An Anatomy of International Trade: Evidence from French Firms

Jonathan Eaton,\(^1\) Samuel Kortum,\(^2\) and Francis Kramarz\(^3\)

March 2008

\(^1\)Department of Economics, New York University, and National Bureau of Economic Research (jonathan.eaton@nyu.edu).

\(^2\)Department of Economics, University of Chicago and National Bureau of Economic Research (kortum@uchicago.edu).

\(^3\)CREST-INSEE and CEPR (kramarz@ensae.fr).
Abstract

We examine the cross section of sales of French manufacturing firms in 113 destinations, including France itself. Several regularities stand out: (1) The number of French firms selling to a market, relative to French market share, increases with market size. (2) Average sales in France rise very systematically with selling to less popular markets and to more markets. (3) Sales distributions are very similar across markets of very different size and extent of French participation. We adopt a model of firm heterogeneity and export participation which we estimate to match moments of the French data using the method of simulated moments. The results imply that nearly half the variation across firms that we see both in market entry can be attributed to a single dimension of underlying firm heterogeneity, efficiency. Conditional on entry underlying efficiency accounts for a much smaller variation in sales in any given market.
1 Introduction

We exploit a detailed set of data on the exports of French firms to confront a new generation of trade theories. The data, from French customs, report the sales of over 200,000 individual firms to over 100 individual markets in a cross section.

We ask how well the model of the export behavior of heterogeneous firms introduced by Melitz (2003) and more concretely specified by Chaney (2008) and Helpman, Melitz, and Yeaple (2004) stands up to these data. Basic elements of the model are that firms’ efficiencies follow a Pareto distribution, demand is Dixit-Stiglitz, and markets are separated by iceberg trade barriers and require a fixed cost of entry. The model is the simplest one we can think of that can square with the facts.

With this basic model in mind, we extract three relationships that underlie the data: (1) how entry varies with market size, (2) how export participation abroad connects with sales at home, and (3) how the distribution of sales varies across markets. Through the haze of numbers we begin to see the outlines of the basic model, and even rough magnitudes of parameters. The basic model fails to come to terms with some features of the data, however: Firms don’t enter markets according to an exact hierarchy and their sales where they do enter deviate from the exact correlations the basic model would insist upon.

To reconcile the basic model with these failures we extend it by introducing market and firm-specific heterogeneity in demand and entry costs. We also incorporate a reduced form version of Arkolakis’s (2007) market access cost. The extended model, while remaining very parsimonious and transparent, is one that we can connect more formally to the data. We describe how the model can be simulated, and estimate its parameters using the method of
simulated moments.

With the parameter estimates in hand we find that the forces underlying the basic model remain very powerful. Simply knowing a firm’s efficiency improves our ability to explain the probability it sells in any market by nearly fifty percent. Conditional on a firm selling in a market, knowing its efficiency improves our ability to predict how much it sells there, by only 6 percent. While these results leave much to be explained by the idiosyncratic interaction between individual firms and markets, they tell us that any theory that ignores features of the firm that are universal across markets misses much.

A feature of all models of monopolistic competition with a fixed cost of entry is that big markets, because they can accommodate more variety, are better places to live. With heterogenous firms this effect is attenuated by the result that entry on the margin is by higher cost producers. Our parameter estimates indicate that in fact firm heterogeneity reduces the benefit of size by a half.

With homogeneous firms, fixed costs dissipate all variable profit. With firm heterogeneity substantial profits remain for firms far from the threshold of entry. Our parameter estimates suggest that fixed costs dissipate about half of variable profit.
2 Three Regularities

Our data, described in Appendix A, are the sales, translated into U.S. dollars, of 229,900 French manufacturing firms to 113 markets in 1986. Among them fewer than 34,035 sell elsewhere than in France. The firm that exports most widely sells to 110 of these destinations. At least one firm sells to every smaller number of destinations.

We assemble our complex data in three different ways that reveal sharp regularities. (1) We show how the number of firms selling in a market varies with the size of the market, first looking at the raw numbers and then the raw numbers normalized by French market share. (2) We then look at sales in France by firms selling (a) to less popular destinations and (b) to more destinations. (3) We look at features of the distribution of sales within individual markets.

2.1 Market Entry

Figure 1a plots the number of French manufacturing firms $N_{nF}$ selling to a market against total manufacturing absorption $X_n$ in that market across our 113 markets. While the number of firms selling to a market tends clearly to increase with the size of the market, the relationship is a cloudy one. Note in particular that more French firms sell to France than its market size would suggest.

---

1Eaton, Kortum, and Kramarz (EKK, 2004), describe the data in detail, partitioning firms into 16 manufacturing sectors. While features vary across industries, enough similarity remains to lead us to ignore the industry dimension here.

2Manufacturing absorption is calculated as total production plus imports minus exports. Aggregate trade data are from Feenstra (2000) while production data are from UNIDO (2000). See EKK (2004) for details.
The relationship comes into focus, however, when the number of firms is normalized by the share of France in a market. Figure 1b continues to report market size across the 113 destinations along the $x$ axis. The $y$ axis replaces the number of French firms selling to a market with that number divided by French market share, $\pi_{nF}$, defined as total French exports to that market, $X_{nF}$, divided by absorption, i.e.,

$$\pi_{nF} = \frac{X_{nF}}{X_n}.$$ 

Note that the relationship is not only very tight, but linear in logs, with a slope of about two-thirds. Correcting for market share pulls France from the position of a large positive outlier to a slight negative one.

If we make the assumption that French firms don’t vary systematically in size from other firms selling in a market, the measure on the $y$ axis can serve as an indicator of the total number of firms selling in a market. We can then interpret Figure 1b as telling us how the number of sellers varies with market size.

Models of perfect and Bertrand competition and the standard model of monopolistic competition without market specific entry costs predict that the number of sellers in a market is invariant to market size. Figures 1a and 1b compel us to abandon these approaches.

The number of firms selling to a market increases with market size, but with an elasticity less than one. A mirror relationship (not shown) is that average sales per firm increases with market size as well, again with an elasticity less than one. We can get a fuller picture by asking how sales per firm rises with market size at different points in the sales distribution there. Figure 4a shows this relationship, reporting the 95th, 75th, 50th, and 25th percentiles of sales in each market (on the $y$ axis) against market size (on the $x$ axis). The upward drift
is apparent across the board, although more weakly for the 25th percentile.

2.2 Export Participation and Size in France

How does a firm’s participation in export markets relate to its sales in France? We organize our firms in two different ways.

First, we rank destinations according to their popularity as destinations for exports. We then look at average sales in France of firms selling to different markets according to their rank in popularity. The most popular destination is of course France itself, where all of our firms sell, followed by Belgium-Luxembourg with 17,699 exporters. The least popular is Nepal, where only 43 French firms sell (followed in unpopularity by Afghanistan and Uganda, with 52 each).

Figure 2a depicts average sales in France on the $y$ axis of those firms selling to the $k$th most popular market, where $k$ is reported on the $x$ axis. The relationship is clearly increasing: Firms that sell to less popular markets sell more in France.

Figure 2a does not tell us about the number of firms in each group. Figure 2b orders firms according to the number in each group, going from least to most popular, on the $x$ axis. The $y$ axis again reports average sales in France. The relationship is tight and linear in logs.

Second, we group firms according to the minimum number of destinations where they sell. All of our firms, of course, sell to at least one market while none sell to all 113 destinations. Figure 2c depicts average sales in France on the $y$ axis for the group of firms that sell to at least $k$ markets with $k$ on the $x$ axis. The similarity with Figure 2a is apparent. Selling to more markets has a very similar positive association with sales in France as selling to less
popular markets.

Figure 2d, analogous to Figure 2b, reports average sales in France of firms selling to $k$ or more markets against the number of firms selling to $k$ or more markets. The similarity with Figure 2b is apparent. Since the axes are the same we can place one figure on top of the other and hold them to compare their slopes. They are nearly the same. The highly linear, in logs, negative relationship between the number of firms that export to a group of countries and their sales in France is highly suggestive of a power law.

We conclude that firms that sell to less popular markets and sell to more markets systematically sell more in France. To what extent are we looking at the same firms in these two exercises? If destinations obey an exact hierarchy, in the sense that any firm selling to the $k + 1$st most popular destination necessarily sold to the $k$th most popular as well, the two groups would coincide. Not surprisingly firms are less orderly in their choice of destinations. A good metric of how far they depart from a hierarchy is elusive. We can get some sense, however, by looking simply at exporters to the top seven foreign destinations. Table 1 reports these destinations and the number of firms selling to each. It also reports the total number that export to at least one of these destinations and the total number of exporters. Note that 4106 of the 34,035 exporters, constituting only 12 percent, don’t sell in the top 7. The last column of the table reports, for each top 7 destination, the marginal probability of selling there conditional on selling somewhere among the top 7.

We use these marginal probabilities to calculate the probability of selling to the markets in the order prescribed by a hierarchy if the probabilities of selling in any market were independent across markets. The first column of Table 2 lists each of the strings of destinations
that obey a hierarchical structure while column 2 reports the number of firms exporting to that string, irrespective of their export activity outside the string. The third column reports the probability that a firm would export to that string if the probabilities of exporting to each destination were independent across markets. Independence implies that only 13.5 percent of exporters would obey the required ordering (for example, selling to Belgium and Germany but not the other five), so that 86.5 would deviate from it (e.g., by selling to Belgium, Germany, and the United Kingdom but nowhere else). In fact more than twice that number, 30.9 percent, adhere to the hierarchy. Column 4 reports the implied number of firms that would sell to each string under independence. Note that many more sell to the short strings and fewer to the long strings than independence would imply. We conclude that a model needs to recognize both a tendency for firms to export according to a hierarchy while allowing them significant latitude to depart from it.

Delving further into the French sales of exporters to markets of varying popularity, Figure 4b reports the 95th, 75th, 50th, and 25th percentile of sales in France (on the y axis) against the number of firms selling to each market. Note the tendency of sales in France to rise with the unpopularity of a destination across all percentiles (less systematically so for the 25th percentile).

2.3 Sales Distributions

Our third exercise is to look at the distribution of sales within individual markets. For all 113 destinations we have plotted the percentiles \( q \) of French sales in that market normalized by their mean in that market, \( x^q_{nt} \), against \( q \). Figures 3a, 3b, and 3c plot the results for
France, the United States, and Ireland, with axes the same. Since there are many fewer firms exporting than selling in France the upper percentiles in the foreign destinations are empty. Nevertheless, stacking one on top of the other and holding them to the light reveals a remarkable similarity across the three.

To interpret these figures as distribution we can write:

$$\Pr [X_n \leq x^n_q] = q.$$  

where $X_n$ is sales in market $n$ relative to the mean. Suppose the distributions were Pareto with parameter $a > 1$. We could then write:

$$1 - \left( \frac{ax^n_q}{a - 1} \right)^{-a} = q$$

or:

$$\ln (x^n_q) = \ln \left( \frac{a - 1}{a} \right) - \frac{1}{a} \ln(1 - q).$$

implying a straight line with slope $-1/a$. At the top percentiles the slope does appear nearly constant at close to $-1$ but at the lower tails it is much steeper, reflecting the presence of suppliers selling very small amounts. This shape is well known in the industrial organization literature looking at various size measures in the home market.\textsuperscript{3} What we find here is that this shape is inherited across markets looking at the same set of potential sellers. The deviation from Pareto is indicative of a log normal distribution. A challenge for modeling is capturing this size distribution of sales within any given market with the stark linear (in logs) relationships in Figures 2b and 2d.

\textsuperscript{3}See Simon and Bonini (1958) and Luttmer (2006), among many, for a discussion and explanations.
3 Theory

In seeking to explain these relationships we turn to a parsimonious model that delivers predictions about where firms sell and how much they sell there. We infer the parameter values of the model from our observations on French firms. To this end we build on Melitz (2003), Helpman, Melitz, and Yeaple (2004), Chaney (2008), and Arkolakis (2007). We begin with a description of demand and market structure in a typical market. We then describe technology and sources of producer heterogeneity.

3.1 Demand, Market Structure, and Entry

A market $n$ contains a measure of potential buyers indexed by $l$ each spending some amount $X^l$. An individual buyer $l$ of good $j$ in country $n$ (who might be using it for final consumption or as an intermediate input into production) maximizes a standard CES aggregator with elasticity of substitution $\sigma > 1$. Denote by $\Omega^l_n$ the set of goods available to buyer $l$. With a continuum of available goods her demand for good $j$ with price $p_n(j)$ is:

$$X^l_n(j) = \alpha_n(j) \left( \frac{p_n(j)}{P^l_n} \right)^{1-\sigma} X^l_n$$

(1)

where $P^l_n$ is the CES price index goods available to her:

$$P^l_n = \left\{ E \left[ \alpha_n(j) (p_n(j))^{1-\sigma} | \forall j \in \Omega^l_n \right] \right\}^{1/(1-\sigma)}.$$

Here $\alpha_n(j) \geq 0$ is a demand shifter specific to good $j$ in market $n$.

We follow Arkolakis (2007) in assuming that, in order to sell to a fraction $f$ of the buyers in country $n$, a producer selling good $j$ must incur a fixed cost:

$$E_n(j) = \varepsilon_n(j) E_n \frac{1 - (1 - f)^{1-1/\lambda}}{1 - 1/\lambda}.$$
Here $\varepsilon_n(j)$ is a cost shock specific to good $j$ in market $n$, $E_n$ is a component of the cost shock faced by all who sell there, and $\lambda > 0$ is a parameter that reflects the increasing cost of reaching a larger fraction of potential buyers.

All potential buyers in the market have the same probability $f$ of being reached by a particular seller that is independent across sellers. Hence they face the same distribution of prices and price index $P_n$. Conditional on selling in a market the producer of good $j$ with unit cost $c_n(j)$ who charges a price $p$ and reaches a fraction $f$ of buyers earns a profit:

$$\Pi_n(j) = \left(1 - \frac{c_n(j)}{p}\right)\alpha_n(j)f\left(\frac{p}{P_n}\right)^{1-\sigma}X_n - \varepsilon_n(j)E_n\frac{1 - (1 - f)^{1-1/\lambda}}{1 - 1/\lambda}$$

where $X_n$ is total expenditure in the market.

To maximize profit a producer will set the standard Dixit-Stiglitz (1977) markup over unit cost:

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

where:

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

and seek a fraction:

$$f_n(j) = 1 - \left[\eta_n(j)\frac{X_n}{\sigma E_n}\left(\frac{\overline{m}c_n(j)}{P_n}\right)^{1-\sigma}\right]^{-\lambda}$$

of buyers in the market where:

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

the ratio of the demand shock to the entry-cost shock. It will choose to sell anything at all if and only if:

$$\frac{d\Pi_n(j)}{df} > 0$$
at $f = 0$ or if:

$$\alpha_n(j) \left( \frac{m c_n(j)}{P_n} \right)^{1-\sigma} \frac{X_n}{\sigma} \geq \varepsilon_n(j) E_n.$$

This condition determines a ceiling unit cost:

$$\bar{c}_n(\eta) = \left( \frac{\eta_n(j) X_n}{\sigma E_n} \right)^{1/(\sigma-1)} \frac{P_n}{\bar{m}} \tag{3}$$

such that a firm will enter market $n$ only if its unit cost $c_n(j)$ and realization of $\eta_n(j)$ satisfy:

$$c_n(j) \leq \bar{c}_n(\eta_n(j)).$$

We can use the expression for (3) to simplify the expression for the fraction of buyers a producer with unit cost $c_n(j) \leq \bar{c}_n(\eta_n(j))$ will reach:

$$f_n(j) = 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)}.$$

Its total sales are then:

$$X_n(j) = \bar{\alpha}_n(j) \left( \frac{m c_n(j)}{P_n} \right)^{1-\sigma} X_n \tag{4}$$

so that its profit is:

$$\Pi_n(j) = \bar{\alpha}_n(j) \left( \frac{m c_n(j)}{P_n} \right)^{1-\sigma} \frac{X_n}{\sigma} - \bar{\varepsilon}_n(j) E_n$$

where:

$$\bar{\alpha}_n(j) = \alpha_n(j) \left[ 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)} \right]$$

and

$$\bar{\varepsilon}_n(j) = \varepsilon_n(j) \frac{1}{1 - 1/\lambda} \left[ 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)} \right].$$

We thus fold the variation across sellers in the range of buyers they reach into the sales and entry-cost shocks.
To summarize, the particular situation of a potential seller of product $j$ in market $n$ is captured by three magnitudes: the unit cost $c_n(j)$ and the demand and entry cost shocks $\alpha_n(j)$ and $\varepsilon_n(j)$. The relevant characteristics of market $n$ that apply across sellers are total purchases $X_n$, the price index $P_n$, and the common component of the entry cost $E_n$. We treat $\alpha_n(j)$ and $\varepsilon_n(j)$ as the realization of producer-specific shocks. In the next section we relate $c_n(j)$ to underlying heterogeneity in the efficiency of different potential producers in different countries, input costs, and trade barriers. We then turn to the determination of the price index and expenditure.

3.2 Producer Heterogeneity

A potential producer of good $j$ in country $i$ has efficiency $z_i(j)$. A bundle of inputs there costs $s_i$, so the unit cost of producing good $j$ is $s_i/z_i(j)$. Finally, countries are separated by iceberg trade costs, so that delivering one unit of a good to country $n$ from country $i$ requires shipping $d_{ni} \geq 1$ units, where we set $d_{ii} = 1$ for all $i$. Combining these terms, the unit cost of delivering one unit of good $j$ to country $i$ from country $n$ is:

$$c_{ni}(j) = s_i d_{ni} / z_i(j).$$

(5)

The measure of potential producers in country $i$ who can produce their good with efficiency at least $z$ is:

$$\mu_i^Z(Z \geq z) = T_i z^{-\theta} \quad z > 0$$

where $\theta$ is a positive parameter where we restrict $\theta > \sigma$.

\footnote{We follow Helpman, Melitz, and Yeaple (2004) and Chaney (2007) in treating the underlying heterogeneity in efficiency as Pareto. Our observations above on patterns of sales by French firms in different markets are...} Using (5), the measure of goods
that can be delivered from country $i$ to country $n$ at unit cost $C \leq c$ is $\mu_{ni}(c)$ defined as:

$$\mu_{ni}(c) = \mu_t^Z \left( Z \geq \frac{w_id_{ni}}{c} \right) = T_i(w_id_{ni})^{-\theta} c^\theta.$$ 

The measure of goods that can be delivered to country $n$ from anywhere at unit cost $c$ or less is therefore:

$$\mu_n(c) = \sum_{i=1}^{N} \mu_{ni}(c) = \Phi_n c^\theta$$

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

Within this measure, the fraction originating from country $i$ is:

$$\frac{\mu_{ni}(c)}{\mu_n(c)} = \frac{T_i(w_id_{ni})^{-\theta}}{\Phi_n} = \pi_{ni}.$$ \hspace{1cm} (6)

where $\pi_{ni}$, which arises frequently in what follows, is invariant to $c$. 

very suggestive of an underlying Pareto distribution. A Pareto distribution of efficiencies can arise naturally from a dynamic process that is a history of independent shocks, as shown by Simon (1956), Gabaix (1999), and Luttmer (2006). The Pareto distribution is closely linked to the type II extreme value (Fréchet) distribution used in Kortum (1977), Eaton and Kortum (2002), and Bernard, Eaton, Kortum, and Jensen (2003). Say that the range of goods is limited to the interval $j \in [0, J]$ with the measure of goods produced with efficiency at least $z$ is given by: $\mu^Z_t(Z \geq z; J) = J \left\{ 1 - \exp \left[ - (T/J) z^{-\theta} \right] \right\}$ (where $J = 1$ in these previous papers). This generalization allows us to stretch the range of goods while compressing the distribution of efficiencies for any given good. Taking the limit as $J \to \infty$ gives the expression above. To take the limit rewrite the expression as $\{1 - \exp \left[ - (T/J) z^{-\theta} \right]\} / J^{-1}$ and apply L'Hôpital’s rule.
3.3 The Price Index and Entry Cutoffs

Under our assumption about the distribution of efficiencies, the price index in market \( n \) is determined by:

\[
\left( \frac{P_n}{m} \right)^{1-\sigma} = E_{\eta_n} [E(\bar{\alpha}_n|\eta_n)c^{1-\sigma}] \\
= \Phi_nE_{\eta_n} \left[ \int_0^{\tau_n(\eta)} E(\bar{\alpha}_n|\eta_n)\theta c^{\theta-\sigma} dc \right] \\
= \Phi_nE_{\eta_n} \left\{ E(\alpha_n|\eta_n) \left[ \int_0^{\tau_n(\eta)} \theta c^{\theta-\sigma} dc - \tau_n(\eta_n)^{-(\sigma-1)} \int_0^{\tau_n(\eta)} \theta c^{\theta-\sigma+\lambda(\sigma-1)} dc \right] \right\} \\
= \Phi_n \left[ \frac{\tilde{\theta}}{\theta-1} - \frac{\tilde{\theta}}{\theta+\lambda-1} \right] E_{\eta_n} \left[ E(\alpha_n|\eta_n)\bar{c}_n(\eta_n)^{\theta-(\sigma-1)} \right].
\]

where \( E_{\eta_n} \) denotes taking the expectation over \( \eta_n \) and where:

\[
\tilde{\theta} = \frac{\theta}{\sigma-1} > 1. \quad (7)
\]

Substituting the expression for the entry hurdle (3) and simplifying gives:

\[
\left( \frac{P_n}{m} \right)^{-\theta} = \kappa_1 \Phi_n \left( \frac{X_n}{\sigma E_n} \right)^{\tilde{\theta}-1} \quad (8)
\]

where:

\[
\kappa_1 = \left[ \frac{\tilde{\theta}}{\theta-1} - \frac{\tilde{\theta}}{\theta+\lambda-1} \right] E \left[ \alpha_n\eta_n^{\tilde{\theta}-1} \right].
\]

The term \( \tilde{\theta} \) comes up repeatedly as it translates unobserved heterogeneity in underlying producer efficiency into observed heterogeneity in sales. A higher value of \( \tilde{\theta} \) implies less heterogeneity in efficiency while a higher value of \( \sigma \) means that given efficiency heterogeneity translates into greater sales heterogeneity. Since we observe sales and not underlying efficiency we are able to identify only the cluster of parameters \( \tilde{\theta} \).

Substituting the price index (8) into (3) we get:

\[
[\bar{c}_n(\eta_n)]^\theta = \left( \frac{\eta_n X_n}{\sigma E_n} \right)^{\tilde{\theta}} \left( \frac{P_n}{m} \right)^{\theta} = \frac{\eta_n^{\tilde{\theta}} X_n}{\kappa_1 \Phi_n \sigma E_n}.
\]
For any $\eta_n(j) = \eta$, a good can be profitably sold in market $n$ only if it can be supplied there at a cost less than $\bar{c}_n(\eta)$.

### 3.4 Implications for Sales

Having solved for the price index we can now write latent sales in market $n$, expression (4), for any firm with $\alpha_n(j)$, $\eta_n(j)$, and $c_n(j)$ as:

$$X^*_n(j) = \tilde{\alpha}_n(j) \left( \frac{m c_n(j)}{P_n} \right)^{1-\sigma} X_n$$  \hspace{1cm} (9)

$$= \alpha_n(j) \left[ 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)} \right] c_n(j)^{1-\sigma} \left( \frac{m}{P_n} \right)^{1-\sigma} X_n$$

$$= \alpha_n(j) \left[ 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)} \right] c_n(j)^{1-\sigma} \left( \frac{X_n}{\sigma E_n} \right)^{-(1-1/\bar{\sigma})} X_n$$

$$= \alpha_n(j) \left[ 1 - \left( \frac{c_n(j)}{\bar{c}_n(\eta_n(j))} \right)^{\lambda(\sigma-1)} \right] c_n(j)^{1-\sigma} \left( \frac{X_n}{\kappa_1^2 \Phi_n \sigma E_n} \right)^{1/\bar{\sigma}} \sigma E_n$$

and the entry condition (3) as:

$$c_n(j) \leq \bar{c}_n(\eta) = \eta_n^{1/(\sigma-1)} \left( \frac{X_n}{\kappa_1^2 \Phi_n \sigma E_n} \right)^{1/\bar{\sigma}}.$$  \hspace{1cm} (10)

Note that, once the price index is solved for, sales in a market increase less than in proportion to $X_n$. The reason is that a larger market attracts more entry, so that the price index is lower.

### 3.5 Implications for Firm Entry

The measure of suppliers who can pass that cost hurdle is $\mu_n(\bar{c}_n(\eta))$. Hence, integrating over the distribution of $\eta_n$, the measure of entrants into market $n$ is:

$$J_n = E[\mu_n(\bar{c}_n(\eta_n))] = \Phi_n E[\bar{c}_n(\eta_n)^{\theta}] = \frac{\kappa_2}{\kappa_1^2 \sigma \Phi_n \sigma E_n} \frac{X_n}{\kappa_1^2 \Phi_n \sigma E_n}$$  \hspace{1cm} (11)
where:

\[ \kappa_2 = E[\eta_n^\theta]. \]

Note that this measure rises in proportion to \( X_n \).

Suppliers to market \( n \) have heterogeneous costs. But, conditional on entry, suppliers from all countries have the same cost distribution in \( n \). To see why, consider good \( j \) in market \( n \) with entry shock \( \eta_n(j) = \eta_n \). For any cost \( c \) less than the entry threshold, the measure of suppliers with \( c_{ni}(j) \leq c \) among those with \( c_{ni}(j) \leq \bar{c}_n(\eta_n) \) is simply \( \mu_{ni}(c)/\mu_{ni}(\bar{c}_n(\eta_n)) = [c/\bar{c}_n(\eta_n)]^\theta \) for any \( c \leq \bar{c}_n(\eta_n) \). Hence for any \( \eta_n \) this proportion does not depend on source \( i \). Since we assume that the distribution of \( \eta_n \) is independent \( i \), different sources will have different measures of suppliers selling in market \( n \), but all who do sell will have the same distribution of unit costs.

Hence, given the constant markup over unit cost, suppliers from any source have the same distribution of prices in \( n \) and, hence, of sales. An implication is that the fraction of entrants into \( n \) coming from \( i \), \( \pi_{ni} \), is also the fraction of spending by country \( n \) on goods originating from country \( i \):

\[ \pi_{ni} = \frac{X_{ni}}{X_n}, \tag{12} \]

where \( X_{ni} \) is \( n \)'s purchases on goods originating from \( i \). This relationship gives us a connection between the cluster of parameters embedded in \( \pi_{ni} \) in (6) above and data on trade shares.

Combining (11) and (12), we get that the measure of firms from country \( i \) selling in country \( n \) is

\[ J_{ni} = \pi_{ni}J_n = \frac{\kappa_2 \pi_{ni}X_n}{\kappa_1 \sigma E_n}. \tag{13} \]

Hence the number of firms is proportional to trade share \( \pi_{ni} \) and to market size \( X_n \) given the
cost of entry $E_n$.

Our analysis is conditional on total expenditure of country $n$, $X_n$. Since the model makes a prediction about $J_n$, the measure of firms selling in market $n$ (including domestic suppliers), average sales per firm $\overline{X}_n$ is:

$$\overline{X}_n = \frac{X_n}{J_n} = \frac{\kappa_1}{\kappa_2} \sigma E_n.$$  

The model predicts that the distribution of sales is invariant to the location of the supplier, so $\overline{X}_n$ is the mean sales of entrants from any source $i$ selling in market $n$.

We manipulate expression (13) to give us an expression for the entry cost (multiplied by $\sigma$) in terms of the number of firms from any source $i$ divided by $i$’s market share:

$$\sigma E_n = \frac{\kappa_2 \pi_{ni} X_n}{\kappa_1 J_{ni}} = \frac{\kappa_2}{\kappa_1} \overline{X}_n.$$  

(14)

Since $\kappa_1$ and $\kappa_2$ are the same in any destination, this last result implies that variation in average sales per firm across markets reflects variation in the general component of the fixed cost of entry, a relationship we exploit in our estimation strategy.

3.6 Specification of the Distribution of the Entry and Sales Shock

We treat the sales and entry shocks $\alpha_n(j)$ and $\varepsilon_n(j)$ as jointly lognormally distributed. It’s more convenient for us to work with the equivalent assumption that $\ln \alpha_n(t)$ and $\ln \eta_n(j)$ are normally distributed with zero means and variances $\sigma_a^2$, $\sigma_h^2$, and correlation $\rho$. Under these assumptions we may write:

$$\kappa_1 = \left[ \frac{\tilde{\theta}}{\tilde{\theta} - 1} - \frac{\tilde{\theta}}{\tilde{\theta} + \lambda - 1} \right] \exp \left\{ \frac{\sigma_a^2 + 2\rho \sigma_a \sigma_h (\tilde{\theta} - 1) + \sigma_h^2 (\tilde{\theta} - 1)^2}{2} \right\}$$

17
and:

$$\kappa_2 = \exp \left\{ \frac{(\theta \sigma_h)^2}{2} \right\}.$$ 

Our bivariate normal assumption for \((\ln \alpha_n(j), \ln \eta_n(j))\) is equivalent to a bivariate normal assumption on \((\ln \alpha_n(j), \ln \varepsilon_n(j))\) with parameters \((0, 0, \sigma^2_a, \sigma^2_e, \rho')\). The mapping between the last two of these parameters and those that we estimate is

$$\sigma^2_e = \sigma^2_a - 2\rho \sigma_a \sigma_h + \sigma^2_h$$

and

$$\rho' = \frac{\sigma_a - \rho \sigma_h}{\sqrt{\sigma^2_a - 2\rho \sigma_a \sigma_h + \sigma^2_h}}$$

The boundary cases are particularly revealing. Suppose \(\rho = 1\). Then \(\sigma_e = \sigma_a - \sigma_h\) and \(\rho' = 1\) if \(\sigma_a > \sigma_h\) while \(\rho' = -1\) if \(\sigma_a < \sigma_h\). Our estimation results below put us at the opposite extreme with \(\rho = -1\). In this case \(\sigma_e = \sigma_a + \sigma_h\) and \(\rho' = 1\) so we can write

$$\varepsilon_n(j) = [\alpha_n(j)]^\zeta,$$

with \(\zeta = (\sigma_a + \sigma_h)/\sigma_a > 1\). The entry cost shock is simply an exaggerated version of the demand shock.

3.7 A Respecification for Simulation

We estimate the parameters of the model by the method of simulated moments. This method requires us to simulate a dataset of artificial French firms for a given set of parameter values. In our simulations we condition on the actual data on: (i) French market share in each of our 113 destinations, \(\pi_{nF}\), (ii) the number of French firms selling there \(J_{nF}\), and (iii) total
expenditure in each market $X_n$. Here we respecify the model in a way that makes simulation more convenient.

To isolate the heterogeneous component of unit costs we transform the efficiency draw of any potential producer in France as:

$$u(j) = T_F z_F(j)^{-\theta}.$$  

What is convenient about this reformulation is that the measure of firms with $U \leq u$ is the measure with $Z_F \geq (T/u)^{1/\theta}$ or simply $u$. Hence $u$ is uniform and its distribution doesn’t depend on any parameters.

We can write the unit cost of a potential French producer with efficiency $u(j)$ in market $n$ as:

$$c_n(j) = w_F d_{nF} = \left(\frac{u(j)}{\pi_{nF}}\right)^{1/\theta} \Phi_n^{-1/\theta}.$$  

We now reformulate the condition for firm $j$’s entry into market $n$ and sales in market $n$ conditional on entry in terms of its $u(j)$, $\alpha_n(j)$, and $\eta_n(j)$. The entry condition (10) is:

$$u(j) \leq \bar{u}_n(\eta) = \Lambda_n \alpha_n(j).$$  

where:

$$\Lambda_n = \frac{\pi_{nF} X_n}{\kappa_1 \sigma E_n}.$$  

which, using (14), becomes:

$$\Lambda_n = \frac{J_{nF}}{\kappa_2}.$$  

Hence $\Lambda_n$ can be constructed with the data on number of French firms selling in $n$ and the parameter values embedded in $\kappa_1$.  

19
Producer $j$’s unit cost relative to the threshold in any market $n$ can be written as:

$$\frac{c_n(j)}{\bar{c}_n(\eta)} = \left( \frac{u(j)}{\bar{u}_n(\eta)} \right)^{1/\theta} = \left( \frac{u(j)}{\Lambda_n \eta_n^{\bar{\theta}}} \right)^{1/\theta}.$$ 

Producer $j$ enters market $n$ if:

$$u(j) \leq \bar{u}_n(\eta) = \Lambda_n \eta_n^{\bar{\theta}}.$$ 

We can write the latent sales in market $n$ of a producer $j$ with $u(j)$, $\alpha_n(j)$, and $\eta_n(j)$ as:

$$X_{nF}^*(j) = \alpha_n(j) \left[ 1 - \left( \frac{u(j)}{\Lambda_n [\eta_n(j)]^{\bar{\theta}}} \right)^{\lambda/\bar{\theta}} \right] u(j)^{-1/\bar{\theta}} \Lambda_n^{1/\bar{\theta}} \sigma E_n.$$ 

(18)

Expressing the model in terms of (15) and (18) allows us to simulate French firms very straightforwardly. Having distilled the model to its statistical essence we now summarize all features of market $n$ that hold across sellers in $X_n$ and $\Lambda_n$. An individual producer $j$ is characterized by its efficiency draw $u(j)$ and its entry shocks $\eta_n(j)$ and sales shocks $\alpha_n(j)$ in each market $n$.

### 3.8 The Model and the Data

We now show how the model can deliver the features of the data about entry and sales described in Section 2. The basic feature of a producer that applies across markets is its underlying efficiency as reflected in $u(j)$, which enters both the entry condition and the expression for latent sales. A firm with a lower $u(j)$ can be expected both to enter more markets and to sell more where it enters.

**Entry.** Without the firm and market specific entry shock $\eta_n(j)$, (15) implies this efficiency variation is all that would matter for entry, dictating a deterministic ranking of destinations
with a less efficient firms (with a higher \( u(j) \)) selling to a subset of markets served by any more efficient firm. The firm and market specific sales shock \( \alpha_n(j) \) allows for independent variation in a firm’s sales in different markets.

To see how our model connects to entry we can rewrite the number of French firms selling in market \( n \), (13) as:

\[
J_{nF} = E_{\eta}[\pi_n(\eta)] = \Lambda_n \kappa_2.
\]

Invoking the definition of \( \Lambda_n \) we get that:

\[
\frac{J_{nF}}{\pi_{nF}} = \frac{\kappa_2}{\kappa_1} \frac{X_n}{\sigma E_n}, \tag{19}
\]

a relationship between the number of French firms selling to market \( n \) relative to French market share and the size of market \( n \) just like the one plotted in Figure 1b. Equivalently we can use expression (14) to give us an expression for \( \sigma E_n \) from average sales per French firm in market \( n \).

**Sales.** To get further insight into what our specification implies for the distribution of sales within a given market \( n \) it is convenient to note that, conditional on a firm’s entry, the term:

\[
v_n(j) = \frac{u(j)}{\Lambda_n \eta_n^{\theta}}
\]

is distributed uniformly on \([0, 1]\). We can then write sales (no longer latent) in market \( n \) as:

\[
X_{nF}(j) = \alpha_n(j) \left[ 1 - v_n(j)^{\lambda/\beta} \right] v_n(j)^{-1/\beta} \frac{\sigma E_n}{\eta_n(j)}
\]

\[
= \varepsilon_n(j) \left[ 1 - v_n(j)^{\lambda/\beta} \right] v_n(j)^{-1/\beta} \frac{\kappa_2}{\kappa_1} \frac{\pi_{nF} X_n}{J_{nF}}. \tag{21}
\]

Since the distributions of \( \varepsilon_n \) and \( v_n \) are identical across markets the distribution of sales in any market \( n \) is identical up to a scaling factor equal to \( \sigma E_n \). Hence we can generate the parallel
sales percentiles in Figure 4a (although we will miss the flatness of the 25th percentile). The variation introduced by $\varepsilon_n$ explains why the sales distribution in a market might inherit the lognormal characteristics apparent in Figures 3. As $\nu_n$ goes to one we can, with finite $\lambda$, generate sales that approach zero, allowing us to capture the curvature of sales distribution at the lower end, as observed in Figures 3. Finally the term $\nu_n^{-1/\tilde{\theta}}$ instills Pareto features into the distribution, a component of which carries across markets.

We can also look at the sales in France of French firms selling to any market $n$; as depicted in Figure 4b. To condition on these firms’ selling in market $n$ we couch the expression in terms of $\nu_n(j)$:

$$X_{FF}(j)|_n = \alpha_F(j) \left[ 1 - \nu_n(j)^{\lambda/\tilde{\theta}} \left( \frac{\Lambda_n}{\Lambda_F} \right)^{\lambda/\tilde{\theta}} \left( \frac{\eta_n(j)}{\eta_F(j)} \right)^{\lambda} \right] \nu_n(j)^{-1/\tilde{\theta}} \left( \frac{\Lambda_F}{\Lambda_n} \right)^{1/\tilde{\theta}} \frac{\sigma E_F}{\eta_n(j)}.$$

Invoking (17) we can rewrite this expression as:

$$X_{FF}(j)|_n = \frac{\alpha_F(j)}{\alpha_n(j)} \varepsilon_n(j) \left[ 1 - \nu_n(j)^{\lambda/\tilde{\theta}} \left( \frac{J_{nF}}{J_{FF}} \right)^{\lambda/\tilde{\theta}} \left( \frac{\eta_n(j)}{\eta_F(j)} \right)^{\lambda} \right] \nu_n(j)^{-1/\tilde{\theta}} \left( \frac{J_{nF}}{J_{FF}} \right)^{-1/\tilde{\theta}} \frac{\kappa_2}{\kappa_1} \pi_{FF}X_F.$$

Thus the distribution shifts with the number of firms having entered market $n$, as shown in Figure 4b. It follows that mean sales in France should shift with the number of firms selling to market $n$, the relationship depicted in Figure 2b, with a slope of $-1/\tilde{\theta}$ on a log scale.

Finally, we can look at the ratio of a firm’s sales in market $n$ to its sales in France (conditional on its selling to both markets) relative to mean sales in each market:

$$\frac{X_{nF}(j)}{X_{FF}(j)} = \frac{\alpha_n(j)}{\alpha_F(j)} \left[ 1 - \nu_n(j)^{\lambda/\tilde{\theta}} \left( \frac{J_{nF}}{J_{FF}} \right)^{\lambda/\tilde{\theta}} \left( \frac{\eta_n(j)}{\eta_F(j)} \right)^{\lambda} \right] \nu_n(j)^{-1/\tilde{\theta}} \left( \frac{J_{nF}}{J_{FF}} \right)^{1/\tilde{\theta}}.$$

To say more about the connection between the model and the data we need to go to the computer.
4 Estimation

To estimate the parameters of the model we proceed in three steps. We first simulate a set of artificial French firms given a particular set of parameter values $\Theta$. We then calculate a set of moments of the artificial firms and compare them with moments from the actual data. We then search for a $\Theta$ that brings the artificial moments close to the actual ones. The estimation algorithm is described in Appendix B.

4.1 Simulation

We create an artificial set of French firms that operate as the model tells them. Each firm has the option to sell in any of 113 markets.

We now refer to an artificial French firm by $j$ and the number of such firms as $S$ (which need not bear any relationship to the actual number of French firms). Our simulation proceeds in 3 stages as described:

1. As we search for parameters we want to hold fixed the realizations of the stochastic components of the model. Hence Stage 1 does not require any parameter values. It involves two steps:

   (a) Draw realizations $v(j)$’s independently from $U[0, 1]$, for $j = 1, \ldots, S$ and put them aside to construct the $v(j)$ in Stage 3.

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

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

   and put them aside to construct the $a_n(j)$ and $h_n(j)$ in Stage 3.
2. Stage 2 requires data for each destination \( n \) on \( X_n \) and French market share \( \pi_{nF} \) as well as a set of parameters \( \Theta \). It involves two steps:

(a) Calculate:

\[
\kappa_1 = \left[ \frac{\tilde{\theta}}{\tilde{\theta} - 1} - \frac{\tilde{\theta}}{\tilde{\theta} + \lambda - 1} \right] \exp \left\{ \frac{\sigma_A^2 + 2\rho \sigma_A \sigma_H (\tilde{\theta} - 1) + \sigma_H^2 (\tilde{\theta} - 1)^2}{2} \right\}.
\]

(b) Calculate:

\[
\sigma E_n = \frac{\kappa_2 \pi_{nF} X_n}{\kappa_1 J_{nF}}
\]

and:

\[
\Lambda_n = \frac{J_{nF}}{\kappa_2}.
\]

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 1b 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:

\[
\pi_n(j) = \frac{J_{nF}}{\kappa_2} \eta_n(j) \tilde{\theta}.
\]

(c) We don’t want to waste time drawing potential producers who end up not selling anywhere. We construct our artificial firms so that they necessarily sell in France.
Hence, we want \( u(j) \leq \bar{\pi}_F(j) \), with \( \bar{\pi}(j) = \bar{\pi}_F(j) \), for each draw \( j \). In other words, \( u(j) \) should be a realization from \( U[0, \bar{\pi}_F(j)] \). Therefore we construct:

\[
u(j) = \nu(j)\bar{\pi}_F(j).
\]

using the \( \nu(j) \) from Stage 1.

(d) To correct for the fact that a potential producer with a low \( \bar{\pi}_F(j) \) is less common our simulated French firm gets a weight \( \omega(j) = \bar{\pi}_F(j) \).

(e) Calculate \( S_{nF}(j) \), where \( S_{nF}(j) = 1 \) if firm \( j \) enters market \( n \) and 0, otherwise as determined by the entry hurdles:

\[
S_{nF}(j) = \begin{cases} 
1 & \text{if } u(j) \leq \bar{\pi}_n(j) \\
0 & \text{otherwise},
\end{cases}
\]

where of course \( S_{FF}(j) \) necessarily equals 1.

(f) Wherever \( S_n(j) = 1 \) calculate sales as:

\[
X_{nF}(j) = \alpha_n(j) \left[ 1 - \left( \frac{u(j)}{\bar{u}_n(j)} \right)^{\lambda/\theta} \right] u(j)^{-1/\theta} \Lambda_n^{\lambda/\theta} \sigma E_n
\]

\[
= \alpha_n(j) \left[ 1 - \left( \frac{u(j)}{\bar{u}_n(j)} \right)^{\lambda/\theta} \right] u(j)^{-1/\theta} \left( \frac{J_{nF}}{\kappa_2} \right)^{1/\theta} \frac{\kappa_2 \pi_{nF} X_n}{\kappa_1 \bar{J}_{nF}}.
\]

Following this procedure we simulate the behavior of \( S \) firms. We know three things about each firm: where it sells, \( S_{nF}(j) \), how much it sells where it does, \( X_{nF}(j) \), and its weight \( \omega(j) \). From these we can calculate any moment that could have been calculated on the actual French data.

### 4.2 Moments

We record:
1. The number of simulated firms entering each market as:

\[ J_{nF} = \frac{1}{S} \sum_{j=1}^{S} \omega(j) S_{nF}(j) \]

giving us 113 moments. This formula repeats for all the other moments with appropriate changes in the indicator \( S_n(j) \).

2. The number of simulated firms that sell in \( n \) but sell below the \( q \)th percentile there for \( q = 0.5, 0.75, \) and 0.95, giving us \( 3 \times 113 \) moments. (Here and in the following two sets of moments, the percentiles are from the actual data).

3. The number of simulated firms that sell in \( n \) and sell below the \( q \)th percentile in France, among all French firms selling in \( n \), for \( q = 0.5, 0.75, \) and 0.95, again giving us \( 3 \times 113 \) moments

4. The number of simulated firms that sell in \( n \) whose ratio of sales in \( n \) to sales in France is below the 75th percentile of French firms selling in \( n \), yielding 113 moments.

5. The number of simulated firms selling to each possible combination of the seven most popular export destinations, giving us \( 2^7 \) moments.

Our objective function is the weighted squared deviation of each simulated moment from the actual one. Note that each moment is a count of firms achieving a relatively rare event (except for numbers selling in France itself). Hence we treat them as realizations of Poisson random variables and weight by the inverse of the actual moment, an estimator of the Poisson parameter and hence of the variance.
5 Results

The best fit is achieved at the following parameter values (with bootstrapped standard errors in parentheses):

\[
\begin{array}{cccccc}
\tilde{\theta} & \lambda & \sigma_a & \sigma_b & \rho \\
2.2305 & 4.0808 & 1.7684 & .3861 & -1 \\
(0.09) & (0.45) & (0.013) & (0.015) \\
\end{array}
\]

using a simulated sample of 500,000 firms.

To see how well the model captures the features of the data described in Section 2 we reproduce pairs of figures (actual and simulated), with common axes within a pair to facilitate comparison. Some of the figures are closely related to the moments that we fit, while others are quite distinct from those moments.

Figures 1c (actual and simulated) show how well the estimated model captures French firm entry into multiple markets. This moment was not used in the estimation, yet the estimated model captures it quite closely. The main deviation is that the model predicts too many firms selling to over 50 markets or more. Figures 2d (actual and simulated) explores the selection effect based on entry into multiple markets. Here the model picks up the distinctive log-linearity between mean sales in France and the number of firms entering multiple markets, but the slope is decidedly flatter than the one in the data. In contrast, Figures 2b (actual and simulated) shows that the model captures the corresponding relationship based on entry into unpopular markets. The reason is likely due to the fact that this moment is closely related to ones used in the estimation.

Likewise, Figures 4c (actual and simulated) shows that the model captures very well the spread of the sales distribution in each market (each figure shows the median and 95th percentiles). Finally, Figures 4d (actual and simulated) shows that the model roughly captures
the much noisier moment related to market-level export intensity variation.

Aside from these Figures we report the number of firms selling to the 7 most popular export destinations in order of their popularity (that is, obeying a hierarchy), both in the actual and simulated data. That is, we report the number of firms selling to Belgium and no other top 7, to Belgium and Germany and no other top 7, etc. with the following results:

<table>
<thead>
<tr>
<th>Country</th>
<th>Actual Number</th>
<th>Simulated Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>3988</td>
<td>4421</td>
</tr>
<tr>
<td>B-DE</td>
<td>863</td>
<td>881</td>
</tr>
<tr>
<td>B-DE-CH</td>
<td>579</td>
<td>412</td>
</tr>
<tr>
<td>B-DE-CH-I</td>
<td>330</td>
<td>249</td>
</tr>
<tr>
<td>B-DE-CH-I-UK</td>
<td>313</td>
<td>257</td>
</tr>
<tr>
<td>B-DE-CH-I-UK-NL</td>
<td>781</td>
<td>454</td>
</tr>
<tr>
<td>B-DE-CH-I-UK-NL-USA</td>
<td>2406</td>
<td>2665</td>
</tr>
<tr>
<td>Total</td>
<td>9260</td>
<td>9339</td>
</tr>
</tbody>
</table>

In our estimated data set 27.2 percent of exporters adhere to hierarchies compared with 29.4 percent in the data and the 13.5 percent implied by simply predicting on the basis of the marginal probabilities.

We now turn to some implications of our parameter estimates.

### 5.1 Sources of Variation

Our model explains variation across firms in their entry and sales as a reflection of differences in their underlying efficiency, which applies across all markets, and idiosyncratic entry and sales shocks in individual markets. We ask how much of the variation in entry and sales can be explained by the universal rather than the idiosyncratic components.
5.1.1 Variation in Entry

We first calculate the fraction of the variance of entry in each market that can be explained by the cost draw $u$ alone, which is the same across markets. Since we do all these calculations only for firms that sell in France it is convenient to condition on a firm’s $\eta_F$. From (15) above the probability that such a firm will sell in market $n$ is:

$$q_n(\eta_F) = \frac{\Lambda_n \kappa_2}{\Lambda_F \eta_F^p}.$$ 

Hence the variance of entry not conditioning on either $\eta_n(j)$ or $u(j)$ is:

$$V_n^U(\eta_F) = q_n(\eta_F)[1 - q_n(\eta_F)].$$ (22)

Conditional on its $u(j)$ a firm enters market $n$ if:

$$\eta_n \geq (u/\Lambda_n)^{1/\bar{\theta}}.$$ 

Since $\eta_n$ is lognormally distributed with mean 0 and variance $\sigma_h$ the probability that this condition is satisfied is:

$$q_n(u, \eta_F) = 1 - \Phi \ln((u/\Lambda_n)/(\bar{\theta} \sigma_h))$$

where $\Phi$ is the standard normal cumulative density. The variance conditional on $u(j)$ is:

$$V_n^C(u, \eta_F) = q_n(u, \eta_F)[1 - q_n(u, \eta_F)].$$

For each destination $n$ we repeatedly draw $\eta_F(j)$ and $u(j)$ in order to calculate:

$$R^E_n(u) = \frac{E_{\eta_F,u} [V_n^C(u, \eta_F)]}{E_{\eta_F} [V_n^U(\eta_F)]}.$$ 

Simulating this term for 200,000 firms selling in our 113 markets delivered an average value of $R_n^E$ of .54 with a standard deviation of .021. Hence we can attribute around 46 percent of
the variation in entry in a market to the core efficiency of the firm rather than their draw of 
\( \eta \) in that market.\(^5\)

### 5.1.2 Variation in Sales

Looking at the firms that enter a particular market, how much does the variation in \( u \) explain the variation in their sales there. Applying expression (18) to a firm \( j \) that actually enters market \( n \), the log of its sales there are:

\[
\ln X_{nF}(j) = \ln \alpha_{n}(j) + \ln \left[ 1 - \left( \frac{u(j)}{\Lambda_n \eta_n(j)^{\theta}} \right)^{\lambda/\theta} \right] - \frac{1}{\theta} \ln u(j) + \ln \left( \Lambda_n^{1/\theta} \sigma E_n \right).
\]

where we have divided sales into four components. Component 4 is common to all firms selling in market \( n \) so does not contribute to variation in sales there. The first component involves firm \( j \)'s idiosyncratic sales shock in market \( n \) while component 3 involves its efficiency shock that applies across all markets. Complicating matters is component 2, which involves both firm \( j \)'s idiosyncratic entry shock in market \( j \eta_n(j) \) and its overall efficiency shock \( u(j) \). We deal with this issue by first asking how much of the variation in \( \ln X_n(j) \) is due to variation in component 3 and then in the variation in components 2 and 3 together.

We simulate sales of 50,000 firms across our 113 markets, and divide the contribution of each component to its sales in each market where it sells. We find that component 3 contributes a .069 share to the variation in \( \ln X_{nF}(j) \), averaging across markets (with a standard error of .0095). Again averaging across markets, the share of components 2 and 3 together in the

---

\(^5\) Not conditioning on \( u \) the probability \( q_n \) that a firm sells in any market \( n \) other than France is small. It is straightforward to show that taking the limit as \( q_n \to 0 \) the term \( R_n^{E} \) is independent of \( n \). Hence the systematic variation in \( R_n^{E} \) across markets is small.
variation of $\ln X_{nF}(j)$ is .26 (with a standard deviation of .043).\footnote{The presence of $\Lambda_n$ in component 2 means that there is variation in the contribution of each component varies across markets, but our simulation indicates that the differences are very small.}

Together these results indicate that the general efficiency of a firm is very important in explaining its entry into different markets, but makes a much smaller contribution to the variation in the sales of firms actually selling in a market. An explanation is that in order to enter a market a firm has already to have a low value of $u$. Hence the various sellers present in a market already have low values of $u$, so that differences among them in their sales are dominated by their market-specific sales shock $\alpha_n(j)$.

This finding does, of course, depend on our parameter estimates. A higher value of $\tilde{\theta}$ (implying, given $\sigma$, less heterogeneity in efficiency) or a higher value of $\sigma_h$, implying more variable entry cost shocks, would lead us to attribute more of the variation in entry to destination-specific factors. Similarly, a higher value of $\tilde{\theta}$ or a higher value of either $\sigma_a$ or $\sigma_h$ would lead us to attribute more to sales and entry shocks in individual markets.

5.2 Productivity

It has been commonly observed that producers who export exhibit higher productivity (according to various measures) than those that don’t.\footnote{See, for example, Bernard and Jensen (19XX), Lach, Roberts, and Tybout (1997), and BEJK (2003).} The same is true for our exporters here, at least as measured according to value added per worker. We have not exploited this feature of the data in our method of simulated moments estimation. Instead we take our parameter estimates and calculate what they imply for the measured productivity for French firms. We then compare these implications with what we see in the data.
We calculate intermediate purchases by firm \( j \) as:

\[
CI(j) = (1 - \beta^{VAR})\overline{m}^{-1}X(j) + (1 - \beta^{FIX})E_F(j)
\]

where \( \beta^V \) is the share of factor costs in variable costs and \( \beta^F \) is the share of factor costs in fixed costs. We then define value added as:

\[
VA(j) = X(j) - CI(j)
\]

where

\[
X(j) = \sum_{n=1}^{N} X_n(j).
\]

We conclude with value added per unit of factor cost as:

\[
v(j) = \frac{VA(j)}{\beta^{VAR}X(j) + \beta^{FIX}E_F(j)}
\]

\[
= \frac{\left[1 - (1 - \beta^{VAR})\overline{m}^{-1}\right]X(j) - (1 - \beta^{FIX})E_F(j)}{\beta^{VAR}\overline{m}^{-1}X(j) + \beta^{FIX}E_F(j)}
\]

\[
= \frac{[\overline{m} - (1 - \beta^{VAR})] - \overline{m}(1 - \beta^{FIX})[E_F(j)/X(j)]}{\beta^{VAR} + \overline{m}\beta^{FIX}[E_F(j)/X(j)]}.
\]

We are not taking into account fixed costs incurred in entering export markets, because they are not in our data.

We focus on a special case in which \( \beta^{FIX} = 1 \) in which case:

\[
v(j) = \frac{[\overline{m} - (1 - \beta^{VAR})]}{\beta^{VAR} + \overline{m} [E_F(j)/X(j)]}.
\]

We calculate this statistic for our French firms and see how it varies according to the number of destinations served. We cannot derive its analytic distribution with multiple markets.

Consider the case of a single market, however, and, for simplicity, set \( \beta^{VAR} = 1 \) as well. The measure of firms selling \( X \geq x \) is given by:

\[
\mu_X[X \geq x] = \left(\frac{x}{\sigma E}\right)^{-\theta}.
\]

32
To translate the distribution of sales into a distribution of productivities we note that the measure of firms with productivity $V \geq v$ is the measure of firms with sales $X \geq \bar{m}E/[(\bar{m}/v) - 1]$

or:

$$\mu_X \left[ X \geq \frac{\bar{m}E}{(\bar{m}/v) - 1} \right] = \left( \frac{\bar{m}}{\sigma ((\bar{m}/v) - 1)} \right)^{-\tilde{\theta}}$$

or, since $\sigma = \bar{m}/(\bar{m} - 1)$,

$$\mu_V[V \geq v] = \left( \frac{(\bar{m}/v) - 1}{\bar{m} - 1} \right)^{\tilde{\theta}}$$

where $v \in [1, \bar{m}]$. This distribution is related to the truncated Pareto.

In general, with $\beta^{FIX} = 1$ and $\beta^{VAR} \in [0, 1)$ we get that the measure of firms with productivity of at least $v$ is:

$$\mu_V[V \geq v] = \left( \frac{\bar{m} + \beta^{VAR} - 1}{v \bar{m} - 1} - \beta^{VAR} \right)^{\tilde{\theta}}.$$  

Figures 6a (actual and simulated) shows how the model does in predicting the rise in productivity corresponding to entry into multiple markets. The model fails to reproduce the spike in the productivity of those firms that export to 40-50 countries.

### 5.3 Out of Sample

We conclude with a look at how well the model holds up out of sample. To do so we examine the cross-section for 1992 (remember that our analysis up to this point has used only data for 1986). Figure 5a shows that the most striking relationship we have uncovered, that between entry into unpopular markets and sales in France, continues to hold up. Figure 5b examines whether the changes in the number of French firms entering a market correlates as it should with changes in mean sales in France of these firms. The predicted relationship shows up clearly, knocked slightly out of whack by the data point for Iraq.
6 Conclusion

To be added
References


A Data Construction

Up to 1992 all movements of goods entering or leaving France were declared to French customs either by their owners or by authorized customs commissioners. These declarations constitute the basis of all French trade statistics. Each movement generates a record. Each record contains the firm identifier, the SIREN, the country of origin (for imports) or destination (for exports), a product identifier (a 6-digit classification), and a date. All records are aggregated first at the monthly level. In the analysis files accessible to researchers, these records are further aggregated by year and by 3-digit product (NAP 100 classification, the equivalent of the 3-digit SIC code). Therefore, each observation is identified by a SIREN, a NAP code, a country code, an import or export code, and a year. In our analysis, we restrict attention to exporting firms in the manufacturing sector in year 1986 and in year 1992. Hence, we aggregate across manufacturing products exported. We can thus measure each firm’s amount of total exports in years 1986 and 1992 by country of destination. Transactions are recorded in French Francs and reflect the amount received by the firm (i.e. including discounts, rebates, etc.). Even though our file is exhaustive, i.e., all exported goods are present, direct aggregation of all movements may differ from published trade statistics, the second being based on list prices and thus exclude rebates.

We match this file with the Base d’Analyse Longitudinale, Système Unifié de Statistiques d’Entreprises (BAL-SUSE) database, which provides firm-level information. The BAL-SUSE database is constructed from the mandatory reports of French firms to the fiscal administration. These reports are then transmitted to INSEE where the data are validated. It includes all firms subject to the “Bénéfices Industriels et Commerciaux” regime, a fiscal regime manda-
tory for all manufacturing firms with a turnover above 3,000,000FF in 1990 (1,000,000FF in the service sector). In 1990, these firms comprised more than 60% of the total number of firms in France while their turnover comprised more than 94% of total turnover of firms in France. Hence, the BAL-SUSE is representative of French enterprises in all sectors except the public sector.

From this source, we gather balance sheet information (total sales, total labor costs, total wage-bill, sales, value-added, total employment). Matching the Customs database and the BAL-SUSE database leaves us 229,900 firms in manufacturing (excluding construction, mining and oil industries) in 1986 with valid information on sales and exports. In 1992, the equivalent number is 217,346. In 1986, 34,035 firms export to at least one country; among them 17,699 export to Belgium, the most popular destination. To match our data with aggregate trade and production data, we restrict attention to 113 countries (including France).

A  Estimation Algorithm

The algorithm we use to fit theoretical moments to their empirical counterparts is simulated annealing. We rely on a version specifically developed for Gauss and available on the web from William Goffe (Simann). Goffe, Ferrier, and Rogers (1994) describe the algorithm. Simulated annealing, in contrast with other optimization algorithms (Newton-Raphson, for instance), explores the entire surface and moves both uphill and downhill to optimize the function. It is therefore largely independent of starting values. Because it goes both downhill and uphill, it escapes local maxima. Finally, the function to optimize does not need to have stringent properties; differentiability and continuity, for instance, are not needed. The version
developed by these authors, and implemented in Simann, possesses some features that make it more efficient (in particular, less time-consuming) than previous implementations of simulated annealing. The program includes a precise explanation of the various parameters that must be set in advance. It also suggests reasonable starting values. The program, as well as the starting values, is available from the authors. Optimizing our admittedly complex function, with 5 parameters, on a standard personal computer takes around one week.

We use bootstrap to compute standard errors for our parameters. We follow the bootstrap procedure suggested by Horowitz (2001) closely. More precisely, we use the bootstrap with recentering as suggested when using a method of moments estimation strategy (Horowitz, 2001, subsection 3.7, pages 3186-3187). Because each bootstrap repetition requires one week for estimation, we used only 10 repetitions. The small number is unlikely to have any effect given the concentration of the bootstrap estimates around the estimated values.
Table 1: Exports to the Seven Most Popular Destinations

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of French Exporters</th>
<th>Marginal Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium-Luxembourg</td>
<td>17,699</td>
<td>0.59</td>
</tr>
<tr>
<td>Germany</td>
<td>14,579</td>
<td>0.49</td>
</tr>
<tr>
<td>Switzerland</td>
<td>14,173</td>
<td>0.47</td>
</tr>
<tr>
<td>Italy</td>
<td>10,643</td>
<td>0.36</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>9,752</td>
<td>0.33</td>
</tr>
<tr>
<td>Netherlands</td>
<td>8,294</td>
<td>0.28</td>
</tr>
<tr>
<td>United States</td>
<td>7,608</td>
<td>0.25</td>
</tr>
<tr>
<td>Total firms exporting to at least one of the top 7</td>
<td>29,929</td>
<td>1.00</td>
</tr>
<tr>
<td>Total firms exporting</td>
<td>34,035</td>
<td></td>
</tr>
</tbody>
</table>

Source: Customs data, year 1986. Marginal probability is the ratio of the number exporting to the destination relative to the number exporting to at least one of these seven countries.
Table 2: Do Firms Obey Country Hierarchies?

<table>
<thead>
<tr>
<th>Exporting to:</th>
<th>Observed Number of Firms</th>
<th>Under Independence</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Predicted Probability</td>
<td>Number of Firms</td>
</tr>
<tr>
<td>1000000</td>
<td>3988</td>
<td>0.037</td>
<td>1119</td>
</tr>
<tr>
<td>1100000</td>
<td>863</td>
<td>0.036</td>
<td>1063</td>
</tr>
<tr>
<td>1110000</td>
<td>579</td>
<td>0.032</td>
<td>956</td>
</tr>
<tr>
<td>1111000</td>
<td>330</td>
<td>0.018</td>
<td>528</td>
</tr>
<tr>
<td>1111100</td>
<td>313</td>
<td>0.009</td>
<td>255</td>
</tr>
<tr>
<td>1111110</td>
<td>781</td>
<td>0.003</td>
<td>98</td>
</tr>
<tr>
<td>1111111</td>
<td>2406</td>
<td>0.001</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>9260</td>
<td>0.135</td>
<td>4052</td>
</tr>
</tbody>
</table>

Source: Customs data, year 1986. All firms export at least to one of these 7 countries: Belgium, Germany, Switzerland, Italy, United Kingdom, Netherlands, United States, in this order. The string of 0s and 1s in column labelled "Exporting to" refer to this order: 1000000 means that these firms only export to Belgium; 1111111 means that these firms export to all seven countries. We ignore in this table export behavior outside these 7 countries.
Figure 1a: Entry of French Firms and Market Size
Figure 2a:
Size in France and Popularity of the Market
Figure 2b: Size in France and Number Entering the Market

average sales in France ($ millions) vs. # firms entering the market
Figure 2c: Size in France and Minimum Number of Markets Penetrated
Figure 2d:
Size in France and Number Entering Multiple Markets

average sale in France ($ millions)
firms selling to k or more markets
Figure 3a: Distribution of French Sales in France
Figure 3b: Distribution of French Exports to Ireland
Figure 3c: Distribution of French Exports to the United States
Figure 4a: Distribution of French Firm Sales, by Market
Figure 4b: Distribution of Sales in France, by Market Penetrated

[Graph showing the distribution of sales in France, with sales percentiles (25, 50, 75, 95) in dollars and the number of firms selling in the market.]
Figure 4c: Distribution of French Firm Sales, by Market

sales percentiles (25, 50, 75, 95) by market ($ millions)

mean sales by market ($ millions)
Figure 4d: Distribution of Export Intensity, by Market

percentiles (25, 50, 75, 95) of exports to sales in France

mean sales by market ($ millions)
Figure 5a:
Size in France and Number Entering the Market, in 1992
Figure 5b: Size in France and in Number Entering the Market Changes between 1986 and 1992
Figure 1c (Actual): Number of Firms Entering Multiple Markets

- **X-axis:** Minimum number of markets penetrated
- **Y-axis:** Number of firms selling to multiple markets

The graph shows a decreasing trend in the number of firms selling to multiple markets as the minimum number of markets penetrated increases.
Figure 1c (Simulated): Number of Firms Entering Multiple Markets

- X-axis: Minimum # of markets penetrated
- Y-axis: # firms selling to multiple markets

The graph shows a decreasing trend as the minimum number of markets penetrated increases.
Figure 2b (Actual): Size in France and Number Entering the Market
Figure 2b (Simulated): Size in France and Number Entering the Market
Figure 2d (Actual):
Size in France and Number Entering Multiple Markets

average sale in France ($ millions)
firms selling to k or more markets
Figure 2d (Simulated): Size in France and Number Entering Multiple Markets
Figure 4c (Actual):
Distribution of French Firm Sales, by Market

sales percentiles (25, 50, 75, 95) by market ($ millions)

mean sales by market ($ millions)
Figure 4c (Simulated): Distribution of French Firm Sales, by Market
Figure 4d (Actual): Distribution of Export Intensity, by Market

percentiles (25, 50, 75, 95) of exports to sales in France

mean sales by market ($ millions)
Figure 4d (Simulated): Distribution of Export Intensity, by Market

percentiles (25, 50, 75, 95) of exports to sales in France

mean sales by market ($ millions)
Figure 6a (Actual): Productivity and Entering Multiple Markets
Figure 6a (Simulated): Productivity and Entering Multiple Markets