (2002) using data from various countries, have shown that producers who export are typically in the minority and tend to be more productive and much larger, even in terms of their domestic sales; yet they usually export only a small fraction of their output.

All of these characteristics suggest that individual producers face substantial hurdles in entering foreign markets. Several theories have emerged in response to these observations. Bernard, Eaton, Jensen, and Kortum (2003) (Henceforth BEJK) develop a Ricardian model of plant-level export behavior while Melitz (2003) and Chaney (2005) provide models based on monopolistic competition. Essential to either explanation are trade barriers that deter many producers who sell at home from entering foreign markets.

Until recently, our ability to gauge a producer's export activity has been limited to observations on how much it exported. We have been in the dark about how exports broke down into sales in individual destinations. Yet this information is critical in understanding the nature of the barriers that individual producers face in selling abroad. In particular, the existing evidence raises three questions that our analysis here seeks to answer: (1) Is the major hurdle to exporting selling beyond the home market, with broad penetration across foreign markets among producers that do export, (as implicit in Roberts and Tybout, CLT, and Melitz), or do exporters appear to incur such costs market by market (as in BEJK and Chaney)? (2)

In either case, does the cost appear to be fixed (as in CLT and Melitz), or increasing in the amount shipped, as with standard "iceberg" transport costs? (3) What market structure does the evidence favor?

Our work makes use of an extensive source of data that provides some insight into the answers to these questions. The starting point is a comprehensive data set of French firms that has been merged with customs data on the value of each firm's shipments to individual destinations (see, Biscourp and Kramarz, 2002). Focusing on manufacturing firms in 1986, these data reveal enormous heterogeneity across both destinations and across firms in the nature of entry into different markets. Nevertheless, we observe some striking regularities. Looking across firms, the size and productivity advantages of exporters extend very seamlessly into size and productivity advantages of firms that export widely, and to less popular destinations. Looking across destinations, the number of French sellers to a destination increases with overall French market share with an elasticity close to one, while the number increases with market size with an elasticity of around two-thirds (see, Eaton, Kortum, and Kramarz, 2004).

We develop a model of firm competition across a wide number of markets that incorporates a fixed cost of entering an individual market as well as the standard iceberg costs that rise in proportion to the amount shipped. Both are needed to explain the increase in the number of sellers with market size and the dominance of home sales even among exporters. Our model nests the Ricardian framework of BEJK and the monopolistic competition (MC) approach of Melitz and Chaney by introducing the range of possible goods as a parameter of the model. When the range is small relative to the number of active producers, the model implies that

multiple producers are competing head to head in different markets of the world, as in BEJK. A large range, however, implies that it is very unlikely that a producer faces a direct competitor anywhere (another firm able to cover the fixed cost of entry there), so monopolistic competition prevails.

We estimate the parameters of the model using data on aggregate production and bilateral trade among France and 112 trade partners as well as moments from our firm level dataset. Our estimation proceeds in two stages.

We first show that, under a simple deterministic formulation of the model with monopolistic competition, several parameter values can be calibrated very directly by observing (1) the relationship between the number of producers selling to at least some given number of destinations and the average sales in France of these producers and (2) average sales per market and the size of the market. The parameterized model fits these relationships tightly and provides ballpark estimates of several parameters close to values delivered by a more sophisticated model. It fails in two dimensions, however. First, it implies a much less heterogeneous sales distribution in any individual market than the data exhibit. Second, it predicts a strict hierarchy of export destinations: That is, a firm selling to the k'th most popular export destination is predicted to sell to the first through $k - 1$'st as well. There are substantial deviations from such a hierarchy in the data.

We then estimate the parameters of a richer model that incorporates both an endogenous range of goods (so that monopolistic competition is only a special case) and firm and market specific shocks to demand and to the fixed cost of entry. We estimate the model by simulated method of moments. That is, given a vector of parameters, we use the model to simulate a

4

population of French firms and their export behavior around the world. We then search over parameter values to make the moments of our simulated dataset match key moments of the actual data.

The outline of our paper is as follows. In the next section we describe our data in detail. Section 3 then presents our theoretical framework. In Section 4 we show how a simple version of the model lines up with some systematic features of the data, and delivers some estimates of the parameters. Section 5 describes our simulation approach and the results of our estimation of the parameters by simulated method of moments.

2   The Data

Our analysis uses both aggregate and firm level data. At the firm level we analyze the sales of 234,300 French manufacturing firms in 113 destinations around the world (including France itself). At the aggregate level we use data on bilateral trade flows in manufactures among these 113 countries, including home sales.

Our firm level data are constructed as follows: We merge data from two French administrative sources. The first is a collection of records of the universe of firms subject to the standard tax system. After additions and controls made at INSEE, the data include all balance-sheet variables, employment, industry affiliation, total sales, and a firm identifier (the Siren identifier). Second, French Customs compile all sales of French firms (also indicating their Siren identifier) in over 200 foreign destinations. Biscourp and Kramarz (2002) provide a thorough description of the two sources.

While the data cover all private sector firms, our focus is on a cross section of manufac-

turing firms from 1986, yielding a sample of 234,300. Since we lack other data (in particular, on domestic production) from many of the smaller destinations, we limit ourselves to 113 destination countries (including France). (Since the entities which we eliminated are very small, they constitute a trivial proportion of France's total export activity.)

As is typically the case, summing across what individual producers report exporting produces a number that is less than what is reported at the aggregate level. In the French case missing exports arise because manufacturing firms sell to nonmanufacturing intermediaries who report the foreign sales, and the connection between producer and destination is lost. Across all destinations, the firm data fail to account for about 20 per cent of total manufacturing exports.[1]

While the raw data themselves are confidential and housed at INSEE, we can construct a rich set of statistics from them. Some of these statistics do not rely on individual export destinations, so can be compared with the analogous statistics from producers located elsewhere (and, in particular, to U.S. producers). Statistics based on individual destinations, however, are to our knowledge unique to these data, providing a new window on the connections between firms and where they sell.

Previous work on the export behavior of individual producers has typically used the plant as the unit of observation. The French data report exports and other features by firm. Obviously differences arise. A firm might own several plants, for example, while a firm might exist that does not own any production unit that corresponds with the definition of a plant. *A priori*, a case can be made for either unit of observation over the other. A firm, for exam-

---

[1]This figure compares with underreporting of about 40 per cent in the U.S. Census of Manufactures. See Bernard and Jensen (1999) for a discussion.

ple, might own several plants with very diverse characteristics. Hence firm-level observations might mask a great deal of the variation in the plant-level data. But observations at the plant level may fail to pick up inputs provided by the headquarters.

3 Theory

Our theory is about competition across N geographically separated destinations in selling a good j, where there are a continuum of possible goods with measure J. In our quantitative analysis, of course, $N = 113$, while the range of potential goods J is a parameter that we estimate.

3.1 Technology

Our model of technology is adapted from Eaton and Kortum (2002) and BEJK (2003). The most efficient potential producer of good j in country i can produce and amount $z_i(j)$ per unit of input, where $z_i(j)$ is the realization of a random variable Z_i drawn from the distribution:

$$F_i(z) = \Pr[Z_i \leq z] = \exp[-(T_i/J)z^{-\theta}],$$

where $T_i > 0$ and $\theta > 1$ are parameters. The parameter θ, which we treat as common across countries and goods, governs the extent of heterogeneity in efficiency, with larger values of θ implying less heterogeneity. The parameter T_i governs the average level of efficiency in country i. It may appear that the measure of goods J should be simply subsumed into T_i, since the ratio is all that matters for the probability distribution of Z_i. A distinct role for J emerges, however, when we look across all the goods in the economy. The measure of goods that can be produced in country i with efficiency greater than z is $J\{1 - \exp[-(T_i/J)z^{-\theta}]\}$.

The cost of an input unit in i is w_i while it requires shipping $d_{ni} \geq 1$ units of a good from i to deliver one unit in n. We normalize $d_{ii} = 1$ for all i. The unit cost of delivering a unit of good j in n from i is thus

$$c_{ni}(j) = \frac{w_i d_{ni}}{z_i(j)}.$$

The lowest cost version of good j in market n costs:

$$c_n(j) = \min\{c_{n1}(j), c_{n2}(j), ..., c_{nN}(j)\}.$$

3.2 The Distribution of Costs

Our distributional assumptions about Z imply that $c_{ni}(j)$ is the realization of a random variable C_{ni} drawn from the distribution:

$$\Pr[C_{ni} \leq c] = 1 - \exp[-(T_i/J)(w_i d_{ni})^{-\theta} c^{\theta}].$$

while $c_n(j)$ is the realization of a random variable C_i drawn from the distribution:

$$\Pr[C_n \leq c] = 1 - \exp[-(\Phi_n/J)c^{\theta}] \tag{1}$$

where:

$$\Phi_n = \sum_{i=1}^{N} T_i(w_i d_{ni})^{-\theta}.$$

The measure of goods that are potentially supplied to country n at a cost less than c is:

$$\mu_n(c) = J\{1 - \exp[-(\Phi_n/J)c^{\theta}]\}. \tag{2}$$

The probability that country i is the low cost supplier of good j to n is:

$$\pi_{ni} = \frac{T_i(w_i d_{ni})^{-\theta}}{\Phi_n}. \tag{3}$$

In addition to the unit cost, we introduce a fixed cost $E_n(j)$ to a firm of selling good j in market n. We assume that this fixed cost can be decomposed into a country component that applies to all goods and a component that varies across goods j. Hence we can write:

$$E_n(j) = E_n \varepsilon_n(j),$$

where $\varepsilon_n(j)$ is the realization of a random variable ε, which we treat as independent of any producer's efficiency $Z_i(j)$.

3.3 Demand and Market Structure

Our assumptions about demand are very standard. Expenditure on good j in market n, available at price $p_n(j)$ there, is:

$$X_n(j) = \alpha_n(j) \left(\frac{p_n(j)}{P_n} \right)^{1-\sigma} X_n \qquad (4)$$

where X_n is total spending and P_n the CES price index:

$$P_n = \left[\int_0^J \alpha_n(j) p_n(j)^{1-\sigma} dj \right]^{1/(1-\sigma)}.$$

where σ is the elasticity of substitution. We restrict $\sigma \in (1, \theta + 1)$.[2]

The term $\alpha_n(j)$ is the realization of a random variable α that is also independent of any producer's efficiency $Z_i(j)$, but may be correlated with ε. We treat goods that are not sold in country n as having an infinite price.

Our assumptions are compatible with a broad range of market structures, at least over certain ranges of parameters. For concreteness we assume here that at most only the lowest

---

[2]We need $\sigma > 1$ both to get a well defined price for the case in which a producer faces no direct competition and to explain why the range of goods provided in different locations can differ. The restriction that $\sigma < \theta + 1$ is needed to ensure that the price index, derived below, is well defined.

unit cost supplier of good j to market n enters, with the fixed cost of entry deterring entry by others. Hence this supplier, conditional on entry, has a monopoly in market n, so charges a markup over unit cost of:

$$\overline{m} = \frac{\sigma}{\sigma - 1}.$$

Hence its unit price is:

$$p_n(j) = \overline{m} c_n(j)$$

if good j is sold in country n at all.

To summarize our assumptions, a potential producer has three characteristics: (1) the country i of its location, (2) the good j it knows how to make, and (3) its efficiency $z_i(j)$ making good j at location i. In turn, each good $j \in [0, J]$ has $2N$ characteristics: (1) the good-specific component of the entry barrier in each location $\varepsilon_n(j)$ and (2) the good-specific shock to demand in each location $\alpha_n(j)$. Each location i is distinguished by (1) its measure of ideas T_i, (2) its input cost w_i, (3) the component of the entry barrier that is common across goods E_i, and its geography relative to other locations reflected in the geographic barriers d_{ni}. The remaining parameters of the model are the range of goods J and the parameter θ, which governs heterogeneity in technology, and σ, which governs heterogeneity in preferences. We now turn to the determination of equilibrium.

3.4 Equilibrium Entry

The lowest cost supplier, conditional on entry, earns a profit, gross of the fixed cost, of $X_n(j)/\sigma$. Conditional on being the low cost supplier of good j in market n, a producer enters

that market if:

$$X_n(j) \geq \sigma E_n \varepsilon_n(j)$$

or if:

$$\eta_n(j)x_n \geq \left(\frac{\overline{m}c_n(j)}{P_n}\right)^{\sigma-1} \tag{5}$$

where:

$$\eta_n(j) = \frac{\alpha_n(j)}{\varepsilon_n(j)}$$

and

$$x_n = \frac{X_n}{\sigma E_n}.$$

Note that η is a positive shock to entry. For any $\eta_n(j) = \eta$, condition (5) for entry determines

a cutoff cost $\overline{c}_n(\eta)$ such that only a supplier with $c_n(j) \leq \overline{c}_n(\eta)$ would enter, where

$$\overline{c}_n(\eta) = \eta^{1/(\sigma-1)}\overline{c}_n \tag{6}$$

and:

$$\overline{c}_n = x_n^{1/(\sigma-1)}\frac{P_n}{\overline{m}}.$$

Integrating across the range of costs in any location n, the price index is:

$$
\begin{aligned}
P_n &= \overline{m}\left\{E_\eta\left[\int_0^{\overline{c}_n(\eta)} E[\alpha|\eta]c^{1-\sigma}d\mu_n(c)\right]\right\}^{1/(1-\sigma)} \tag{7}\\
&= \overline{m}\left\{E_\eta\left[E[\alpha|\eta]\int_0^{\eta^{1/(\sigma-1)}\overline{c}_n} c^{\theta-\sigma}\exp[-(\Phi_n/J)c^\theta]\theta\Phi_n dc\right]\right\}^{1/(1-\sigma)}
\end{aligned}
$$

(since we treat α and η as independent of z, and hence c).

Equation (6) defines a positive relationship between $\overline{c}_n$ and P_n while equation (7) defines

a negative one. Together they determine $\overline{c}_n$ and P_n.

To solve for each we define the variable:

$$\widetilde{P}_n = \left(\frac{P_n}{\overline{m}}\right)^{1-\sigma} \Phi_n^{-1/\widetilde{\theta}} \tag{8}$$

where

$$\widetilde{\theta} = \frac{\theta}{\sigma - 1}.$$

The $\widetilde{P}_n$ solves:

$$\widetilde{P}_n = E_\eta \left[E[\alpha|\eta] J^{1-1/\widetilde{\theta}} \Gamma \left(1 - 1/\widetilde{\theta}, \frac{\left(\eta x_n/\widetilde{P}_n\right)^{\widetilde{\theta}}}{J} \right) \right], \tag{9}$$

where $\Gamma(a,x) = \int_0^x t^{a-1} e^{-t} dt$ is the incomplete gamma function.[3] Using this new term, the price index can then be written:

$$P_n = \overline{m} \left[\widetilde{P}(x_n) \right]^{-1/(\sigma-1)} \Phi_n^{-1/\theta}$$

while the number of entrants is:

$$J_n = J \left(1 - E_\eta \left[\exp \left\{ - \left(\eta x_n/\widetilde{P}(x_n) \right)^{\widetilde{\theta}} /J \right\} \right] \right). \tag{10}$$

Suppliers to market n have heterogeneous costs, and, from (3) above, a supplier from country i is more likely to sell in country n the larger π_{ni}. But, conditional on entry, suppliers from all countries have the same cost distribution in n and, given the constant markup, have the same distribution of prices and, hence, of sales. An implication is that the probability π_{ni} that a firm from i is the supplier of some particular good j, is also the fraction of spending by

---

[3]Combining (6) and (7), applying the change of variable $s = \Phi_n c^\theta$, and rearranging gives:

$$\widetilde{P}_n = E_\eta \left[E[\alpha|\eta] \int_0^{\left(\eta x_n/\widetilde{P}_n\right)^{\widetilde{\theta}}} s^{-1/\widetilde{\theta}} \exp(-s/J) ds \right].$$

Applying the change of variable $t = s/J$ we obtain (9).

country n on goods from country i. We can thus relate π_{ni} to data on import shares, that is:

$$\pi_{ni} = \frac{X_{ni}}{X_n} \qquad (11)$$

where X_{ni} is n's purchases from i.

Since, conditional on entry, suppliers have the same cost distribution in market n regardless of their origin, the measure of firms from source i in market n, J_{ni}, should equal a fraction π_{ni} of the total number, so that:

$$J_n = \frac{J_{ni}}{\pi_{ni}}.$$

We use this relationship to infer the total number of sellers to a market from the number of French firms selling there and French market share.

3.5 Two Special Cases

Before turning to the general solution we consider two special cases close to those in the existing literature.

3.5.1 Pure Ricardian Competition

Say that $E_n = 0$ and $J = 1$ as in EK (2002). Since there is no entry barrier, the cutoff is infinite while the price index is:

$$P_n = \overline{m}\,[E(\alpha)]^{-1/(\sigma-1)}\,[\Gamma(1 - 1/\widetilde{\theta})]^{-1/(\sigma-1)}\Phi_n^{-1/\theta}$$

which, setting $\overline{m} = 1$ (since they assume perfect competition), reduces to the expression in EK (2002). Only the lowest unit cost supplier of good j to market n sells there.

3.5.2 Monopolistic Competition

Let $J \to \infty$. We then get a price index:

$$P_n = \overline{m} \left(1 - 1/\widetilde{\theta}\right)^{1/\theta} a_1^{-1/\theta} x_n^{-(1-1/\widetilde{\theta})/(\sigma-1)} \Phi_n^{-1/\theta} \tag{12}$$

where:

$$a_1 = \left\{ E_\eta \left[E[\alpha|\eta] \eta^{\widetilde{\theta}-1} \right] \right\}.$$

The cutoff is

$$\overline{c}_n = \left(1 - 1/\widetilde{\theta}\right)^{-1/\theta} \left(\frac{\Phi_n}{a_1 x_n}\right)^{-1/\theta}. \tag{13}$$

Taking the measure (2) as $J \to \infty$, the measure of entrants with cost less than or equal to c is Pareto with parameters Φ_n and θ:

$$\lim_{J \to \infty} \mu_n(c) = \Phi_n c^\theta. \tag{14}$$

Taking the limit of (10) as $J \to \infty$ the measure of entrants is:

$$J_n = \left(1 - 1/\widetilde{\theta}\right) x_n E[\eta^{\widetilde{\theta}}]. \tag{15}$$

which rises in proportion to x_n.

If, in addition, we shut down market-specific sales and entry shocks by setting $\alpha_n(j) = \varepsilon_n(j) = 1 \ \forall n, \ j$, our formulation is monopolistic competition with potential sellers having a Pareto distribution of efficiencies, as in Chaney (2005). Firms in any source are identical except for their efficiencies z. An implication is that there is a hierarchy of destinations.

Setting $a_1 = 1$ in (13) above, a necessary and sufficient condition for a firm from i to sell in market n is that it have a domestic cost below:

$$\overline{c}_{ni} = \overline{c}_n / d_{ni}$$

14

For each source i we can rank destinations n according to $\bar{c}_{ni}$, where $\bar{c}_i^{(1)} \geq \bar{c}_i^{(2)} \geq \bar{c}_i^{(3)} \geq ... \geq \bar{c}_i^{(k)} \geq ... \geq \bar{c}_i^{(N)}$. Hence any firm that sells to the k'th ranked market has a domestic cost c below $\bar{c}_i^{(k)}$ which is also below $\bar{c}_i^{(k')}$ for all $k' < k$. Hence it must sell to these markets as well. Hence in this special case each source i should have a hierarchy of destinations, with more efficient firms selling to destinations further down the hierarchy.

This special case yields simple specifications for (1) the distribution of sales in any market, (2) the number of firms entering a market, and (3) the relationship between a firm's sales in any particular market and the number of markets where it sells.

1. **The sales distribution.** To sell in market n a firm must sell at least σE_n to overcome the entry hurdle. The distribution of its sales there is:

$$F_n(x) = 1 - \Pr[X \geq x | X \geq \sigma E_n] = 1 - \Pr\left[C \leq \left(\frac{x}{X_n}\right)^{1/(\sigma-1)} \middle| C \leq \left(\frac{\sigma E_n}{X_n}\right)^{1/(\sigma-1)} \right]$$

which, from (14), is:

$$F_n(x) = 1 - \left(\frac{x}{\sigma E_n}\right)^{-\widetilde{\theta}} \quad x \geq \sigma E_n. \tag{16}$$

while mean sales are:

$$\bar{x}_n = \frac{\sigma E_n}{1 - 1/\widetilde{\theta}}.$$

That is, the sales distribution is Pareto with slope $\widetilde{\theta}$.

2. **Entry.** From (15), the measure of firms selling in market n is simply:

$$J_n = \left(1 - 1/\widetilde{\theta}\right) x_n \tag{17}$$

3. **Sales in a Market and Number of Markets Served.** Consider the sales of a firm from country i selling in market n. Its sales in that market are drawn from the

15

distribution $F_n(x)$ above and its cost in market n must be below $\bar{c}_n$. If the firm sells in markets that are less popular than n, its cost in market n must be lower still, implying higher sales in n. Denote by $J_{ni}^{(k)}$ the measure of firms from i selling in n that also sell in at least k less popular markets than n. This measure is decreasing in k. From (16) above, to sell in at least k less popular markets, sales in n must be at least:

$$\underline{x}_{ni}^{(k)} = \sigma E_n \left(\frac{J_{ni}^{(k)}}{J_{ni}^{(0)}} \right)^{-1/\widetilde{\theta}}$$

while the mean sales in market n of firms from i selling to k less popular destinations than n is:

$$\bar{x}_{ni}^{(k)} = \frac{\sigma E_n}{1 - 1/\widetilde{\theta}} \left(\frac{J_{ni}^{(k)}}{J_{ni}^{(0)}} \right)^{-1/\widetilde{\theta}} \tag{18}$$

The model delivers a precise relationship between a firm's sales in any given market and the number of less popular markets it sells in.

3.6   Application to French Firms

Our particular focus is on the model's implications for observations on French firms and their export activity. Furthermore, we assume that French firms are observed only if they sell in the French market. Our model does not impose this last requirement, so we will interpret our data on French firms as being a truncated sample.

If a French firm producing good j were to enter market n, it would sell:

$$X_n^*(j) = \alpha_n(j) \left(\frac{\overline{m} c_{nF}(j)}{P_n} \right)^{1-\sigma} X_n \tag{19}$$

where $c_{nF} = w_F d_{nF}/z_F(j)$. To enter the market it has to overcome two distinct hurdles. First,

its operating profits need to overcome the cost of entry, meaning that:

$$X_n^*(j) \geq \sigma E_n \varepsilon_n(j), \tag{20}$$

what we call the entry hurdle. Second, it must be the lowest cost supplier of good j to market n, meaning that:

$$c_{nF}(j) < \widetilde{c}_n(j) = \min_{i \neq F}\{c_{ni}(j)\}, \tag{21}$$

what we call the competition hurdle.

It is important to remember that we treat the $\eta_n(j)$ as applying to all potential sellers of good j in market n, regardless of source. Hence if a French firm passes the entry hurdle in destination n so does any other seller with a lower unit cost in that market. Hence the entry hurdle never protects a French firm from a lower unit cost competitor.

4   Quantification I: Calibrating Monopolistic Competition

The special case of the model with monopolistic competition and no shocks to sales or the entry barrier, is particularly simple to calibrate.

A strict implication, as discussed, is a market hierarchy. In fact, every firm in our sample sells in France, so that this first element of the hierarchy is not violated. But 48 percent of the over 35 thousand firms that export in our sample don't sell in Belgium, the most popular foreign destination. Looking at the top seven destinations, 73 percent of exporters violate the hierarchy by skipping more popular destinations. But looking at this figure another way, 27 percent of firms sell to a string of destinations that satisfies the market hierarchy. Such

strings constitute only 7 of 128 ($= 2^7$), or 5.5 percent, of possible configurations of sales to 7 markets.[4]

Moreover, organizing firms according to the least popular market they serve and according to the number of markets they serve gives very similar results. Figure 1 graphs the number of firms selling to k or more markets against the number of firms selling to the k'th most popular market. The relationship suggests a rough one-to-one correspondence. Furthermore, as we discuss below, selling in less popular markets has very similar implications for sales in France than selling in many markets. Hence, while the strict implication of market hierarchies is violated, we think that there are enough features of the data consistent with this implication that this simple version of the model is worth exploring further to see what it has to say about parameters of interest. We go on to examine the three relationships discussed above, in reverse order.

4.1 Sales in France by Exporters to Multiple Destinations

Figure 2a plots average sales in France of French firms selling to k or more markets against the number of firms selling to k or more markets. It is the observational analogue of (18) above. Note that the relationship is tight, and approximately linear on a logarithmic scale, as the theory implies. Moreover, its slope is -2/3, suggesting a value of $\widetilde{\theta}$ of 1.5. Hence monopolistic competition combined with an assumption that efficiencies have a Pareto distribution fits the the relationship between French firms' sales in France and the number of export destinations

---

[4]The most popular foreign destinations, in order, are Belgium, West Germany, Switzerland, Italy, United Kingdom, the Netherlands, and USA. With its prediction of a hierarchy of destinations, the basic model cannot explain how, for example, we could ever observe a French firm selling in Italy but not in Belgium.

that they serve.

Returning to the issue of hierarchies, Figure 2b plots sales in France of French firms selling to the k'th most popular market against the frequency of firms selling there. Note that the relationship is very similar. French firms that sell to unpopular markets sell systematically more in France, just as French firms that sell to many markets sell more in France.

4.2   Entry and the Price Index

Under monopolistic competition, as well as for a wide range of other market structures, the number of French firms selling to a destination, divided by French market share, provides an estimate of the total number of firms selling there. We thus use (17) to infer σE_n across our 113 destinations, using J_{nF}/π_{nF} as a proxy for J_n. That is, we calculate:

$$\sigma E_n = (1 - 1/\widetilde{\theta})\overline{x}_{nF}$$

where $\overline{x}_{nF} = X_{nF}/J_{nF}$ is mean sales of French firms in market n (using our estimate of $\widetilde{\theta} = 1.5$).

Figure 3 plots our estimate of σE_n against total market size X_n (measured as manufacturing absorption, home production plus imports minus exports) on a logarithmic scale. Note that the relationship is linear, upward sloping, and quite tight. A linear regression of $\ln(\sigma E_n)$ against $\ln X_n$ has an R^2 of .71. The slope is .36 with a standard error of .02.

We can use our estimates of σE_n and $\widetilde{\theta}$, along with data on X_n and conjectures about σ, to infer the contribution of x_n to the price index, using (12) (setting $a_1 = 1$) Specifically we calculate the term:

$$\widetilde{P}_n(\sigma) = x_n^{-(1-1/\widetilde{\theta})/(\sigma-1)} = \left[\frac{X_n}{(1-1/\widetilde{\theta})\overline{x}_{nF}}\right]^{-1/[3(\sigma-1)]} = \left[\frac{J_{nF}}{(1-1/\widetilde{\theta})\pi_{nF}}\right]^{-1/[3(\sigma-1)]}$$

19

for each destination n using various values of σ. Figure 4 plots this component of the price index against market size X_n on a logarithmic scale. A regression of $\ln \widetilde{P}_n$ against $\ln X_n$ has an R^2 of .89 and yields a regression coefficient of $-.046$ (standard error .002). The implication is that a doubling of market size is associated with a decline in the price index of $-.046$ percent due to increased entry.[5] Size has a modest but notable effect on welfare through increased variety.[6]

4.3 The Sales Distribution

In the simple case of monopolistic competition, the distribution of sales in any market is given by (16), a Pareto distribution with parameter $\widetilde{\theta}$. From above, the relationship between a firm's sales in France and number of markets its serves implies that $\widetilde{\theta} = 1.5$.

Figure 4 plots the average sales distribution of French firms across destinations (distinguishing among markets according to whether France's total exports there are large, medium, or small)..Two things don't fit. First, the relationship is nonlinear at the lower end of the distribution, violating the implication that the sales distribution should be linear in logarithms. Second, the slope at the upper end is too steep, with a slope closer to -1 than -2/3. The

---

[5]Recalculating $\widetilde{P}_n^e(\sigma)$ for $\sigma = 3$ yields an elasticity of the price index with respect to market size of -11 percent.

[6]Taking $\widetilde{\theta} \to \infty$ delivers the standard formulation of monopolistic competition with homogeneous firms. Note that in this case market size has a greater effect on welfare through variety as the elasticity of $\widetilde{P}^e(\sigma)$ with respect to size is $-1/(\sigma - 1)$. Heterogeneity in technology attenuates the effect of size on welfare as larger markets attract higher-cost firms on the margin. A point made by Ghironi and Melitz (2004) is that, with technological heterogeneity, the average price of a good sold will be higher in a large market, even though the true price index is lower due to greater variety of goods.

sales distribution is more skewed than what is implied by the value of $\widetilde{\theta}$ inferred from the size advantage in France of prolific exporters implies.

We conclude that the model of monopolistic competition does a good job of picking up the relationship between exports in any given market and the number of markets served. It provides hints about the cost of entry (σE_n) and about the ratio of the heterogeneity parameters $(\widetilde{\theta} = \theta/(\sigma - 1))$. But it does not explain aspects of entry (with its prediction of a strict hierarchy of destinations) and it understates the curvature and heterogeneity of sales in any given market.

5 Quantification II: Simulated Method of Moments

We now turn to the estimation of a more general model to assess its ability to grapple with these feature of the data. We generalize the case above by allowing for destination-specific shocks to entry and to sales, and by treating the range of goods J as a parameter to be estimated.

5.1 Stochastic Specification

In all of the quantitative analysis, it is convenient to isolate the stochastic component of $c_{ni}(j)$ by introducing the variable:

$$u_i(j) = (T_i/J)\,(w_i d_{ni})^{-\theta} c_{ni}(j)^{\theta}.$$

Our assumptions imply that $u_i(j)$ is the realization of a random variable U_i drawn from the unit exponential distribution:

$$\Pr[U_i \leq u] = 1 - \exp(-u). \tag{22}$$

This definition allows us to express $c_{ni}(j)$ as:

$$c_{ni}(j) = (w_i d_{ni})(T_i/J)^{-1/\theta} u_i(j)^{1/\theta}.$$

Invoking expression (3) for the trade share, implies:

$$c_{ni}(j) = (\Phi_n/J)^{-1/\theta} \left(u_i(j)/\pi_{ni} \right)^{1/\theta} \tag{23}$$

We can express our competitiveness hurdles (21) in terms of the $u_i(j)$'s and data on trade shares as:

$$u_F(j) < \widetilde{u}_n(j) = \min_{i \neq F}\{\pi_{nF} u_i(j)/\pi_{ni}\}$$

if the firm is competitive in market n. We use $\widetilde{u}(j)$ to denote the vector of competitiveness hurdles across all markets n. As mentioned above, a necessary condition for a French firm to appear in our data is that the firm sells in France. Thus, it must pass the competitiveness hurdle for the French market, i.e. $u_F(j) < \widetilde{u}_F(j)$.

Aside from the $u_i(j)$'s, our model has the stochastic components $\alpha_n(j)$ and $\eta_n(j)$. We assume that these components have the joint bivariate lognormal distribution:

$$\begin{bmatrix} \ln \alpha \\ \ln \eta \end{bmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\alpha^2 & \rho \sigma_a \sigma_h \\ \rho \sigma_a \sigma_h & \sigma_h^2 \end{pmatrix} \right].$$

Under this distributional assumption, the expression (9) for $\widetilde{P}(x_n)$ becomes:

$$\widetilde{P}(x_n) = \exp\left(\frac{\sigma_a^2(1-\rho^2)}{2} \right) J^{1-1/\widetilde{\theta}} E_\eta \left[\eta^{\rho \sigma_a/\sigma_h} \Gamma \left(1 - 1/\widetilde{\theta}, \left(\frac{\eta x_n}{\widetilde{P}(x_n)} \right)^{\widetilde{\theta}} J^{-1} \right) \right].$$

22

The expression (19) for the latent sales of a French firm in market n (actual sales if it enters that market) can be simplified by exploiting (23) and (8):

$$X_n^*(j) = \alpha_n(j)X_n \left(\frac{Ju_F(j)}{\pi_{nF}}\right)^{-1/\widetilde{\theta}} \widetilde{P}(x_n)^{-1}.$$

which is, in logs:

$$\ln X_n^*(j) = \Gamma_n - \widetilde{\theta}^{-1}\ln u_F(j) + \ln \alpha_n(j),$$

where the vector Γ summarizes firm-invariant country-level variables with representative element:

$$\Gamma_n = \ln(X_n/\widetilde{P}(x_n)) + \widetilde{\theta}^{-1}\ln(X_{nF}/X_n) - \widetilde{\theta}^{-1}\ln J. \tag{24}$$

Finally, we can express our entry hurdles (20) as:

$$\ln u_F(j) < \ln \overline{u}_n(j) = \widetilde{\theta}\left[\Gamma_n - \ln(\sigma E_n) + \ln \eta_n(j)\right]$$

if the firm covers the fixed cost of entering market n. Thus, for a French firm to enter market n it must be that $u_F(j) < \widetilde{u}_n(j)$ and $u_F(j) < \overline{u}_n(j)$.

To illustrate the role of the parameter J it is useful to consider the special case of $E_n = 0$. In that case the mean sales of a French firm in market n, conditional on entry, is simply X_n/J. The parameter J is the scale factor between aggregate magnitudes and firm-level magnitudes. In the extreme case of $J \to \infty$, on the other hand, the mean sales of entrants varies in proportion to σE_n. More generally, J also enters the model in a more subtle way. Notice that as J gets larger, the entry hurdle gets increasingly difficult to pass. The competitiveness hurdle, on the other hand, is invariant to J. Thus J parameterizes the relative importance of the two hurdles.

The parameters of the model are σ_a^2, σ_h^2, ρ, $\widetilde{\theta}$, J, and, for each country, σE_n. We do not try to estimate the full set of σE_n's. Based on our results for the simple case of monopolistic

competition we specify $\sigma E_n = \gamma X_n^\phi$, giving us two new parameters γ and ϕ. The vector Θ of 7 parameters of the model is then:

$$\Theta = \left(\begin{array}{ccccccc} \tilde{\theta} & J & \phi & \gamma & \sigma_a^2 & \sigma_h^2 & \rho \end{array} \right)'.$$

For a value of Θ we can simulate a dataset of firms competing in each of 113 markets around the world, following the procedure described in the appendix. We can then extract firms located in France from that simulated dataset, and observe their entry and sales in markets around the world. Moments generated by these simulated data can then be compared with the actual data.

5.2   Estimation by Simulated Method of Moments

To estimate Θ we seek a value that generates a simulated dataset that approximates the actual data in the following moments:

1. The number of French firms entering each of 113 destinations.

2. The fraction of simulated firms selling the amount sold by the actual 5th percentile of sales in each country.

3. The fraction of simulated firms selling the amount sold by the actual 75th percentile of sales in each country.

4. The fraction of simulated firms selling the amount sold by the actual 95th percentile of sales in each country.

5. The number of firms selling to b or more markets.

6. Average sales in France of firms selling to b or more markets.

7. Average exports of firms selling to b or more markets.

8. The number of firms selling to subsets of the 7 most popular destinations.

Using the amoeba algorithm we searched over values of Θ to minimize the difference between these moments of our simulated data and the actual data.

5.3   Parameter Estimates

The procedure yielded the following estimates for Θ:

$$\begin{array}{ccccccc} \widetilde{\theta} & J & \phi & \gamma & \sigma_\alpha^2 & \sigma_\eta^2 & \rho \\ 1.5 & 13*10^6 & .30 & .0034 & 1.8 & 0.95 & -.32 \end{array}$$

Note first that our estimation yields the same value of $\widetilde{\theta}$ as that provided by the calibration of the simpler model. Our estimate of the elasticity of entry with respect to market size is only slightly lower. Moreover, the value of J is enormous. Our simulations almost never delivered multiple potential suppliers of the same good, consistent with monopolistic competition. The richer model's predictions about how entry varies with market size, and about sales in France of firms that sell to b or more markets are very similar to those of the simpler model. Figures 5, 6, and 7 compare some other moments of our simulated data with moments of the actual data. Figure 5 reports the actual and simulated number of firms selling to at least b countries, for $b = 1, 2, 4, 8, 16, 32$, and 64. Figures 6 and 7 report sales in France and export sales, respectively, according to this same classification. Figure 8 shows how our simulated data pick up on the number of firms entering into different markets.

What about the dimensions in which the simpler model failed? Remarkably, the richer model, in which country-specific shocks to entry $\eta_n(j)$ can generate deviations from a hierarchy of export markets, delivers a simulated data set of firms in which 27 percent follow the proper hierarchy among the top 7 destinations, the same fraction as in the actual data. Figure 9 plots the distribution of sales in France. The richer model, in which loglinear country-specific shocks to sales lead to a more skewed distribution of sales in the upper tale and introduces curvature in the lower tail.

We conclude that a quite simple model of monopolistic competition, with technological heterogeneity and good and country-specific shocks to entry and to sales, can pick up the basic features of the micro-level data very well.

References

Aw, Bee Yan, Chung, Sukkyun, and Roberts, Mark J. (1998), "Productivity and the Decision to Export: Micro Evidence from Taiwan and South Korea." *NBER Working Paper No. 6558.*

Bernard, Andrew B., Eaton, Jonathan, Jensen, J. Bradford, and Kortum, Samuel (2003), "Plants and Productivity in International Trade." *American Economic Review 93*: *1268-1290.*

Bernard, Andrew B. and Jensen, J. Bradford (1995), "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987." *Brookings Papers on Economic Activity: Microeconomics*: 67-119.

Bernard, Andrew B. and Jensen, J. Bradford (1999) "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics 47*: 1-25.

Biscourp, Pierre and Kramarz, Francis (2002), "Les Firmes Francaises et le Commerce Exterieur: Une Etude Descriptive sur la Periode 1986-1992," mimeo, CREST.

Chaney, Thomas (2005), "Distorted Gravity: Heterogeneous Firms, Market Structure, and the Geography of International Trade." mimeo, University of Chicago.

Clerides, Sofronis, Lach, Saul, and Tybout, James R. (1998), "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics 113*: 903-947.

Eaton, Jonathan and Kortum, Samuel. (1999), "International Technology Diffusion: Theory and Measurement." *International Economic Review 40*: 537-570.

Eaton, Jonathan and Kortum, Samuel (2002), "Technology, Geography, and Trade." *Econometrica 70*: 1741-1780.

Eaton, Jonathan, Kortum, Samuel, and Francis Kramarz (2004), "Dissecting Trade: Firms, Industries, and Export Destinations." *American Economic Review, Papers and Proceedings*, 94: 150-154.

Feenstra, Robert C., Lipsey, Robert E. and Bowen, Henry P. (1997), "World Trade Flows, 1970-1992, with Production and Tariff Data." *NBER Working Paper No. 5910.*

Ghironi, Fabio and Marc J. Melitz (2004), "International Trade and Macroeconomic Dynamics with Heterogeneous Firms," *Boston College Working Paper No. 599.*

Hallward-Driemeier, Mary, Iaross, Giusepe, and Sokoloff, Kenneth (2002), "Exports and Manufacturing Productivity in East Asia: A Comparative Analysis with Firm-Level Data." NBER WP #8894

Helpman, Elhanan, Melitz, Marc J., and Yeaple, Stephen R. (2004), "Export versus FDI with Heterogenous Firms," forthcoming, *American Economic Review.*

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

OECD (1995), *The OECD STAN Database.*

Roberts, Mark and Tybout, Robert (1997), "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs." *American Economic Review 87*: 545-564.

Sutton, John.(1991), *Sunk Costs and Market Structure.* (Cambridge, MA: The MIT Press).

UNIDO (1999), *Industrial Statistics Database.*

6 Appendix: The Simulation Algorithm

Given a vector of parameters we can simulate the behavior of our sample of French firms. We will describe our algorithm as if we were simulating an arbitrary number French firms, indexed by $j = 1, \ldots S$. But, we can easily scale the results to be comparable to our data on all French firms selling in France. We introduce notation here for the data on French firms, indexed by $j = 1, \ldots, J_{data}$. We define the indicator $S_n(j) = 1$ if we observe firm j selling in destination n, and $S_n(j) = 0$ if we observe no sales. In the destinations n in which we observe sales (where $S_n(j) = 1$), we let $y_n(j)$ be the natural logarithm of the firm's sales (and arbitrarily set $y_n(j) = 0$ when $S_n(j) = 0$). Thus: $y_n(j) = S_n(j) \ln X_n(j)$.

1. Stage 1 does not require any parameter values and uses data only on the world bilateral trade matrix (expressed as shares of the importer's absorption) with representative element $\pi_{ni} = X_{ni}/X_n$. It involves four steps.

   (a) Draw $v_i(j)$'s independently from $U[0, 1]$, for $i = 1, \ldots, 113$ and $j = 1, \ldots, S$.

   (b) For $i \neq F$ calculate $S \times 112$ values of:

   $$u_i(j) = -\ln\left[1 - v_i(j)\right]$$

   (c) Use the $u_i(j)$'s and the π_{ni}'s to construct $S \times 113$ competitiveness hurdles:

   $$\widetilde{u}_n(j) = \min_{i \neq F}\{\pi_{nF}u_i(j)/\pi_{ni}\}.$$

   (d) Independently draw $S \times 113$ realizations of $a_n(j)$ and $h_n(j)$ from:

   $$\begin{bmatrix} a_n(j) \\ h_n(j) \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

30

2. Stage 2 requires data for each destination n on X_n and π_{nF} as well as a set of parameters. It involves three steps,

   (a) Fix values for $\widetilde{\theta}, \gamma, \phi, \sigma_a^2, \sigma_h^2, \rho,$ and J.

   (b) Calculate $\widetilde{P}(x_n)$ for each destination n as the solution to:

   $$\widetilde{P}(x_n) = \exp\left(\frac{\sigma_a^2(1-\rho^2)}{2}\right) J^{1-1/\widetilde{\theta}} E_\eta\left[\eta^{\rho\sigma_a/\sigma_h}\Gamma\left(1-1/\widetilde{\theta}, \left(\frac{\eta x_n}{\widetilde{P}(x_n)}\right)^{\widetilde{\theta}} J^{-1}\right)\right]$$

   (c) Calculate:

   $$\Gamma_n = \ln(X_n/\widetilde{P}(x_n)) + \widetilde{\theta}^{-1}\ln\pi_{nF} - \widetilde{\theta}^{-1}\ln J.$$

   for each destination n.

3. Stage 3 combines the simulation draws from Stage 1 and the parameter values and destination variables from Stage 2. It involves seven steps.

   (a) Use the draws from 1d and the parameter values from 2a to construct $S \times 113$ realizations for each of $\ln\alpha_n(j)$ and $\ln\eta_n(j)$ as:

   $$\begin{bmatrix} \ln\alpha_n(j) \\ \ln\eta_n(j) \end{bmatrix} = \begin{bmatrix} \sigma_a\sqrt{1-\rho^2} & \sigma_a\rho \\ 0 & \sigma_h \end{bmatrix}\begin{bmatrix} a_n(j) \\ h_n(j) \end{bmatrix}$$

   (b) Construct the $S \times 113$ entry hurdles:

   $$\overline{u}_n(j) = \exp\left\{\widetilde{\theta}\left[\Gamma_n - \ln(\sigma E_n) + \ln\eta_n(j)\right]\right\}.$$

   (c) Construct S joint hurdles $\overline{u}(j)$ faced by a French firm in its home market:

   $$\overline{u}(j) = \min\{\overline{u}_F(j), \widetilde{u}_F(j)\}.$$

   Note that a French firm will sell in France if and only if it passes both the entry hurdle and the competitiveness hurdle there, i.e. iff $u_F(j) \leq \overline{u}(j)$.

(d) Construct S probability weights:

$$p[\bar{u}(j)] = 1 - \exp[-\bar{u}(j)].$$

If we were to construct $u_F(j)$ in a manner parallel to how we constructed all the other $u_i(j)$'s in step 1 b, $p[\bar{u}(j)]$ would be the probability of the French firm selling in France.

(e) We actually construct $u_F(j)$, based on the draw $v_F(j)$ from step 1 a, so that we necessarily obtain a French firm selling in France. To do so, we set:

$$u_F(j) = -\ln\{1 - p[\bar{u}(j)]v_F(j)\}.$$

Our simulated French firm gets a weight $p[\bar{u}(j)]$ in the sample.

(f) Calculate $S_n(j)$ as determined by the competition and entry hurdles:

$$S_n(j) = \begin{cases} 1 \text{ if } u_F(j) \leq \tilde{u}_n(j) \text{ and } u_F(j) \leq \bar{u}_n(j) \\ 0 \text{ otherwise.} \end{cases}$$

(g) Wherever $S_n(j) = 1$ calculate log sales as:

$$\ln X_n(j) = \Gamma_n - \tilde{\theta}^{-1}\ln u_F(j) + \ln \alpha_n(j).$$

Following this procedure we simulate the behavior of S firms. In generating statistical moments from this simulated sample, we need to keep track of two issues. First, when summing across firms, we must apply the sampling weight $p[\bar{u}(j)]$ to firm j. Second, if we want to mimic the scale of the French data, we need to apply a scaling factor of J/S. In this way the choice of S matters only for the variance of the resulting simulated moments.
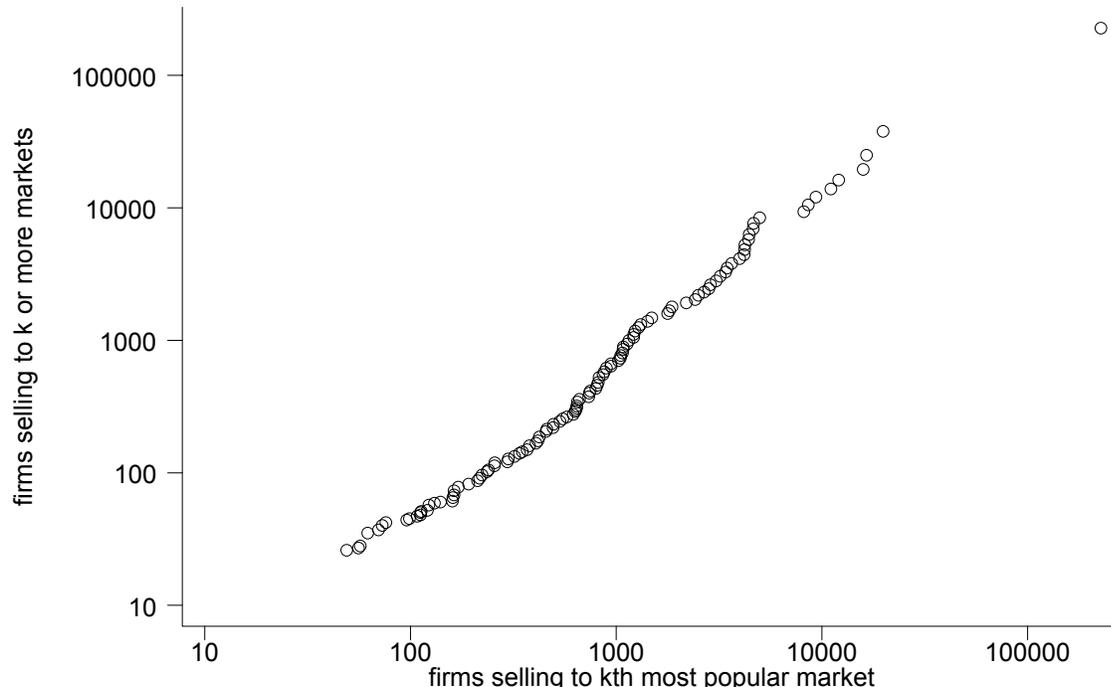
Figure 1: Market Hierarchy for French Firms
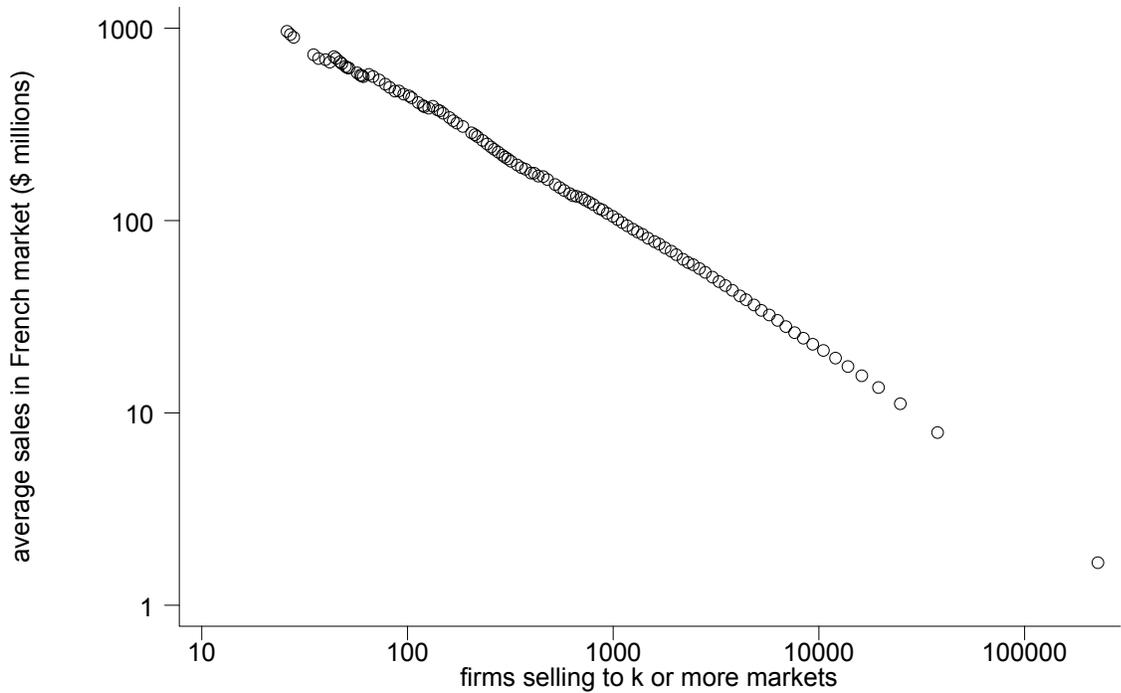
Figure 2a: Firm Size and Frequency of Multiple Markets
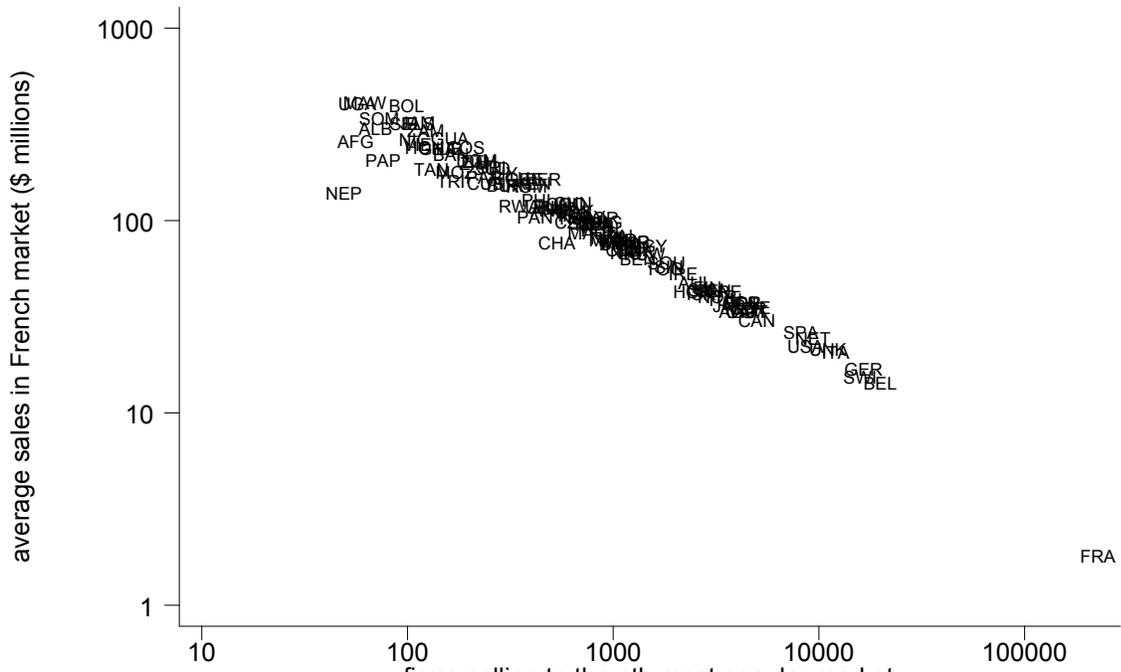
Figure 2b: Firm Size and Popularity of Market
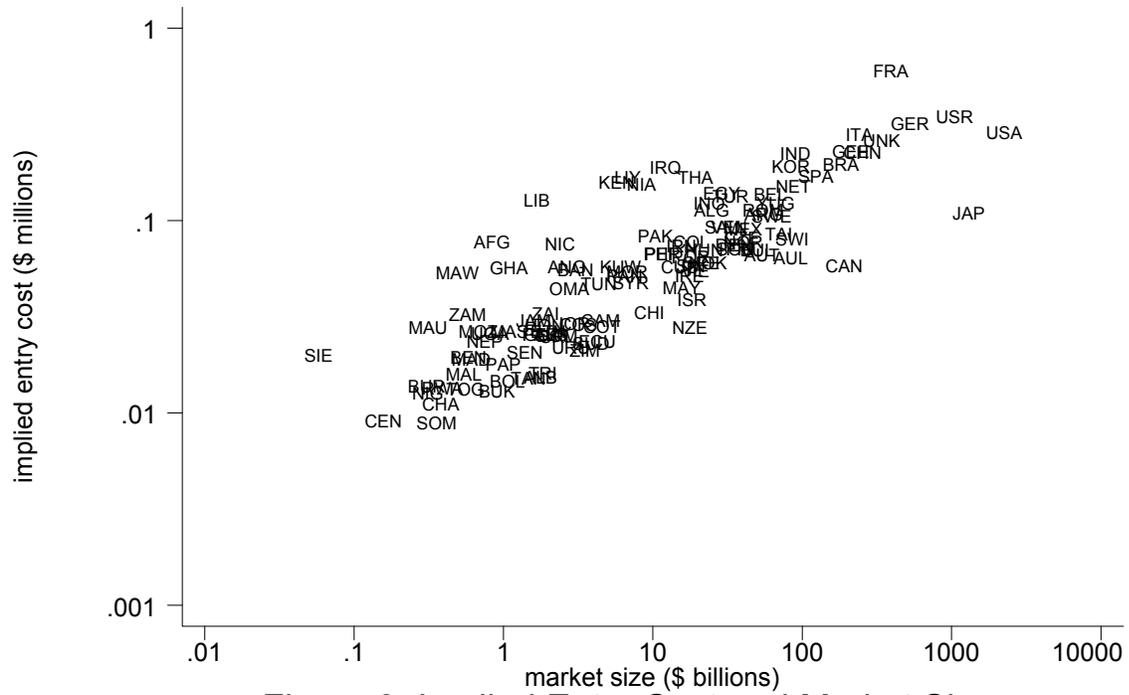
Figure 3: Implied Entry Cost and Market Size

stata™
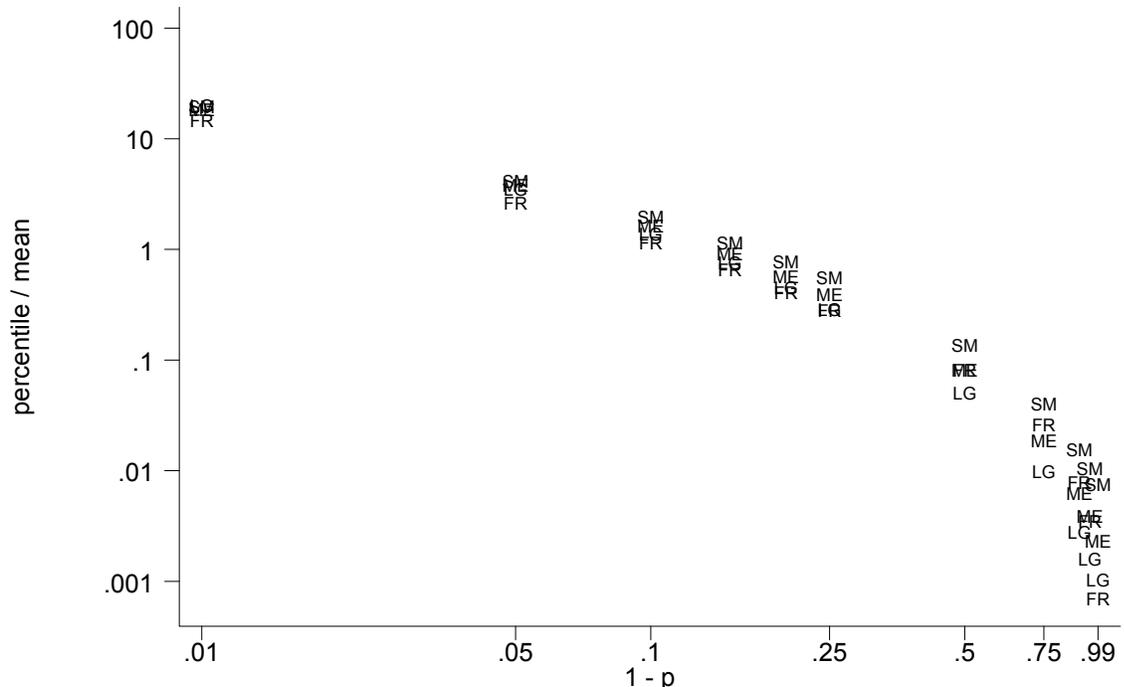
Figure 4: Distribution of Sales, by Market Size

Figure 5: Firms Exporting to B or More Countries
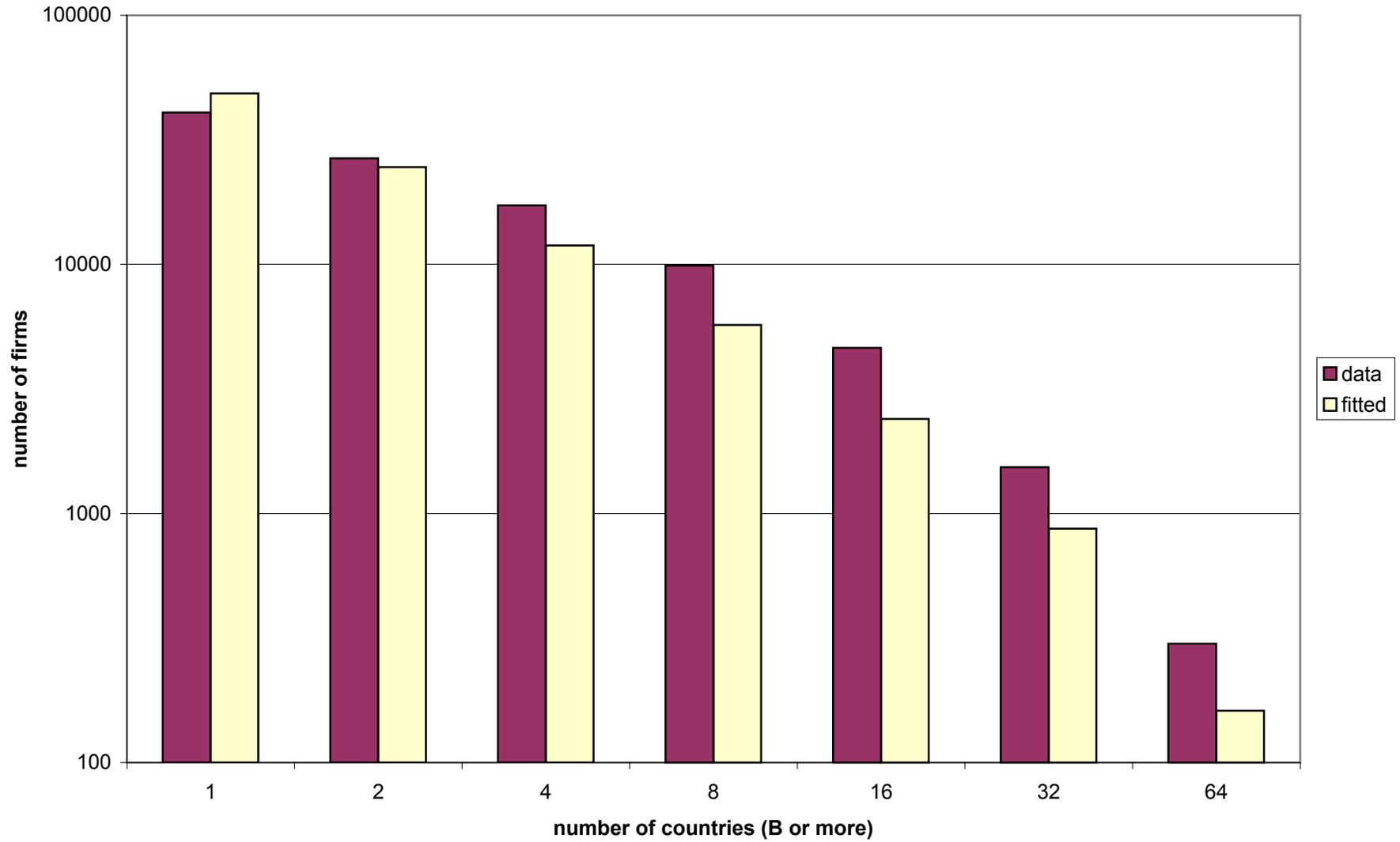
Figure 6: French Sales of Firms Exporting to B or More Countries
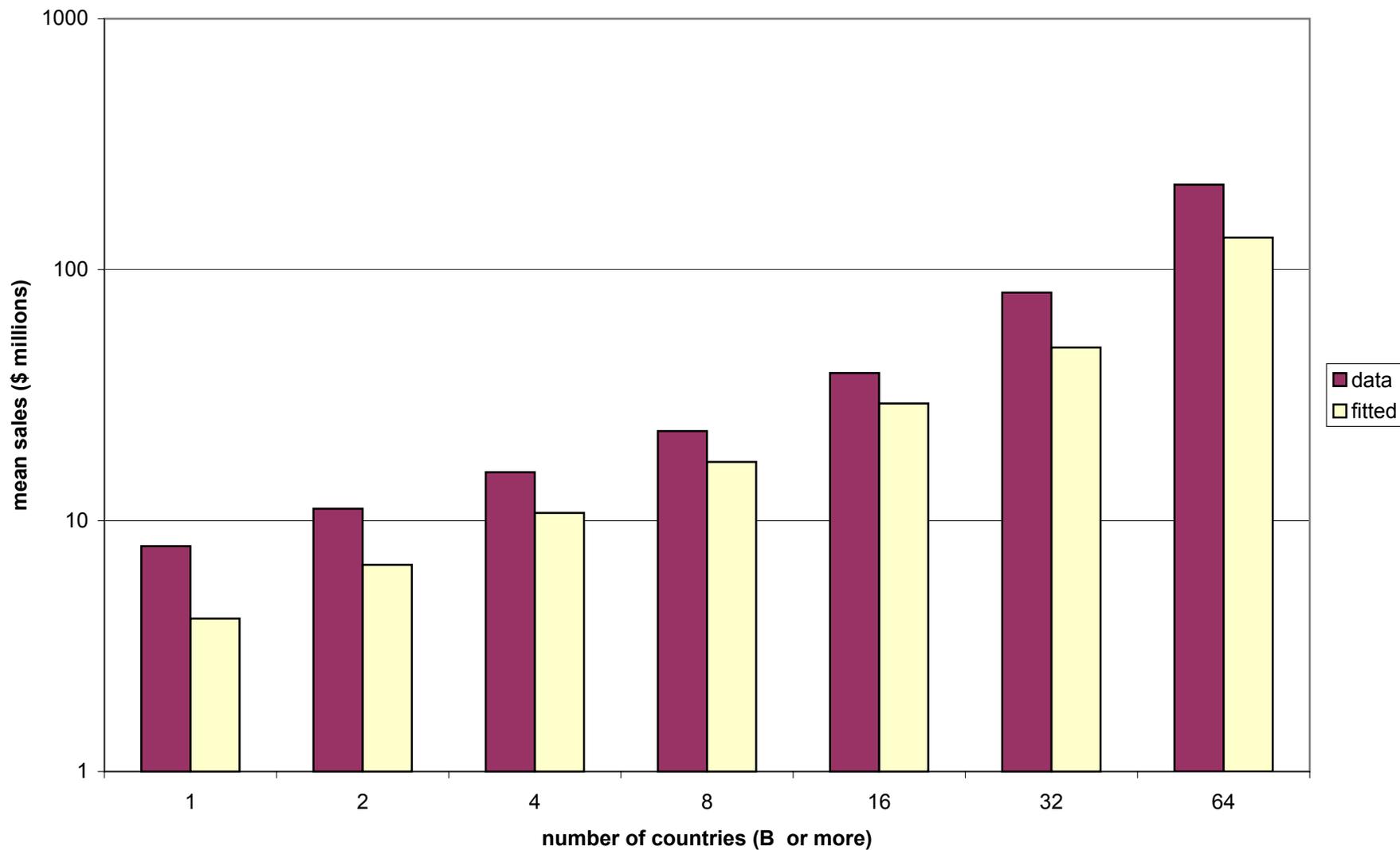
**Figure 7: Exports of Firms Selling to B or More Countries**
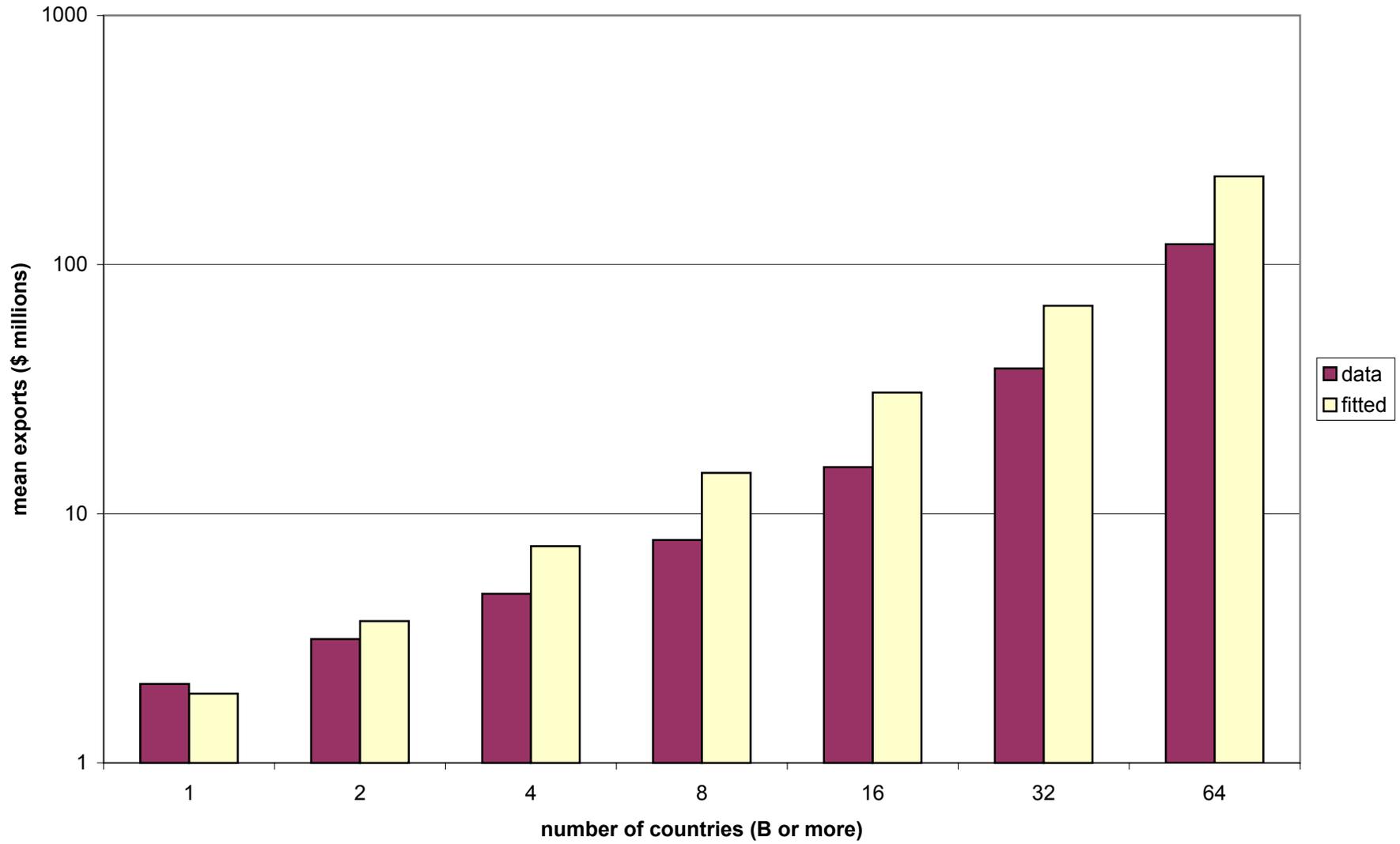
Figure 8: French Firm Entry by Country
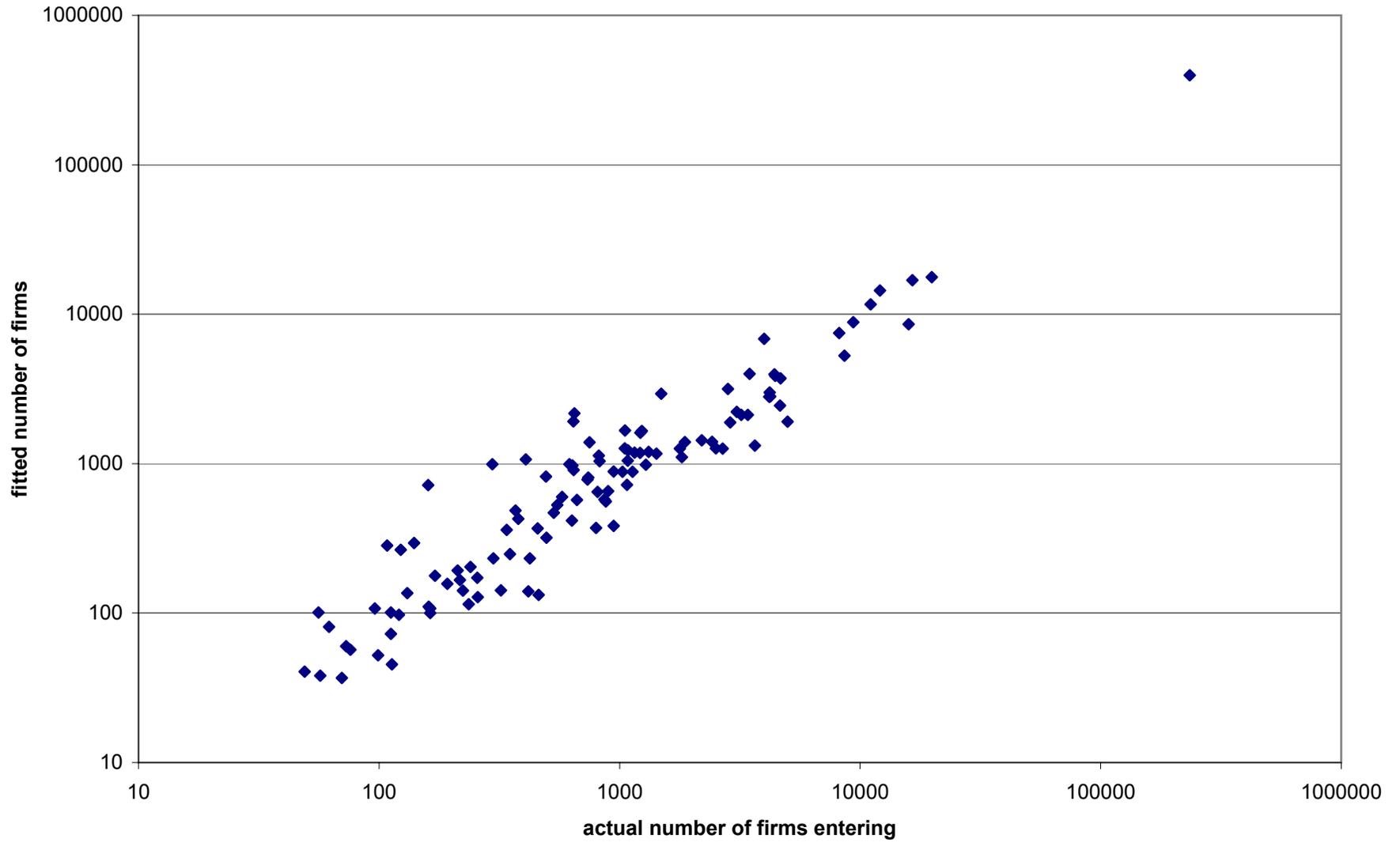
Figure 9: Fitted Sales Distribution in France