A Search and Learning Model of Export Dynamics*

Jonathan Eaton[a,e], Marcela Eslava[b], David Jinkins[c],
C. J. Krizan[d], and James Tybout[c,e]

May 22, 2020

Abstract

Customs record data reveal a number of patterns in relationships Colombian firms have with their U.S. buyers. We interpret these patterns in terms of a continuous-time model in which heterogeneous sellers search for buyers in a market. Success in selling to a buyer reveals information to the seller about the appeal of her product in the market, affecting her incentive to search for more buyers. Fit using the method of simulated moments, the model replicates key patterns in the customs records and allows us quantify several types of trade costs, including the search costs of identifying potential clients and the costs of maintaining business relationships with existing clients. It also allows us to estimate the effect of previous exporting activity on the costs of meeting new clients, and to characterize the cumulative effects of learning on firms' search intensities. Finally, we use our fitted model to explore the effects of these trade costs and learning effects on aggregate export dynamics.

*We gratefully acknowledge support from the National Science Foundation (Grant SES-0922358), the United States Census Bureau, and Banco de la República de Colombia. We also thank Monica Hernández, Gustavo Caballero and Camilo Acosta for excellent research assistance, as well as Enrique Montes for expert data advice. This paper was written in part by Census Bureau staff. It has undergone a more limited review than official Census Bureau publications. All results were reviewed to ensure confidentiality. Any views, findings and opinions in the paper reflect the views of the authors and do not reflect the views of the U.S. Census Bureau.

1 Introduction

Research on exporting has been digging deeper into microeconomics data to understand the barriers that producers face in entering foreign markets and their implications for export dynamics. Firm-level datasets have provided insights first into the costs of exporting at all, and then, as data became available, to penetrating individual markets (Das et. al, 2007; Arkolakis, 2010; Eaton et al., 2011; Morales et al., 2019). We take this analysis one step forward by examining exporters' relationships with individual buyers in a market, both descriptively and through the lens of a dynamic model. In doing so we quantify exporting costs, link them to particular types of information frictions, and explore their dynamic implications.

The first type of friction in our model is standard: exporters must engage in costly search to identify potential clients abroad. Since search costs are convex in match rates, this forces firms to gradually build their portfolio of foreign buyers. The second type of friction arises from sellers' limited knowledge of foreign buyers' tastes. Potential exporters are unsure about the appeal of their products in foreign markets, but they gradually learn about this as they meet potential customers. Therefore, exporters with appealing products intensify their marketing efforts as they receive positive reinforcement. The final type of friction has to do with buyers learning about sellers. Exporters that have already established a large number of business relationships are relatively visible to other buyers, so for a given level of spending on search, they meet relatively more potential customers. This means that new exporters add clients relatively slowly, and it makes it profitable for exporters with large client bases to replace the many business relationships they lose to attrition.

We base our analysis on the cross-sectional and temporal variation in shipment-level customs records from the United States, which report importer and exporter identifiers. We begin by summarizing the main patterns in these data that we want our model to explain, including the dynamics of seller matching patterns and the life-cycle revenue trajectories generated by indiviudal matches. Then we develop a dynamic search and matching model in which exporters expand by adding to their client base in destination markets. Finally, we fit this model to our customs data and use it to make inferences about the importance of several types of infomation frictions.

In addition to quantifying the value of information, this last stage involves several counterfactual experiments. These are designed to answer several questions: How much intangible capital is tied to foreign business connections, and how does this capital vary across different types of exports? Controlling for exporters' productivity and product appeal, how big a role does market visability play in driving success? And finally, to what extent do information frictions cause exporting patterns to deviate from a sequence of static equilibria? **(JT: list**

of questions to be updated; summary of results here)

1.1 Relation to literature

(JT: section needs to be updated) While we look at the evolution of firms' sales in a particular market, our analysis is related to the literature on the dynamics of firm size in general. The model explains the size distribution of firm sales through two interacting mechanisms. One, as in Melitz (2003), Bernard et al. (2003), Luttmer (2007), and Irarrazabal and Opromolla (2006), is firm efficiency: More efficient firms sell more to a given set of buyers by having a lower price or a higher quality product. A second is that some firms have larger networks of buyers than others, as in Jackson and Rogers (2007) or Chaney (2011).

Investments in building a client base constitute a type of sunk cost, so our model also relates to the export hysteresis literature (Dixit, 1989; Baldwin and Krugman, 1989; Das, et al., 2007; Alessandria and Choi, 2007; Alessandria et al., 2010), where firms pay a one-shot start-up cost to break into new markets. But unlike these formulations, our sunk costs are incurred on the client margin rather than the country margin, and they pay off in terms of market knowledge and reputation as well as revenue streams. These features of our model allow us to explain why new exporters who don't exit tend to rapidly expand, and why established exporters' sales are relatively stable. They also explain why many firms export for short periods on a very small scale.

Our formulation is also related to the two-period learning models developed by Rauch and Watson (2003) and Albornoz et al (2012). In the former, importers experiment with foreign suppliers by placing trial orders with them, and they gain access to a supplier network if they establish a successful business relationship. In the latter, firms choose to experiment in markets with low entry costs in order to learn about their product's appeal elsewhere. Like our model, these formulations provide interpretations for the fact that when new exporters survive, their exports tend to grow rapidly.[1]

Finally, in allowing firms to attract more buyers by incurring greater costs, our analysis relates to Drozd and Nozal (2012) and Arkolakis (2009, 2010). By positing that firms face marketing costs that are convex in the number of foreign clients they service, Arkolakis also accounts for small-scale exporters and the age-dependence of export growth rates. However, since all exporting relationships last a single period in his models and learning is absent, Arkolakis's models do not explain the irreversibilities observed in firms' exporting behavior, nor do they speak to the duration of matches.

[1]Ruhl and Willis (2008) also note this pattern in plant-level export data and show that market entry costs are insufficient to explain it.

2 Firm-Level Trade: Transaction Level Evidence

2.1 Data

The empirical motivation for our model comes from a comprehensive data set that describes all imports by buyers in the United States from Colombian exporters (as well as other origins) during the period 1992-2009. The source is the U.S. Census Bureau's Longitudinal Foreign Trade Transactions Database (LFTTD). Each record includes a date, the US dollar value of the product shipped, a 6-digit harmonized system product code, a quantity index, and, critically, ID codes for both sellers and buyers. These IDs allow us to identify the formation and dissolution of business relationships between individual buyers in the U.S. and sellers in Colombia, hereafter referred to as "matches."[2]

To identify foreign exporters, the U.S. import transactions records include a manufacturer's identification code.[3] This field is an amalgamation of the manufacturer's country, company name, street address, and city. Anecdotal information from customs brokers indicates that commonly used software constructs it automatically as the name and address information is entered in other fields. So this variable is sensitive to differences in the way exporters' names and addresses are recorded as they pass through customs, and shipments from the same exporter can appear to originate from distinct Colombian firms. To gauge the importance of this problem, we have conducted various checks on the matches that are based on this variable; these are explained in Appendix 2.

We limit our analysis to transactions between non-affiliated trade partners, and we consider only imports of manufactured goods. The latter restriction notably excludes oil and coffee exports, which constitute the bulk of trade between the two countries and are dominated by a few Colombian sellers.[4] Our final data set of manufacturing transactions spans the years 1992-2009. It contains 26,625 unique Colombian exporters, 12,921 unique U.S. importers, and 42,767 unique trading pairs. Value data have been deflated to 1992 prices using the U.S. CPI. Since we exclude a number of large HS codes from our data, as well as affiliated trade, and because we also lose information due to disclosure restrictions, the total value covered by our data is not comparable to total Colombian exports to the U.S. Table

[2] There are two ways to track U.S. importers in the LFTTD: Employment Identification Numbers (EINs) and the firm identifiers in the Longitudinal Business Database ("alphas"). Though an EIN does not necessarily identify a complete firm, it is unique to a firm, and there is one associated with every import transaction. Alphas map to entire firms, but the match rate between trade transactions and alphas is only about 80 percent (Bernard, Redding, and Schott, 2009). To maximize the coverage of our sample, we use Employment Identification Numbers (EIN) to identify U.S. buyers.

[3] This variable is based on Block 13 of CBP form 7501, the import declaration form and customs brokers are required to input the data.

[4] Colombian commercialization of coffee is centralized to an important degree by the National Federation of Coffee Growers. A few players also dominate oil exports.

?? in Appendix **??** compares patterns in our sample to patterns in official aggregates from both the U.S. and Colombia.

In addition to U.S. customs records, we use establishment level survey data from Colombia's national statistics agency (Departmento Administrativo Nacional de Estadistica, or DANE). These data provide annual information on the sales volumes, exports, and other characteristics of all Colombian manufacturing plants with at least 10 workers. Because they have been widely analyzed, we do not discuss summary statistics for this data set herein. Later, however, when estimating our search and learning model, we use such statistics to characterize the size distribution of Colombian firms, the fraction of Colombian plants that export and, among these firms, the relationship between exports and domestic sales.

2.2 Exporter cohort maturation

Following Brooks (2006) and Eaton et al. (2008), Table 1 summarizes the typical cohort maturation process for Colombian exporters of manufactured goods to the United States. It is based on observed evolution patterns among cohorts of firms that entered the market each year between 1993 and 1999, and it exploits U.S. customs records from 1992 through 2009.

To interpret the figures in this Table, imagine for a moment that they describe a particular cohort, say, those firms that first entered the U.S. market in 1993. Then the second row of the Table would imply that only 29 percent of these firms continued exporting though 1994 (column 1), yet these survivors generated 11 percent more export revenue in 1994 than the entire cohort did in 1993 (column 2) because sales per surviving cohort member were 3.77 times as large in 1994 as sales per cohort member in 1993 (column 3). Other rows would have analogous interpretations, each normalized relative to the cohort's entry year.

The actual interpretation for Table 1 differs from this one only in that it is an average of all of the cohort-specific tables we can construct using cohorts observed for at least 10 years.[5] Taking averages saves space but does not affect the basic message, since maturation patterns vary little across cohorts (Appendix tables A.1-A.3).

What are the messages? First, column 1 of Table 1 shows the rate of decline in cohort membership is especially high between the first and second year, with more than 70 percent of firms dropping out. But conditional on making it to the second year, the survival probability is much higher, with an attrition rate around 40 percent the second year, and further declines occur thereafter. Thus, in terms of numbers, the most recent cohort is always larger than any previous one, and exporters with more than 15 years of market tenure are rare. For

[5]Similar tables for Colombian exports of all goods and to all destinations appear in Eaton, et al, 2008.

Table 1: **Average aggregates by cohort age**

	Actual data		
Cohort age	Exporters	Total Exports	Average Exports
1 year	1	1	1
2 years	0.29	1.11	3.77
3 years	0.18	0.93	5.03
4 years	0.14	0.67	4.66
5 years	0.12	0.63	5.18
6 years	0.10	0.51	4.99
7 years	0.08	0.50	5.72
8 years	0.08	0.45	5.91
9 years	0.07	0.39	5.58
10 years	0.06	0.40	6.58

Notes: Figures for cohorts aged 2-10 are expressed relative to corresponding figures for one-year-old cohorts.

example, firms that were exporting to the United States in 1992 account for less than five percent of the firms exporting to the United States towards the end of the sample.

Second, column 2 shows that the rapid initial decline in its membership is not accompanied by a similar collapse in total cohort sales. The relative stablility of total sales means, of course, that sales per firm are growing substantially. Indeed, as can been seen in column 3, sales per surviving exporter more than triple from the first to the second year, increase again in the cohort's third year, and show no strong tendency to grow further or shrink thereafter.

2.3 Patterns of buyer-seller matches

We next use the data to characterize the buyer-seller matches that took place during 1992-2009.

2.3.1 Monogamous and polygamous matches

The number of Colombian exporters appearing in our sample grew from 2,232 in 1992 to 3,300 in 2009, a growth of 2 percent per annum, while the number of U.S. importing firms grew by 3 percent per annum (Appendix 1, Table **??**). The number of Colombian exporter-U.S. importer pairs (representing at least one transaction between them in a year) also grew at an annual rate of 2 percent. Roughly 80 percent of matches are monogamous in the sense that the buyer deals with only one Colombian exporter and the exporter ships to only one buyer in the United States. However, since the remainder of the matches are polygamous, the average Colombian exporter was involved in relationships with around 1.3 U.S. firms while the average U.S. buyer was involved with around 2.3 Colombian firms. Both figures declined slightly over the period.

2.3.2 Transition Probabilities

Like firms' exporting stints (Table 1), most of their buyer-seller matches are short-lived. Of the 3,087 matches that existed at the beginning of the period, 70 percent didn't make it to 1993. But, of those that made it into the next year, almost 50 percent made it into the next year. Similarly, of the relationships that existed in 2005, 57 percent started that year but of those that started before, 37 percent had been around at least three years before. Of the 3,210 matches identified in 1992, less than 25 endure (are present every year) throughout the period.

Table 2: Transition Probabilities, Number of Clients

t \t+1	Out	Dormant	1	2	3	4	5	6-10	11+
Out	.	.	0.932	0.055	0.009	0.002	0.001	0.001	0.000
Dormant	.	.	0.876	0.100	0.015	0.008	.	.	0.000
1	0.539	0.080	0.321	0.048	0.010	0.002	.	0.001	.
2	0.194	0.077	0.375	0.241	.	0.024	0.009	0.004	.
3	0.090	0.042	0.220	0.271	0.210	0.092	.	0.027	.
4	0.059	.	0.129	0.216	0.215	0.184	0.083	0.095	.
5	.	.	0.095	0.184	0.181	0.181	0.126	0.178	.
6-10	.	.	0.039	0.073	0.089	0.123	0.157	0.419	0.073
11+	.	0.000	0.000	0.000	.	.	.	0.432	0.526

Table 2 reports the probability with which a Colombian firm participating in certain number of relationships with buyers transits into a different number of relationships the following year. (Confidentiality restrictions prevent us from reporting numbers for cells that are too sparsely populated.) This table reports the annual average for 1992-2009 across all industries. A firm that stops exporting but re-appears as an exporter sometime later in our sample period is considered to have gone "dormant", while those exporters that drop to zero foreign sales for the extent of our sample are considered to have gone "out" of exporting. Those that have never been observed to export constitute the pool of potential entrants.

Among first-time exporters, 93.2 percent sell to only one firm. Of these, 62 percent don't export the next year, and only about six percent go on to establish a larger number of relationships. For firms with three relationships in a year, about twelve percent enter into a larger number of relationships the next year. Hence there is an enormous amount of churning at the lower end. Even for firms with a large number of relationships the most likely outcome is to have fewer the next year.

2.3.3 Ergodic degree distribution

We can ask what this pattern of entry and growth implies about the ergodic distribution of relationships. If we assume that entrants in a year replace exiting firms, the ergodic distribution implied by this transition matrix is given by Table 3.

Table 3: Ergodic Client Distribution Implied by Transitions

	1	2	3	4	5	6-10	11+
Erg Distribution	0.792	0.112	0.031	0.016	0.009	0.022	0.016
Data	0.778	0.116	0.043	0.021	0.011	.	.

For purposes of comparison, the year-specific average share of Colombian firms in each group is reported as well. Note that the ergodic distribution implied by the transition matrix is very close to the cross-sectional distribution in the data, suggesting that over the period we observe the process has been quite stationary. Interestingly, both distributions are very nearly Pareto, reflecting the coexistence of many small scale exporters with a few "super-exporters."

2.3.4 Match maturation

The survival probability of new matches increases with initial sales volume. Table 4 sorts observations on matches according to their size in their first year of existence and reports year-to-year separation rates. In addition to the very low survival rates, two patterns stand out. First, those matches that begin with sales in the top quartile among all new matches are more likely to survive than matches that begin with smaller sales volumes. Second, survival probabilities improve after the initial year.

Further features of the match maturation process are evident in Figure **??**, which shows average annual sales per match, broken down by initial size quartile. For each size quartile, matches are further distinguished according to their total life span: less than one year (life=0), 1 to 2 years (life=1), and so forth. And for each cluster of bars, the left-most bar corresponds to sales in the initial year of the match's existence, the next bar corresponds to sales during the second year of the match's existence, and so forth.

The first message of these graphs is that initial sales are a good predictor of sales in subsequent years, conditioning on survival. Those matches with first-year sales in the smallest

Table 4: Separation Rates, by Age of Match and Initial Sales

	1 year	2 years	3 years	4 years	5+ years
Quartile 1	82.9	63.2	57.3	55.0	49.7
Quartile 2	75.6	58.4	49.4	46.8	43.7
Quartile 3	67.7	52.1	44.6	40.8	37.6
Quartile 4	52.1	44.5	40.3	39.2	36.7

quartile systematically generated the lowest annual sales in subsequent years, and more generally, first-year sales are monotonically related to annual sales in subsequent years. (Note the different scales of the vertical axes in different panels of Figure **??**.) Second, sales tend to jump from the first to the second year, in large part because observations on a match's first year correspond to less than a full calendar year. (There is an analogous effect at work in the final year of a match's life.) Looking at complete-year observations reveals a tendency for annual sales to grow among matches that start small and survive, but no such tendency among matches that start in the largest quartile. Finally, looking across matches with different life spans, those that survive more years tend to have higher sales in all (full) years than matches that fail relatively quickly. This pattern is robust across matches in the different quartiles for initial sales.

3 A Model of Exporting at the Transactions Level

We now develop a model of exporter behavior consistent with the patterns reviewed above. Buyer-seller relationships form and disband at irregular intervals. Similarly, export shipments are discrete events distributed unevenly through time. To capture these features of the data, and to allow agents to update their behavior each time their circumstances change, we formulate our model in continuous time, treating all of the exogenous processes in our model as Markov jump processes.

Explaining the evolution of a firm's exports and domestic sales requires modeling both its sales to existing buyers and the evolution of its portfolio of clients. We can treat these two components sequentially. We first consider the relationship between a seller and an individual buyer. Having characterized the seller's profits from a relationship with an individual buyer, we then turn to her learning about the popularity of her product, i.e., the chance that a potential buyers likes her product. Finally, we characterize her search for buyers.

3.1 A Seller-Buyer Relationship

This section characterizes the profit streams that sellers generate from successful business relationships. The expressions we develop here describe relationships between domestic firms and foreign buyers, but with appropriate relabelling of market-wide variables they apply equally to relationships between domestic firms and domestic buyers.

3.1.1 Profits from a single shipment

Several features of our model are standard. First, at any time t seller j can hire workers at a wage w_t in real local currency units, each of whom can produce $\varphi_j \in \{\varphi^1, .., \varphi^{N_\varphi}\}$ units of output.[6] Hence seller j's unit cost in local currency is w_t/φ_j. If she sells at price p_{jt} in foreign currency her unit profit in local currency is

$$p_{jt}/e_t - w_t/\varphi_j, \tag{1}$$

where e_t is the exchange rate. Second, goods markets are monopolistically competitive and each producer supplies a unique differentiated product.

Once buyer i has agreed to form a business relationship with seller j, he periodically places sales orders with j. For j, an order from i that arrives at time t generates revenue:

$$X_{ijt} = \left(\frac{p_{jt}}{P_t}\right)^{1-\eta} y_{ijt}\overline{X}_t, \tag{2}$$

where $\eta > 1$ is buyers' elasticity of demand, p_{jt} is the price of seller j's product, $\overline{X}_t$ is the average spending level among all potential foreign buyers, P_t is the relevant price index for all competing products in the foreign market, and $y_{ijt} \in \{y^1, .., y^{N_y}\}$ is a time-varying demand shifter idiosyncratic to the ij relationship.[7]

For simplicity, and to keep the analysis as close as possible to other heterogenous firm models, we assume that the seller posts a non-negotiable price, charging the optimal markup over unit cost:[8]

$$p_{jt} = \frac{\eta}{\eta - 1}\frac{e_t w_t}{\varphi_j} \tag{3}$$

By (1), (2), and (3), an order from buyer i at time t therefore generates the following profits for seller j:

[6] We treat φ as time-invariant to facilitate model identification. Other sources of idiosyncratic temporal variation in sales will be discussed shortly.

[7] Not all buyers necessarily face the same range of goods and hence the same aggregate price index P. We treat idiosyncratic components of the price index as P as reflected in y_{ijt}.

[8] An alternative specification would introduce bilateral bargaining between buyer and seller.

$$\pi_{ijt} = \frac{1}{\eta}\frac{\overline{X}_t}{e_t}\left(\frac{e_t w_t \eta/(\eta-1)}{\varphi_j P_t}\right)^{1-\eta} y_{ijt}.$$

We can combine all the macroeconomic variables affecting the profit of any seller from this source selling in this destination, along with constants, as:

$$x_t = \frac{1}{\eta}\frac{\overline{X}_t}{e}\left(\frac{e_t w_t \eta/(\eta-1)}{P_t}\right)^{1-\eta},$$

where $x \in \{x^1, .., x^{N_x}\}$ is general to all potential buyers in the foreign market. Suppressing subscripts on state variables, this allows us to write the profits from a sale as:

$$\pi_\varphi(x, y) = x\varphi^{\eta-1}y, \tag{4}$$

In what follows, (4) is all we take from our specification of preferences and pricing behavior into the dynamic analysis. Any set of assumptions that deliver this simple multiplicative expression for a firm's profit from a sale would serve us equally well.

3.1.2 Relationship dynamics

At any point in time, each seller maintains business relationships with an endogenous number of buyers. These relationships form as a consequence of a search process that will be characterized in the following section, and they dissolve for several reasons. First, there is a constant exogenous hazard δ that any particular relationship will terminate, which could be due to the demise of the buyer or the buyer no longer finding the seller's product useful. Second, after each sale to a particular buyer, the seller evaluates whether it is worth sustaining her relationship with him. Doing so keeps the possibility of future sales to him alive, but it also means paying the fixed costs F of maintaining the account, providing technical support, and maintaining client-specific product adjustments.[9]

When deciding whether to maintain a particular business relationship, the seller knows her own type, φ, the macro state, x and profits from the current sale, $\pi_\varphi(x, y)$ to the buyer in question. She can therefore infer this buyer's current y value and calculate the value of her relationship with him to be:

$$\widetilde{\pi}_\varphi(x, y) = \pi_\varphi(x, y) + \max\left\{\widehat{\pi}_\varphi(x, y) - F, 0\right\}.$$

[9]For instance, Colombian producers of construction materials interviewed for a related project (Domínguez et al, 2013) referred that it is frequent for foreign buyers to request adjustments in the specifications of products or packages. In turn, these require adjustments in the production process that are costly to maintain.

Here $\widehat{\pi}_\varphi(x,y)$ is the expected value of continuing a relationship that is currently in state (x,y). Clearly the seller terminates this relationship if $\widehat{\pi}_\varphi(x,y) < F$.

If a seller pays F to keep a relationship active, and if the relationship does not end anyway for exogenous reasons, one of several events will next affect it: with hazard λ^b the buyer will place another order, with hazard $q^X_{xx'}$ x will jump to some new marketwide state $x' \neq x$, or with hazard $q^Y_{yy'}$ y will jump to some new buyer-specific shock $y' \neq y$.[10] Let τ_b be the random time that elapses until one of these events occurs. Given that x and y are Markov jump processes, τ_b is distributed exponentially with parameter $\lambda^b + \lambda^X_x + \lambda^Y_y$, where

$$\lambda^X_x = \sum_{x' \neq x} q^X_{xx'} \tag{5}$$

and

$$\lambda^Y_y = \sum_{y' \neq y} q^Y_{yy'}, \tag{6}$$

are the hazards of transiting from x to any $x' \neq x$, and from y to any $y' \neq y$, respectively. Then assuming the seller has a discount factor ρ, the continuation value $\widehat{\pi}_\varphi(x,y)$ solves the Bellman equation:

$$\begin{aligned}
\widehat{\pi}_\varphi(x,y) &= \mathbf{E}_{\tau_b}\left[e^{-(\rho+\delta)\tau_b}\frac{1}{\lambda^b + \lambda^X_x + \lambda^Y_y}\left(\sum_{x' \neq x} q^X_{xx'}\widehat{\pi}_\varphi(x',y) + \sum_{y' \neq y} q^Y_{yy'}\widehat{\pi}_\varphi(x,y') + \lambda^b\widetilde{\pi}_\varphi(x,y)\right)\right] \\
&= \frac{1}{\rho + \delta + \lambda^b + \lambda^X_x + \lambda^Y_y}\left(\sum_{x' \neq x} q^X_{xx'}\widehat{\pi}_\varphi(x',y) + \sum_{y' \neq y} q^Y_{yy'}\widehat{\pi}_\varphi(x,y') + \lambda^b\widetilde{\pi}_\varphi(x,y)\right)
\end{aligned}$$

Before a seller has met her next buyer, she does not know what state y this buyer will happen to be in. So when choosing her search intensity for new business relationships, she must base her decisions on the ex ante expected pay-off to forming a new business relationship. Given the market state x, a type-φ seller calculates this expected value as:

$$\widetilde{\pi}_\varphi(x) = \sum_s \Pr(y^s)\widetilde{\pi}_\varphi(x,y).$$

where $\Pr(y^s)$ is the probability that a randomly selected buyer is currently in state $y^s \in \{y^1, .., y^{N_y}\}$.[11]

[10] Since sales in the data are discrete events rather than flows, we model the buyer's purchases accordingly. We think of the buyer not as making use of the products continually but in discrete spurts. For example, the buyer might be a producer of a product that it makes in batches. At the completion of each batch it buys inputs for the next batch.

[11] Here we take the probabilities $\Pr(y^m)$ to be the ergodic distribution of y implied by the transition hazards $q^Y_{yy'}$. We could assume that the distribution at the time of the first purchase is different from the ergodic one.

For the purposes of the search model that follows, all that matters about an individual relationship is $\widetilde{\pi}_{\varphi}(x)$, and this object can be estimated directly from data on the revenue streams generated by matches. Nonetheless, the history of a seller's interactions with a given buyer affects its overall sales trajectory and hence matters for our characterization of aggregate export dynamics.

Hereafter, we will denote the expected value of a relationship with a foreign buyer by $\widetilde{\pi}_{\varphi}^{f}(x)$ and the expected value of a relationship with a home market buyer by $\widetilde{\pi}_{\varphi}^{h}(x)$. These two objects are calculated in the same way, but since expenditure levels $(\overline{X}_t)$ and price indices (P_t) differ across markets, and no exchange rate factor e is necessary for domestic profit calculations, each has its own process for the market-wide state variable, x. These market-wide demand shifters are denoted x^f and x^h below.

3.2 Learning about Product Appeal

In each market, sellers conduct market-specific searches for buyers. When searching in market $m \in \{h, f\}$, each recognizes that some fraction $\theta^m \in [0,1]$ of the potential buyers she meets there will be willing to do business with her. An encounter with one of these willing buyers generates an expected profit stream worth $\widetilde{\pi}_{\varphi,x}^{m}$, while an encounter with any of the remaining potential buyers does not generate a sale then or subsequently.

Each seller's θ^h and θ^f values are drawn before she has met any clients. These draws remain fixed through time, inducing permanent cross-market differences in her product's popularity. All θ^m draws are independently beta-distributed across sellers and markets:

$$b(\theta^m|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} (\theta^m)^{\alpha-1} (1-\theta^m)^{\beta-1},\ m \in \{h, f\},$$

where $\Gamma(\phi) = \int_0^\infty z^{\phi-1}e^{-z}dz$ is the gamma function (needed to ensure that the distribution has the proper limits). However, the independence of θ^h and θ^f does not mean sellers' domestic and foreign sales are likewise independent. Rather, cross-market correlation in sales will be induced by the firm type φ, which can be viewed as capturing aspects of product appeal that are common to both markets.[12]

Benchmark model: Sellers are presumed to have already met many potential customers in the domestic market, and thus to have learned their θ^h draws. But sellers typically have far less experience abroad, so in the benchmark version of our model, we allow them to

[12] The firm effect is similarly interpreted to reflect both productive efficiency and product appeal in Melitz (2003) and many other papers based on CES demand systems. However in the present context, the global aspects of product appeal captured by φ are qualitatively distinct from the market-specific product appeal effects captured by θ. The former determines the amount of a product each buyer purchases, given that he is interested, while the latter determines what fraction of potential buyers are willing to place orders with the seller, should they happen to meet her.

still be learning about their θ^f draws. Specifically, each seller recognizes that for any given θ^f, the probability a random sample of n potential foreign buyers will yield a customers is binomially distributed:

$$q\left[a|n,\theta^f\right] = \binom{n}{a}\left[\theta^f\right]^a\left[1-\theta^f\right]^{n-a}.$$

So after she has met n^f potential buyers abroad, a^f of whom were willing to buy her product, a seller's posterior beliefs about her θ^f draw are distributed:

$$p(\theta^f|a^f,n^f) \propto q\left[a^f|n^f,\theta^f\right]\cdot b(\theta^f|\alpha,\beta)$$

where the factor of proportionality is the inverse of the integral of the right-hand side over the support of θ^f. Since the beta distribution is the conjugate prior for the binomial, a firm's expected success rate after a successes in n trials has a convenient closed-form representation:

$$\overline{\theta}^f_{a,n} = E\left[\theta^f|a^f,n^f\right] = \int_0^1 \theta p(\theta|a^f,n^f)d\theta = \frac{a^f+\alpha}{n^f+\alpha+\beta}. \quad (7)$$

This posterior mean converges to $p\lim\left(\frac{a^f}{n^f}\right) = \theta^f$ as n gets large.

No-learning model: As an alternative to our benchmark model, we consider the possibility that sellers already know their product's popularity in *both* markets, so that $p(\theta^f|a^f,n^f)$ is a degenerate distribution and $\overline{\theta}^f_{a,n} = \theta^f$. In this version of the model, sellers' matching histories only affect their search intensities by affecting their visibility in each market, as we will discuss shortly. Our no-learning model is not nested by the benchmark model, it is simply a different characterization of the role of information in driving search policies.

3.3 Searching for Buyers

To complete our characterization of firms' behavior, we now consider sellers' search intensities in each market. Each seller continuously chooses the market-specific hazard s^m, $m \in \{h,f\}$, with which she encounters a potential buyer, recognizing that this involves the instantaneous flow cost $c^m(s^m,a^m)$, where $c^m(s^m,a^m)$ is increasing and convex in s^m.[13] Whether $c^m(s^m,a^m)$ increases or decreases in the number of successful matches, a^m, depends upon the relative

[13] Interviews conducted with Colombian exporters revealed a variety of activities firms pursue to meet potential buyers abroad (Domínguez, et al, 2013). Ranked roughly in terms of decreasing cost, these included maintaining a foreign sales office; paying the exports promotion office to organize visits with prospective clients abroad, and sending their sales representatives to those visits; sending sales representatives abroad to visit potential clients on their own; attending trade fairs; paying a researcher to search the web for foreign firms that purchase products similar to their own; paying browsers to ensure that their site appear near the top of a search for their product type; maintaining a web site in English. Interviewees also reported that relatively low-cost activities, such as traveling to trade fairs, or translating their websites to English, led to relationships with one or two clients every few years. Establishing a larger network of clients required much more costly activities.

strength of several forces and will be left for the data to determine. Costs might fall with a^m because encounters with interested buyers increase the seller's visibility and enhance her opportunities to meet additional potential buyers. Alternatively, costs might rise if the pool of easy-to-reach buyers becomes "fished out," as in Arkolakis (2007).

We can now describe optimal search behavior, beginning with the foreign market. Recall that when the foreign market state is x^f, a type-φ seller expects the value of a new business relationship will be $\widetilde{\pi}_{\varphi}^f(x^f)$. Further, she believes the next match will yield such a relationship with probability $\overline{\theta}_{a,n}^f$. Combined with search cost function $c^f(s^f, a^f)$ and the jump process for x^f, these objects imply sellers' optimal search policy abroad.

To characterize this policy, let τ_s^f be the random time interval until the next foreign search event, which could be either a change in the marketwide state x^f or an encounter with a potential buyer. Then, suppressing market superscripts, the optimal search intensity s for a type-φ firm with foreign market search history (a, n) solves the following the Bellman equation:

$$V_{\varphi}(a,n,x) = \max_{s} \mathbf{E}_{\tau_s} \left[-c(s,a) \int_0^{\tau_s} e^{-\rho t} dt + \frac{e^{-\rho \tau_s}}{s + \lambda_x^X} \cdot \left(\sum_{x' \neq x} q_{xx'}^X V_{\varphi,}(a,n,x') \right. \right.$$
$$\left. \left. + s \left[\overline{\theta}_{a,n} (\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1,n+1,x) + (1 - \overline{\theta}_{a,n}) V_{\varphi}(a,n+1,x) \right] \right) \right]$$

(Recall that λ_x^X is given by (5).) Taking expectations over τ_s yields:

$$\begin{aligned} V_{\varphi}(a,n,x) \quad &= \quad \max_{s} \frac{1}{\rho + s + \lambda_x^X} \left[-c(s,a) + \sum_{x' \neq x} q_{xx'}^X V_{\varphi,}(a,n,x') \right. \\ &\left. + s \left\{ \overline{\theta}_{a,n} \left[\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1,n+1,x) \right] + (1 - \overline{\theta}_{a,n}) V_{\varphi}(a,n+1,x) \right\} \right] \end{aligned} \tag{8}$$

Applying the multiplication rule for differentiation and using expression (8) for $V_{\varphi}(a,n,x)$, the optimal search intensity s^* satisfies:

$$\frac{\partial c(s^*,a)}{\partial s} = \overline{\theta}_{a,n} \left[\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1,n+1,x) \right] + (1 - \overline{\theta}_{a,n}) V_{\varphi}(a,n+1,x) - V_{\varphi}(a,n,x) \tag{9}$$

That is, the marginal cost of search must equal the expected marginal benefit of a match, which includes the expected value of the associated profit stream, $\overline{\theta}_{a,n} \widetilde{\pi}_{\varphi}(x)$, and the expected value of the information generated.

Now consider the home market. Since we assume sellers have already learned their true success rates at home, θ_j^h, new encounters do not influence expectations, and we need not condition the value function or the expected success rate on search histories. **[JT: I put an**

a **argument in V()**] Again suppressing market superscripts, the Bellman equation collapses to:

$$\begin{aligned} V_{\varphi}(x, a) &= \max_{s} \frac{1}{\rho + \lambda_x^X} \Bigg[-c(s, a) + \sum_{x' \neq x} q_{xx'}^X V_{\varphi}(x', a) \\ &\quad + s \left\{ \theta_j^h \left[\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1, x) \right] + (1 - \theta_j^h) V_{\varphi}(a, x) \right\} \Bigg] \end{aligned} \tag{10}$$

and the first-order condition is simply:

$$\frac{\partial c(s^*, a)}{\partial s} = \theta_j^h \left[\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1, x) - V_{\varphi}(a, x) \right].$$

The marginal cost of search equals the expected profit from a successful relationship times the probability of success. Of course, this condition also describes foreign market search in the no-learning version of the model.

4 An empirical version of the model

4.1 The search cost function

To implement our model empirically, we impose additional structure in several respects. First, we specify a functional form for our search cost function. Generalizing Arkolakis (2007) to allow for network effects, we write these costs as:

$$c^m(s^m, a^m) = \kappa_0^m \frac{[(1 + s^m)]^{\kappa_1} - 1}{\kappa_1 \left[1 + \ln(1 + a^m)\right]^{\gamma}}. \tag{11}$$

where $m \in \{h, f\}$. Several properties of this function merit note. First, marginal costs fall at a rate determined by γ with the number of successful matches a seller has already made, so $\gamma > 0$ implies “network” effects and $\gamma < 0$ implies "congestion" effects.[14] Second, a seller who is not searching in a particular market incurs no search cost: $c^m(0, a^m) = 0$. Third, given the cumulative number of successful matches, a, the marginal cost of search increases with s at a rate determined by κ_1 : $\frac{\partial c^m(s^m, a^m)}{\partial s^m} = \kappa_0^m (1 + s^m)^{\kappa_1 - 1} / \left[1 + \ln(1 + a)\right]^{\gamma}$. Fourth, we allow the cost function scalar to vary across markets, since the cost of maintaining any given level of visibility is likely to be higher in foreign markets. Finally, since a^m is the cumulative number of successes in market m, network effects endure, even if a firm is not actively searching.

[14] To contain the dimensionality of the computational problem we solve, we assume that firms with more than a^* buyers have (i) exhausted their learning effects, and (ii) reap no additional network effects at the margin from further matches. We choose a^* to exeed the observed maximum a for 99 percent of sellers in the foreign (United States) market. Also, we set $a = a^*$ for all sellers in their home (Colombian) market.

4.2 Processes for exogenous state variables

Next we impose more structure on the exogenous state variables, φ, x^h, x^f, y^h and y^f. All are assumed to have zero means in logs, and the net effect of these normalizations is undone by introducing scalars Π^h and Π^f into the home and foreign profit functions, respectively:

$$\begin{aligned} \pi_\varphi^f(x^f, y^f) &= \Pi^f x^f \varphi^{\eta-1} y^f, & (12) \\ \pi_\varphi^h(x^h, y^h) &= \Pi^h x^f \varphi^{\eta-1} y^f & (13) \end{aligned}$$

More substantively, we impose that the cross-firm distribution of φ is log normal with variance parameter σ_φ, and we treat all of the Markov jump processes (x^h, y^h, x^f, y^f) as independent Ehrenfest diffusion processes. The idiosyncratic match shocks, y^f and y^h, are assumed to share the same distribution, but we allow the x^f and x^h processes to differ. Among other things, the latter accommodates the fact that the exchange rate affects aggregate demand and price indices in the two markets differently.

Any variable z generated by an Ehrenfest process can be discretized into $2g+1$ possible values, $g \in I^+ : z \in \{-g\Delta,\ -(g-1)\Delta,\ ..,\ 0, ..,\ (g-1)\Delta,\ g\Delta\}$. Further, it jumps to a new value with hazard λ_z, and given that a jump occurs, it goes to z' according to:

$$z' = \begin{cases} z + \Delta \\ z - \Delta \\ \text{other} \end{cases} \text{with probability} \begin{cases} \frac{1}{2}\left(1 - \frac{z}{g\triangle}\right) \\ \frac{1}{2}\left(1 + \frac{z}{g\triangle}\right) \\ 0 \end{cases}.$$

Thus, given a grid size g, the intensity matrices $Q^X = \left\{q_{ij}^X\right\}_{i,j=1,N^X}$ and $Q^Y = \left\{q_{ij}^Y\right\}_{i,j=1,N^Y}$ that were introduced in section 3.1 are each block-diagonal and characterized by a single parameter, Δ.

5 Estimation

5.1 Stage 1: estimating observable jump processes

Shimer (2005) shows that if z follows a continuous time Ehrenfest diffusion process, it asymptotes to an Ornstein-Uhlenbeck process with mean zero as the fineness of the grid increases:[15]

$$dz = -\mu z dt + \sigma dW.$$

[15]Specifically, replacing the parameter vector (λ, g, Δ) with $(\lambda/\epsilon, g/\epsilon, \Delta\sqrt{\epsilon})$, $\epsilon > 0$, leaves the autocorrelation parameter μ and the instantaneous variance parameter σ unchanged. But as $\epsilon \to 0$, the innovation dW approaches normal.

Table 5: Market-wide Demand Shifters

	Parameter	*value*
home macro state jump hazard	λ^{x_h}	1.200
foreign macro state jump hazard	λ^{x_f}	1.215
home macro state jump size	Δ^{x_h}	0.003
foreign macro state jump size	Δ^{x_f}	0.053

Here $\mu = \lambda_z/g$, $\sigma = \sqrt{\lambda_z}\Delta$, and W follows a Weiner process. Accordingly, since it is possible to observe proxies for x^f and x^h, these can be viewed as discrete time observations on underlying Ornstein-Uhlenbeck processes, and the parameters of these processes can be econometrically estimated. Then, given μ and σ, estimates of Δ and λ for these processes can be inferred.

Measuring x^f as real expenditures on manufacturing goods in the U.S., and measuring x^h as real expenditures on manufacturing goods in Colombia, we obtain the results reported in Table 5.[16] They imply that x^f and x^h both jump 1.2 times per year, on average. However, jumps in the U.S. market tend to be much larger, essentially because they reflect movements in the real exchange rate as well as movement in dollar-denominated expenditures.

5.2 Stage 2: Indirect inference

Our data are relatively uninformative about the rate of time discount ρ and the demand elasticity η, so we do not attempt to estimate either one. For the former we follow convention and assume $\rho = 0.05$. For the latter, following many previous trade papers, we fix the demand elasticity at $\eta = 5$. Also, to limit the size of the estimated parameter vector, we take the exogenous match failure rate to be the observed match failure rate among matches at least 3 years old ($\delta = 0.326$), we take the search cost function to be quadratic in search intensity ($\kappa_1 = 2$), and we assume that the hazard rate for the buyer is once per quarter ($\lambda_y = 4$).[17]

All of the remaining parameters we estimate jointly using the transactions data summarized in Section 2 above. These parameters include the market size scalars ($\Pi^h, \Pi^{f\cdot}$), the fixed costs of maintaining a match (F^h, F^f), the parameters of the product appeal distributions (α,β), the dispersion of the productivity distribution (σ_φ), the jump size for the idiosyncratic

[16]Our foreign market size mesaure is the OECD time series on American GDP in 'Industry, including energy' adding imports and subtracting net exports of manufactures. Our home market size measure is real Colombian expenditures on manufacturing goods, taken from DANE. We converted all of the data used for the estimation into real 1992 US dollars, deflating nominal US dollars with the consumer price index available on the US Bureau of Labor Statistic website. We used an official Colombian Peso - US Dollar exchange rate time series downloaded from the Central Bank of Colombia to convert Pesos to nominal US Dollars

[17]These last three parameters could in principle be estimated, and in earlier drafts we have done so. However, they have not appeared to be well-identified.

buyer shocks $(, \Delta_y)$, the hazard rate for shipments (λ_b), the network/congestion parameter (γ), and the market-specific cost function scaling parameters $(\kappa_0^h,\, \kappa_0^f)$. For notational convenience we collect these parameters in the vector Λ :

$$\Lambda = \left(\Pi^h, \Pi^{f\cdot}, F^h, F^f, \alpha, \beta, \sigma_\varphi, \Delta_y, \lambda_b, \gamma, \kappa_0^h, \kappa_0^f\right)$$

We construct our estimator for Λ using the method of indirect inference (Gouriéroux and Monfort, 1996). That is, for each candidate Λ vector, we use the model to simulate the foreign and domestic transactions of an artificial sample of producers. Then, using these simulated data, we estimate a set of reduced-form regressions that summarize the relationships we want our model to capture. Finally, looking across candidate Λ vectors, we choose the one that makes the regression coefficients from simulated data correspond as closely as possible to the corresponding regression coefficients based on sample data. Algebraically, our estimator is

$$\hat{\Lambda} = \min_{\Lambda} \left[\bar{m} - m(\Lambda)\right]' W \left[\bar{m} - m(\Lambda)\right],$$

where $\bar{m}$ is a column vector of regression coefficients obtained from sample data, $m(\Lambda)$ is the analogous vector of regression coefficients based on data simulated at Λ, and W is a compatible non-singular weighting matrix. Setting $W^{-1} = var(\bar{m} - m(\Lambda))$ maximizes the efficiency of this estimator, but any non-singular W will yield consistent estimates. We use a block-diagonal version of $var(\bar{m} - m(\Lambda))$, with each block corresponding to the moments from a particular regression.

The regressions themselves are reported in Tables 6, 7 and 8. In each table, the data-based regression estimates are reported, and their standard errors are reported below them in parentheses. To facilitate comparisons between the sample and the simulated data, and with no loss of information, we have replaced the intercept of each regression with the mean value of the dependent variable in cases where that was possible.[18] We now briefly describe these regressions and our reasoning in choosing them.

Search policy. The first regression in Table 6 summarizes the effects of firms' market experiences on their search intensity (s). Roughly speaking, this equation can be viewed as a second order approximation to the foreign market policy function (9)–a central object in our model. The dependent variable is a proxy for a firm's foreign market search intensity after n successful matches, namely, the inverse of the time interval between firm j's n^{th} and $n+1^{st}$ match. And the right-hand side is a second-order translog function of this firm's cumulative number of successes (a) and cumuluative success rate $(\frac{a}{n})$. To deal with firms that have had no successes, we add 1 to a and to $\frac{a}{n}$ before taking logs.

[18]Several regressions were done in real pesos within the Colombian national statistical agency (DANE). We are not confident that they can be expressed in units that are strictly comparable to the real dollar units in which U.S. customs records were expressed.

The unit of observation here is an exporter-specific new match, and we define a new match to occur whenever an exporter makes a shipment to a buyer it has not dealt with before. We view this first shipment as a sample of the exporter's merchandise, so we only consider this match to be successful if it results in at least one additional shipment. This interpetation of the data means we can use customs records to directly infer the cumulative number of successes for each firm j (a_{nj}) after each of its $n \in \{1, ..., N_j\}$ matches, and the associated cumulative success rates $(\frac{a}{n})_{nj}$.

Interpreting the coefficient estimates for this regression is problematic, both because it includes second order terms and because we have not controlled for the highly nonlinear firm effects generated by φ and θ^f. But evaluation of this equation on a grid of success rates and cumulative successes gives us a crude sense for the relationships implied by our estimates. This exercise show that search intensity shows little sensitivity to success rates, but it strongly increases with cumulative successes.

Separation policy. Equation (ii) captures a second basic feature of firms' exporting behavior: their match termination policies. Here the unit of observation is seller j's i^{th} match in year t, and the dependent variable $D^{exit\ match}$ takes a value of unity when this match is in its final year.[19] Our model implies that matches are more likely to terminate when the idiosyncratic demand shock z_{ijt} and/or the firm's productivity level φ_j is low. Neither variable is directly observable, so we use several of their correlates as explanatory variables: current match sales, X^f_{ijt}, age of the match, A_{ijt}, and export market tenure, Δ_{ijt}. All variables are expressed in logs and, given the patterns revealed by Table 4, we allow firms in their first year of exporting ($D^{new\ to\ mkt} = 1$) to experience particularly high failure rates.[20]

Equation (ii) helps to identify the fixed costs of maintaining a foreign match, F^f. That is, conditioning on sales, X^f_{ijt}, matches are more likely to survive when fixed costs are low. Failure rates are also affected by the volatility of z_{ijt}, which is governed by the jump size, λ_y.

Not surpisingly, estimates of equation (ii) reflect the same patterns that we noted in connection with Table 7. Matches in their first year are relatively likely to fail, as are matches that start with relatively small sales volumes. The results also show that exporters with more experience in foreign markets tend to have longer-lived relationships, a feature of the data that our model captures with cross-firm variation in productivity levels, φ.

Match success rates The remaining regressions in Table 4 concern the distribution of success rates, θ. Equation (iii) summarizes the average success rate among active exporters

[19] Only active matches are included in the sample.

[20] Note, however, that in Table 7, matches that die after a single shipment are treated as having existed for less than one year, while our model-based estimates treat these cases of single shipments as unsuccessful meetings that did not lead to business relationships.

and its relation to the cumulative number of meetings an exporter has had (n). Accordingly it is informative about $\alpha/(\alpha+\beta)$ and selection due to learning. Equation (iv) describes dispersion in success rates–i.e., the squared residuals from equation (iii)–among exporters with different experience (n) levels. Both regressions suggest that selection takes place as firms acquire market tenure, since success rates are higher among experienced (high-n) firms, and the dispersion in success rates among such firms is lower.

Client distributions and shipment frequencies. The next set of regressions appears in Table 7. Equation (v) summarizes the information on client distributions in Table 3. Specifically, letting $\Phi(\ell)$ be the fraction of exporters with no more than ℓ active clients, column (v) reports the regression of $\ln(1-\Phi(\ell))$ on $\ln \ell$ and $(\ln \ell)^2$.[21] We choose this functional form because earlier studies have found that exporters' foreign client distributions are approximately Pareto, implying that the relationship between $\ln(1-\Phi(\ell))$ and $\ln \ell$ is approximately linear. Note that our data confirm a nearly-Pareto client distribution, as the coefficient on the quadratic term is quite small (-0.055).

Equation (v) helps to identify the cost function parameters $(\kappa_0^h, \kappa_0^f, \gamma)$ because the client distribution largely reflects firms' search intensities. In particular, the network effects captured by the parameter γ determine how much of a search cost discount large (big a) firms enjoy, and thus the "fatness" of the right-hand tail of the client distribution $\Phi(\cdot)$.

Equation (vi), the other regression in 7, simply establishes the mean log number of shipments per year per continuing match. It serves as a target for the shipment arrival hazard and obviously helps to identify λ_b.

Match- and firm-level sales Regressions that characterize the time series properties of firms' exports, cross-firm dispersion in exports, and patterns of correlation between exports and domestic sales are collected in Table 8. These equations are particularly informative about the parameters $\left(\Pi^h, \Pi^{f\cdot}, F^h, F^f, \sigma_\varphi, \Delta_y\right)$. Equation ($vii$) is an AR1 in log match revenues, conditioned on match age and a dummy to control for first-year effects. By the logic reviewed in section 5.1 above, the root (0.826) and root mean square error (1.208) in this AR1 identify the jump size, Δ_y and the cross-firm variance in productivity, σ_φ, up to selection effects. Also, together with equation (ii), the mean log annual revenue per match (10.67) essentially pins down the profit function scalar and the fixed cost of maintaining a foreign match (Π_f, F^f).

The last four equations in Table 8 involve domestic sales. Since we don't observe firms' individual matches in the domestic market, these regressions describe establishment-level

[21] By construction, the intercept of the (non-parametric version of) this regression line must be zero.

panel data merged with Colombian customs records.[22] Equations ($viii$) is an AR1 for home sales, and is thus informative about the extent which firms adjust their domestic connections and their associated match specific sales in response to idiosyncratic shocks. Accordingly, the coefficients in this equation are particularly helpful in identifying κ_0^h and F^h, and the mean squared error helps identify σ_φ and $\alpha/(\alpha+\beta)$. Equation (ix) is a simple projection of firm level exports on firm-level domestic sales. It serves to distinguish market-specific variation in revenues from variation in revenues that is common to both markets. Thus the estimated parameters of this equation, including its mean squared error, are informative about the variance of productivity shocks (σ_φ^2), which are common to both markets, relative to the variance of market-specific appeal draws, θ^h and θ^f.[23]

[22] More precisely, regressions $viii$ through x in Table XXc are done using a combination of the Colombian Annual Manufacturing Survey (AMS) and Colombian administrative records of exports transactions. The data used cover 1993-2007. Exports are merged into the AMS using firm identifiers. This is done because the AMS has no export information for 1993-1999, and because the dynamics of aggregate exports reported in the EAM starting in 2004 differ substantially from aggregate reports from other sources.

[23] Given the average success rate, $\alpha/(\alpha+\beta)$, the variances of θ^h and θ^f depend only on $\alpha+\beta$.

Table 6: Match hazards, success rates, and endurance

	(i) $\ln(\lambda_{ij})$	(ii) $D_{ijt}^{exit\ match}$	(iii) $\frac{a_{ij}}{n_{ij}}$	(iv) $u^2_{a_{ij}/n_{ij}}$
mean, dep. variable	1.767 (0.621e-2)	0.395 (0.319e-2)	0.413 (0.153e-2)	0.091 (0.26e-4)
$\ln(1+n_{ij})$	–	–	0.093 (0.003)	-0.056 (0.000)
$\ln(1+n_{ij})^2$	-0.818 (0.113)	–	–	–
$[\ln(1+a_{ij})]^2$	0.312 (0.017)	–	–	–
$\ln(1+\frac{a_{ij}}{n_{ij}})$	-1.132 (0.296)	–	–	–
$\left[\ln(1+\frac{a}{n})\right]^2$	2.451 (0.396)	–	–	–
$\ln(1+a_{ij})\cdot\ln(1+\frac{a_{ij}}{n_{ij}})$	-0.708 (0.134)	–	–	–
$D_{ijt}^{new\ to\ mkt}$	–	0.034 (0.011)	–	–
$\ln X_{ijt}^f$	–	-0.031 (0.002)	–	–
$\ln A_{ijt}$	–	-0.054 (0.009)	–	–
$\ln \Delta_{jt}$	–	-0.028 (0.007)	–	–
observations (rounded)	38,500	23,500	35,800	35,800

Notes: Unit of observation, columns i, iii and iv: seller j's i^{th} match. Unit of observation, column ii: seller j's i^{th} match in its t^{th} year. λ_{ij} = inverse of time interval between commencement of match i and commencement of the next one for exporter j $D_{ijt}^{exitmatch} = 1$ if exporter $j's$ i^{th} match dies in year t. a_{ij} = cumulative number of successes for exporter j at time of match i. $D_{ijt}^{newtomkt} = 1$ if exporter $j's$ i^{th} match is in its first year. $\ln A_{ijt}$ = log age of exporter $j's$ i^{th} match. $\ln \Delta_{jt}$ = log age of exporter j in year t. X_{ijt}^f = foreign sales volume generated by exporter $j's$ i^{th} match.

Table 7: Client distribution and shipment frequency

	(v) $\ln(1-\Phi(\ell))$	*(vi)* $\ln(s_{ijt})$
mean, dep. variable	-5.973 (2.173)	0.971 (0.004)
$\ln(\ell)$	-1.8813 (0.1123)	-
$(\ln \ell)^2$	-0.0545 (0.0211)	-
sample restrictions	$\ell > 0$	$s_{ijt} > 0$

Notes: ℓ: number of active clients; $\Phi()$ = cumulative distribution of exporters in terms of ℓ; s_{ijt} = number of shipments per year to client i by exporter j in year t

Table 8: Home and foreign sales regressions

	(vii) $\ln X^f_{ijt}$	*(viii)* $\ln X^h_{jt}$	*(ix)* $\ln X^f_{jt}$	*(x)* D^f_{jt}	*(xi)* $\frac{X^f_{jt}}{X^f_{jt}+X^h_{jt}}$
mean, dep. variable	10.665 (0.002)	–	–	0.102 (0.003)	0.127 (0.002)
R_{ijt-1}	0.328 (0.018)	-	-	-	-
$\ln X^f_{ijt-1}$	0.826 (0.004)	-	-	-	-
$\ln X^h_{jt-1}$	-	0.976 (0.029)	-	-	-
$\ln X^h_{jt}$	-	-	0.323 (0.110)	-	-
$\ln \Delta_t$	0.063 (0.014)	-	-	-	-
root mse	1.2079	0.4621	2.1665	0.303	0.243
sample restrictions	$X^f_{ijt}, X^f_{ijt-1} > 0$	$X^h_{jt}, X^h_{jt-1} > 0$	$X^f_{jt}, X^h_{jt} > 0$	$X^h_{jt} > 0$	$X^f_{jt}, X^h_{jt} > 0$

Notes: $R_{ijt} = 1$ if exporter $j's$ i^{th} match is in its first year. $\ln \Delta jt$ = log age of exporter j. X^f_{ijt} = foreign sales volume generated by exporter $j's$ i^{th} match. X^f_{jt} = total foreign sales volume generated by firm j. X^h_{jt} = total home sales volume generated by firm j. $D^f_{jt} = 1$ if firm j is an exporter.

Finally, equations (x) and (xi) describe the relative importance of home versus foreign sales. The former gives the share of firms that participate in the foreign market and thereby speaks to the relative return to maintaining foreign versus domestic business connections, that is (Π^f, F^f, κ_0^f) versus (Π^h, F^h, κ_0^h). The latter gives the average share of exports to the U.S. in total sales of exporting firms. Accordingly, it largely reflects the number of clients in each market, and thus responds especially to differences between κ_0^f and κ_0^h.

Sensitivity analysis As suggested by Andrews et al. (2017), we check which moments are important using the sample analog to the matrix $(G'WG)^{-1}G'W$ where $G = \frac{-\partial[m(\Lambda)]}{\partial\Lambda'}$ is the Jacobian for the vector of simulated moments. "Intuitively, this matrix is a local approximation to the mapping from moments to estimated parameters." (Andrews, et al., 2017, p. 1555) Evaluated at our benchmark estimates (to be discussed), we obtain the results reported in Appendix 3. They are discussed therein in detail; here we summarize the patterns that emerge.

First, most parameters respond to many moments rather than one or several. Limiting our attention to elasticities with absolute value greater than 0.1, most parameters show significant responses to at least 5 moments, and several (F^f, F^h, γ) respond to more than 15. All parameters respond to at least 2. The moments that affect the most parameters are those generated by the match sales autogression (equation vii), the shipping rate regression (equation vi), the domestic sales autoregression (equation $viii$), the regression explaining the variance in success rates (equation iv), and the fraction of firms that export (equation x).

5.3 Parameter estimates

Table 9 reports estimates of the structural parameter vector Λ for both the benchmark and the no-learning model. Although our estimator exploits month-to-month transitions in the customs records, all hazards are normalized so that the unit of time is one year. Thus, for example, our estimate of δ implies that on average, matches last roughly 4 months (one-third of a year) before separating for exogenous reasons. Most parameter estimates are similar for both models, though the benchmark model fits the data better. [Report non-nested test here?] We therefore focus our discussion on the results for this model, turning later to the main distinguishing features of the no-learning results.

Benchmark parameter estimates Active matches generate an average of $\lambda_b = 15.43$ shipments per year, and the profits associated with these shipments vary widely across firms and macro conditions. Evaluating the gross profit-per-shipment functions (12) and (13) at our estimated values for Π^h, Π^f and the parameters governing realizations for φ, x, and y, we find that gross profits per shipment (before fixed costs) for a firm at the median productivity level matched to a median buyer are essentially zero. Accordingly, these firms are not active. On the other hand, a firm with productivity 1.9 standard deviations above the mean earns gross profits per shipment ranging from $4 to $42, depending upon what state its buyer is in. In the domestic market, the analogous figures range from $45 to $405. Further, a firm with the highest productivity matched to the best possible buyer in the most favorable macro state earns $31,512 in gross profits per export shipment and $281,570 in profit per domestic shipment. Of course, firms almost never attain these maxima, and when they do they are very unlikely to repeat their performance. This is consequence of the short expected life span of matches, and the fact that buyers' demands change an average of $\lambda_y = 4$ times per year.

These figures may seem small, but for several reasons they are consistent with the data. First, the productivity distribution for exporting firms come from the right-hand tail of the unconditional productivity distribution. Thus those firms with productivity 1.9 standard deviations above the mean unconditional mean of φ are actually the smaller exporters. Second, since revenues per shipment are $\eta = 5$ times profits per shipment, and since an average of $\lambda_b = 15.43$ shipment occur per year, expected annual revenues from a match that survives the entire year are $\eta \cdot \lambda_b = 77.15$ times as large as profits per shipment for that match.

Turning to the fixed cost estimates, note that both are quite small (F^f=\$0.30, F^h=\$0.03). These costs thus have no affect on major exporters. Nonetheless, they affect the fraction of exporting firms by keeping fringe players that would otherwise sell tiny amounts out of foreign markets.

The profit and cost function scalars are much more important. The model assigns lower search costs to the home market ($\kappa_0^h = 859.0$ versus $\kappa_0^f = 3,079.7$) and larger profits per sale ($\ln \Pi^h = -3.88$ versus $\ln \Pi^f = -6.16$). Both patterns help explain the small of output exported to the U.S. among these firms (Table 8, regression xi). And the two sets of scalars are separately identified by their different effects on match arrival rates (Table 4, regression i) and revenues from ongoing matches (Table 8, regressions vii and $viii$). The benchmark model also implies that search costs fall significantly as firms acquire market visibility through successful matches ($\gamma = 0.383$). As mentioned earlier, identification of this visibility effect comes largely from the shape of the client distribution (Table 7, regression v).

So what are the costs of making new matches? For a firm with no prior successful matches in the foreign market, a search intensity sufficient to yield an average of one new match per year costs $c^f(1, 0) = \$1,539$, but an expected yield of four new matches—about one successful match for a firm with average product appeal—costs $c^f(4, 0) = \$24,637$. The analogous figures in the home market are $c^h(1, 0) = \$428$ and $c^h(4, 0) = \$6,848$. But having an established reputation is helpful. A firm that has already made 2 successful foreign

Table 9: **Structural parameter estimates**

		Benchmark model		No-learning model	
	Parameter	*value*	*std. error*	*value*	*std. error*
log domestic profit scalar	$\ln \Pi^h$	-3.879	(0.1364)	-3.460	(0.0725)
log foreign profit scalar	$\ln \Pi^f$	-6.135	(0.1993)	-6.273	(0.0759)
fixed cost, domestic	F^h	0.027	(0.0047)	0.037	(0.0064)
fixed cost, foreign	F^f	0.296	(0.0428)	0.301	(0.0359)
First θ distribution parameter	α	0.571	(0.0454)	0.581	(0.0703)
Second θ distribution parameter	β	1.894	(0.2320)	4.661	(0.2107)
demand shock jump size	Δ^y	1.882	(0.2222)	1.951	(0.1810)
shipment order arrival hazard	λ_b	15.426	(0.1991)	15.431	(0.1428)
std. deviation, log firm type	σ_φ	1.386	(0.0095)	1.401	(0.0051)
network effect parameter	γ	0.383	(0.0485)	0.508	(0.0479)
home search cost function scalar	$\ln \kappa_0^h$	11.722	(0.1486)	12.408	(0.0850)
foreign search cost function scalar	$\ln \kappa_0^f$	13.002	(0.0095)	13.666	(0.1373)
fit metric		10.806		11.346	

matches could expect to pay only $c^f(4,2) = \$20,142$ for the next one—roughly 20 percent less than the cost of the first one. Similarly, a firm that has already made two successful home market matches could expect to pay $c^h(4,2) = \$5,598$ for the third. These reputation effects are nontrivial, and other things equal, they create a cost advantage for well-established firms.

Given match payoffs and search costs, firms' search intensity is determined by their expected success rates. Their (unobserved) actual rates are drawn from a beta distribution, which we estimate to have mean $\alpha/(\alpha+\beta) = 0.23$ and standard deviation $\sqrt{\alpha\beta/\left[(\alpha+\beta)^2(\alpha+\beta+1)\right]} = 0.23$. Hence, before they acquire export market experience, firms expect that roughly 1 in 4 new encounters with potential buyers will lead to business relationships. And since new exporters are uncertain about their θ^f draws, they expect to learn a good deal from the outcomes of their early matches.

No-learning parameter estimates Recall that our no-learning model differs from the benchmark model in that it presumes each firm j already know the fraction of the foreign population of buyers that is willing to do business with it, θ_j^f. This assumption implies that low-appeal firms never bother to invest much in foreign market searches, and firms that do search abroad have no incentive to do so relatively intensively when they are new to the market. That is, the information value of matches is no longer present.

The last two columns of Table 9 present parameter estimates based on this version of the model. Most parameters are similar, but the estimate of the network effect is larger ($\gamma = 0.50$ versus $\gamma = 0.38$) and the estimates of search costs are higher ($\kappa_0^h = 859$ and $\kappa_0^f = 3,079$ versus $\kappa_0^h = 1,826$ and $\kappa_0^f = 5,982$). Higher search costs scalars and larger network effects appear to help the model explain the observed ratio of small, gradual growth and eventual dominance of high-θ entrants without relying on learning effects. However, the no-learning version of the model does significantly worse than the benchmark version according to Rivers and Vuong's (2002) test statistic for non-nested comparisons: **XX.** Accordingly, we will focus our discussion on the benchmark model for the remainder of the paper.

6 Analysis of results

6.1 Model fit

Appendix 3 juxtaposes the data-based moments, $\bar{m}$, with their simulated counterparts, $m(\Lambda)$. Generally, the patterns in the data are replicated by our model, though not all of the model-based equation estimates correspond closely to their data-based counterparts. In particular, average exporting rates, match-specific sales dynamics, and the client distribution are well-captured by the model, as are most mean values of dependent variables. However the model fails to generate the association between success rates and firms' search intensities that we observe in the data. A more detailed summary of the fit can be found in Appendix 3.

Since we have not targeted the patterns described in section 1 when estimating our model, it is instructive to check how well they are replicated by our model. Tables 10, 11 and 12 below provide answers. In the top panel of table 10, the information in tables 1-3 is collapsed by averaging across exporting cohorts for which we observe at least 10 years of data. (These cohorts were born in the years 1997 through 2002.) Values of each aggregate for 2-year olds, 3-year olds, and so on are expressed as fractions of the corresponding values for 1-year olds. For example, the data tell us that, on average, only 29 percent of the exporters who began exporting in year t were still exporting in year t+1, and only 5 percent of those that began exporting in year t were present in year t+9. Likewise, among Colombian exporters that survive in the U.S. market for 10 years, average exporters per firm are 6.58 times as large as they are among exporters that are in their first year of exporting.

The bottom panel of table 10 shows corresponding figures based on model-simulated data. Qualitatively, the patterns in the actual and the simulated data match up. For both data sets, the largest drops in the number of exporters occur during the first two years, thereafter cohort size drops gradually. Likewise, total exports rise early in cohort's life, and decline thereafter. Finally, exports per surviving firm grow rapidly over time, reflecting both the exit of small-scale firms and client accumulation among survivors. It should be noted, however, that the "average exports" and "total exports" series based on actual data vary less dramatically with cohort age than the simulated data. Also, in the data-based figures, the drop in cohort membership is more dramatic during the first year. In sigifnicant part, these discrepancies reflect the fact that the data-based figures were constructed by treating

Table 10: Cohort evolutions: data vs. model

	Actual data		
Cohort age	Exporters	Total Exports	Average Exports
1 year	1	1	1
2 years	0.29	1.11	3.77
3 years	0.18	0.93	5.03
4 years	0.14	0.67	4.66
5 years	0.12	0.63	5.18
6 years	0.10	0.51	4.99
7 years	0.08	0.50	5.72
8 years	0.08	0.45	5.91
9 years	0.07	0.39	5.58
10 years	0.06	0.40	6.58

	Simulated data		
Cohort age	Exporters	Total Exports	Average Exports
1 year	1.00	1.00	1.00
2 year	0.61	1.73	2.84
3 years	0.35	1.34	3.81
4 years	0.19	1.81	9.50
5 years	0.10	2.29	22.74
6 years	0.06	2.12	34.43
7 years	0.05	1.89	39.69
8 years	0.04	1.69	43.23
9 years	0.03	1.89	63.69
10 years	0.02	1.46	65.17

Notes: Figures for cohorts aged 2-10 are expressed relative to corresponding figures for one-year-old cohorts.

the first shipment between a buyer and a seller as establishing a match, while the model does not.[24]

Table 11 compares the match exit rates observed in the actual data with those observed in the simulated data. These are broken down by match age, and by the size of the match's first-year sales. As with the figures in Table 6, this comparison is imperfect because of the differences in the way matches are defined in the two data sets. Nonetheless, the relatively high failure rates among first-year matches are replicated by the model, as is the tendency

[24] Specifically, since the data-based series count single-shipment buyer-seller encounters as matches, these series inflate the one-year-old firm and total export counts, while depressing mean exports among one-year olds. And the one-year old figures appear in the denominator of the ratios for all other years. It would of course be possible to generate model-based figures that include single-shipment buyer-seller meetings in the match counts. We plan to do this for the next draft of this paper.

Table 11: Match separation rates

	Actual data			
Match age	Quartile 1	Quartile 2	Quartile 3	Quartile 4
1 year	82.9	75.6	67.7	52.1
2 years	63.2	58.4	52.1	44.5
3 years	57.3	49.4	44.6	40.3
4 years	55	46.8	40.8	39.2
5+ years	49.7	43.7	37.6	36.7
	Simulated data			
Match age	Quartile 1	Quartile 2	Quartile 3	Quartile 4
1 year	0.60	0.88	0.89	0.63
2 years	0.27	0.29	0.31	0.27
3 years	0.30	0.32	0.33	0.30
4 years	0.31	0.28	0.20	0.32
5+ years	0.28	0.30	0.36	0.36

Notes: Figrues are percentages of the exporters in each age-initial size category that do not export during the following year.

Table 12: Exporter distribution by number of buyers

	share of exporters	
Number of buyers	actual data	simulated data
1	0.79	0.77
2	0.11	0.10
3	0.03	0.05
4	0.02	0.03
5	0.01	0.02
6-10	0.02	0.03
11+	0.02	0.01

Notes: Blah blah blah

for matches that begin from the largest sales quartile to fail less frequently than others. However, the drop in failure rates with cohort age is concentrated among two-year matches occurs solely between cohorts' first and second year in the simulated data, while it is a more gradual process in the actual data. Also, unlike the actual data, the exporters that begin in the smallest size quartile exhibit failure rates as low as those of the largest exporters. how explain this?

Finally, Table 12 reports the distribution of client counts across exporters in the actual versus simulated data. Overall the two distributions match up well, though the actual data contain more exporters with two clients (and fewer with more than two clients) than the model predicts.

6.2 The value of relationships

6.2.1 The value of clients

Two forces in our model make exporting decisions forward looking. First, each successful business relationship improves an exporter's visibility and reduces the cost of finding additional potential clients. Second, each match–successful or unsuccessful–conveys information about the scope of the market for the exporter's product. With Bayesian updating (equation 7), this means that early matches generate particularly valuable signals and worth pursuing even if they are not expected to generate significant earnings.