A Search and Learning Model of Export Dynamics*

Jonathan Eaton[†], Marcela Eslava[‡], David Jinkins[§], C. J. Krizan[¶] and James Tybout[‖]

(first draft: February 2014)

May 22, 2024

Abstract

Exporting abroad is much harder than selling at home, and overcoming hurdles to exporting takes time. Our goal is to identify specific barriers to exporting and to measure their importance. We develop a model of firm-level export dynamics that features costly customer search, visibility effects in finding buyers, and learning about product appeal. Fitting the model to customs records of U.S. imports of manufactures from Colombia we replicate patterns of exporter maturation. A potentially valuable intangible asset of a firm is its customer base and knowledge of a market. Our model delivers some striking estimates of what such assets are worth. Totaling across active exporters, the loss from total market amnesia (losing its current U.S. customer base along with its accumulated knowledge of product appeal) is US$ 14.2 billion, more than twice the value of annual Colombian manufacturing exports to the United States. About a quarter of this amount is from the loss of future sales to existing customers while the rest is from the cost of relearning about product appeal in the market and reestablishing visibility abroad. The frictions we estimate slow down the trade response to shocks. The 10-year response of total export sales to an exchange rate shock exceeds the 1-year response by 48 percent.

*We acknowledge the significant contribution of our late coauthor Jonathan Eaton to this research project. His work and insight were invaluable to its development, and we miss him deeply. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1109. 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, Zi Wang, Camilo Acosta and María Paula Alvarez for excellent research assistance, as well as Enrique Montes for expert data advice. Any views expressed are those of the authors and not those of the U.S. Census Bureau or the NBER. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release (DRB approval requests 5458 and 5550).
Corresponding author: David Jinkins, Porcelaenshaven 16A 1st floor, 2000 Frederiksberg, Denmark, email: dj.eco@cbs.dk

[†]Pennsylvania State University and NBER

[‡]Universidad de los Andes

[§]Copenhagen Business School

[¶]United States Department of Labor. This author's contribution largely occurred while employed by the U.S. Census Bureau's Center for Economic Studies.

[‖]Pennsylvania State University and NBER

1 Overview

Quantitative models of global economic activity rely on trade costs to explain why trade flows between countries constitute only a fraction of total production. But aggregate data provide little insight into what these costs are. A recent literature has turned to firm-level evidence to dig deeper into the nature of barriers to trade and how they influence the evolution of trade flows, proposing alternative mechanisms. Our objective is to advance this literature by studying trade flows at the buyer-seller match level, using firms' match histories, time intervals between new matches, and match-specific sales trajectories to make inferences about the role played by different trade barriers. In doing so, we aim to help bridge the gap between the literature on firm export dynamics and the literature on firm-to-firm connections.

Our primary dataset is the population of U.S. customs records describing imports of Colombian manufactured products over the period 1992-2009. This dataset allows us to identify individual U.S. importers ("buyers"), the Colombian exporters who supply them ("sellers"), and their interactions over time. We can observe, for example, Colombian firms entering the U.S. market in any given month, and see how each of their match-specific relationships evolves thereafter. We can also track counts of their successful and failed matches with U.S. firms. To complement these customs records from the US, we use in parallel firm-level (but not match level) data on the domestic activity of Colombian manufacturing firms, including those that do not export, merged with Colombian exports transactions data that indicate exports by firm and by destination.[1] This allows us to connect firms' exports to their domestic sales. Finally, we can decompose Colombia's aggregate exports into the contributions of individual cohorts of exporters.

These features of our data allow us to estimate a dynamic model that quantifies the relative roles of different types of trade barriers. Beyond allowing for fixed costs of sustaining matches, our model admits two frictions that affect firms' export behavior: endogenous search costs, and imperfect knowledge of foreign market conditions. It also allows for two forces that moderate these frictions: "learning effects," which improve market knowledge, and "visibility effects," which reduce search and matching costs for firms with an established market presence.

The estimated model implies that the fixed costs of maintaining a match are unimportant. By contrast, learning and search costs are crucial in explaining why few firms engage in exporting and the evolution of match counts among those that do export. It also implies a modest role for historic market visibility in determining the costs of searching for new buyers. The combined effect of learning and (modest) visibility effects implies that accumulating exporting experience is highly valuable, specially for new exporters.

Simulations of our model yield a number of additional findings. First, luck (in particular early success) can play an important role in determining firms' perseverance abroad. For

[1] Firms' individual matches in the domestic market could, in principle, have been inferred from Colombia's value-added tax records. However, these data are unavailable to researchers.

instance, accumulating five foreign buyers takes four years for a firm for which its three first matches evolved into more than one shipment. But a firm with identical productivity and product appeal takes 37 years if the first two matches generated only a single shipment. Second, a firm's accumulated knowledge of foreign market conditions, including its portfolio of foreign clients, can constitute an important component of its intangible capital. The estimated value of foreign market experience to Colombian exporters in our model is more than twice as high as annual Colombian manufacturing export revenues. Most of this value is attributable to knowledge about foreign market appeal accumulated through export experience. Finally, the aggregate export response to a change in the exchange rate process is slow. The 10-year response of exports to an exchange rate shock is 48% higher than the export response after 1-year.

1.1 The model

Our model is a single-agent continuous time representation of firm behavior. It characterizes two fundamental determinants of firms' sales in their home and foreign markets: the search process through which they identify potential buyers, and the evolution of their buyer-specific relationships, once formed. Below we briefly summarize each component.

As in most endogenous search models, firms must incur costs to connect with possible buyers. To reduce the expected time between new encounters, they must spend more. In our formulation, search costs also depend on sellers' exporting histories. As an exporter accumulates clients, its increased visibility may reduce the cost of finding more potential buyers.[2] This is the "visibility effect" in search.

Taking stock of search costs, firms choose their search intensity at each moment on the basis of the expected payoff from a new encounter. This depends in turn on the probability that the next potential buyer they meet will want their product, and the expected present value of the earnings stream that will result if it does.

A firm enters a foreign market knowing its own efficiency but imperfectly informed about its product's popularity there.[3] With each potential client it meets, it receives a signal about its product's appeal and it updates its behavior accordingly. This is the "learning effect." A series of buyers who want to do business signals a high level of buyer enthusiasm, encouraging the firm to search more intensely for new buyers, while a series of rejections indicates lack of buyer interest, leading the firm to reduce its search effort.

The present value of the earnings stream generated by a foreign relationship depends

[2] Additional clients may be also be harder to reach, so that having more existing customers means a higher cost of adding new ones. Our model allows for either possibility, but since we find that firms with more clients have an easier time adding new ones, we refer to the effect of existing customers on matching costs as "visibility effects."

[3] Efficiency and appeal to costumers are two distinct attributes of firms. Using data for the same Colombian manufacturing firms characterized in this paper, Eslava et al (2024) document that appeal is much more important than efficiency in determining firms' revenue growth overall (i.e. not only in exporting). Greater uncertainty about appeal to foreign vs. domestic costumers is a natural explanation of trade barriers.

upon the seller's efficiency, real exchange rate realizations, match-specific shock realizations, and the match's duration. These variables are mostly exogenous to the seller, but sellers optimally choose to end relationships when their expected future earnings stream falls below the fixed costs of retaining the customer. Unlike those in most trade models, "fixed exporting costs" are specific to each relationship rather than the foreign market.

1.2 Data features

Using the method of indirect inference, we target a number of data features to identify our model. Key for identification of the search cost function are the distribution of match success rates and the time interval between matches for exporters, each conditioned on firms' match histories. Matches that result in more than one shipment are deemed "successes." Key features for identification of firms' productivity distributions include the share of firms that export, the distribution of match-specific sales in the foreign market, and the cross-firm correlation of foreign and home sales. Finally, key for identifying the match-specific shocks and the fixed costs of match maintenance are match sales auto-regressions and the correlation of match death rates with match sales. The model overall fits these moments and others well.

To motivate our model, and to establish the relevance of our findings for other countries, we document some features of our data that we do not directly target. First, we show that patterns of exporter cohort maturation are typical of those reported in earlier studies: most new exporters exit the foreign market within their first year, but survival rates improve thereafter, with larger exporters tending to survive longer. Second, turning to the cross-exporter match count distribution, we confirm that our data exhibit the same "fat tail" shape that others have found. Finally, exploiting the time dimension of our data at the match level, we report several data features less commonly noted: match-counts are quite volatile through time for any given exporter, match separation rates are especially high among newer and/or smaller matches, and after their first year, matches show no systematic tendency to generate growing or shrinking sales. Our model qualitatively replicates all of these untargeted data features.

1.3 Relationship to the literature

A large body of empirical work characterizes firm dynamics in open economies. Alessandria et al. (2021) provide a recent review. We identify and briefly discuss how our paper contributes to the most relevant strands below.

Beginning with Dixit (1989) and Baldwin and Krugman (1989), one major strand of the literature has generated forward-looking exporter behavior by assuming that firms must incur a one-time sunk cost to break into a foreign market, and they must pay a per-period fixed cost to stay there.[4] Our model is only loosely related to this formulation. It incorporates fixed costs, but these are incurred to sustain individual matches rather than to sustain foreign market presence. Similarly, our model incorporates sunk costs in the sense that the costs of finding a client are incurred only once for any given match. But matches eventually fail, so firms that wish to maintain a foreign market presence indefinitely must incur ongoing search costs.

A second strand of the literature has modeled export dynamics as a process of foreign customer accumulation (Drozd and Nozal, 2012; Fitzgerald et al., forthcoming; Rodrigue and Tan, 2019; and Pivetau, 2021). We depart from these papers by exploiting match-level data to directly observe customer counts, rather than relying on sales to the destination market as a proxy for the size of firms' foreign customer base. Beyond providing a direct measure of a key object of interest, this approach allows us to study the evolution of individual matches, including their failure rate after the initial shipment. Critically, it also provides us with a direct measure of exporters' search intensity, namely, the time interval between their encounters with potential new clients.

Third, many papers have incorporated learning into firms' exporting decisions. In some, firms learn about foreign demand conditions (Nguyen, 2012; Ceberos, 2016; Li, 2018; Berman et al., 2019). In others, they learn about the reliability of their foreign customers or contract enforcability abroad (Aeberhardt et al., 2014; Albornoz et al., 2012; Araujo et al., 2016). In still others, firms learn how to penetrate foreign markets (Schmeiser, 2012; Chaney, 2014) or simply how to expand their sales in markets they have recently entered (Timoshenko, 2015; Ruhl and Willis, 2017; Monarch and Schmidt-Eisenlohr, forthcoming). These papers all offer evidence that learning is important, though they mostly formulate models in which learning is the main source of dynamics. We add to this literature by nesting learning about demand conditions in a more general model that admits other reasons for dynamic behavior. Further, while nearly all of these papers treat the unit of analysis as the firm-year-country-destination, we treat each new match within a foreign market as the delivering an informative signal.

Fourth, our work contributes to the literature on firm-to-firm international trade. Much of this is summarized by Bernard and Moxnes (2018). More recent contributions include Heise (2019), Bernard and Dhingra (2019), Sugita et al. (2023), Monarch (2022), Eaton et al. (2022a, 2022b), and Alviarez et al. (2022). Most of these papers assume frictionless matching up to an exogenous fixed cost to focus on assortative matching patterns, supplier switching, and/or pricing strategies.[5] Our focus is instead on micro-foundations of the matching process, so we assume random matching and exogenous mark-ups.

Fifth, our model delivers estimates of shipment frequencies and shipment sizes, and thus relates to earlier studies that do the same (Alessandria, et al., 2010; Kropf and Saure, 2014;

[4]Alessandria et al. (2021) thoroughly explore the properties of sunk cost/fixed cost models and provide references.

[5]Eaton et al. (2022a, 2022b) are exceptions. They also invoke search frictions and visibility effects to characterize buyer-seller matches, but both papers omit learning, match maintenance costs, and dynamic match-specific earnings streams.

Hornok and Koren, 2015; Bekes et al, 2017, Blum et al. 2019). But unlike some of these studies, it does not incorporate an optimization problem that generates a trade-off between these two variables, and we do not attempt to estimate fixed shipment costs. (Implicitly, frequencies are chosen by buyers and taken as exogenous by sellers.) We do let shipment hazards and average sizes vary across markets (home versus foreign), but without shipment-level data describing firms' domestic transactions we are forced to base these cross-market differences on estimates from Alessandria, et al. (2010).

Sixth, because we aggregate the behavior of individual firms to simulate export transition paths, our paper is related to papers that micro-found export dynamics in general equilibrium. Examples include Alessandria and Choi (2007, 2014, 2019), Ruhl (2008), Atkeson and Burstein (2010), Drozd and Nozal (2012), Burstein and Melitz (2013), Alessandria et al. (2014), Impullitti et al. (2013), Arkolakis (2015), Handley and Limao (2017), and Fajgelbaum (2020). Of course, by basing our simulations on a single-agent model, we miss the feedback effects on market aggregates that these models capture. Subject to this caveat, the payoff to our approach is that we are able to simultaneously quantify the role of standard search costs, learning, and visibility effects in shaping export responses.

Finally, our work contributes to the small literature that infers the value of firms' intangible capital from observable features of their life-cycles. Atkeson and Kehoe (2005) use an industrial evolution model to do this, treating firms' discounted lifetime earnings net of factor costs as offsetting (on average) the cost of intangible capital. Closer to our inferences, Monarch and Schmidt-Eisenlohr (forthcoming) infer the value of an exporter's match with an importer as the discounted expected profit stream it generates. Our approach differs from theirs in that we use firms' value functions to measure the capital losses firms would suffer if their portfolios of foreign buyers were to disappear. (These losses come partly from the search costs of replacing them, and partly from the earnings losses that accrue until a comparable portfolio is restored.)

2 Firm-Level Trade: Transaction-Level Evidence

In this section we document a number of facts in our data which will motivate our modeling choices. The exporting behavior of Colombian firms is typical and confirms facts known from the literature. For example, there is intensive churning and a plethora of one-time exporters with small shipments, and that the distribution of client numbers is strongly right-skewed.[6] In addition, our match-level data offer insights regarding the dynamics of the distribution of number of clients and the dynamics of single matches.

[6] Bernard et al. (2017), Bernard and Moxnes (2018), and Alessandria et al. (2021) reference and discuss many of the more recent studies. Studies of particular relevance to our paper are cited in the discussion below.

2.1 Data

We base our analysis on comprehensive data from the U.S. Census Bureau's Longitudinal Foreign Trade Transactions Database (LFTTD), which covers all commercial shipments into and out of the United States, extracting shipments from Colombia during 1992-2009. Each transaction record includes a date, the US dollar value of the product shipped, a 6-digit harmonized system product code, a quantity index, and, critically, an ID for both seller and buyer.

These IDs allow us to identify the formation and dissolution of business relationships ("matches") between an individual buyer in the U.S. and seller in Colombia. To identify the U.S. importer we use the buyer's Employment Identification Number (EIN).[7] To identify the Colombian exporter we used the manufacturer's identification code.[8]

We limit ourselves to transactions between non-affiliated trade partners and consider only imports of manufactures.[9] Our final data set, spanning the years 1992-2009, 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.[10]

In addition to U.S. customs records, we use establishment-level survey data from Colombia's national statistics agency, Departmento Administrativo Nacional de Estadistica (DANE), merged with exports transactions data from the Colombian tax and customs administration (DIAN). The DANE data provide annual information on the total sales and other characteristics of all Colombian manufacturing plants with at least 10 workers. [11] The merge with the DIAN transactions data allows us to break up total sales into domestic and exports, and the latter into total exports and exports to the US. We use these data to characterize the size distribution of Colombian plants, the fraction of Colombian plants that export to

[7]There are two ways to track U.S. importers in the LFTTD: 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 an EIN associated with every import transaction. An alpha maps to an entire firm, but the match rate between trade transactions and alphas is only about 80 percent (Bernard, Jensen, and Schott, 2009). We use EIN's to maximize coverage.

[8]This variable is based on Block 13 of CBP form 7501, the import declaration form. Customs brokers are required to input the data. 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 the code automatically from the name and address information entered in other fields. So this variable is sensitive to differences in how 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've conducted various checks on the matches based on this variable. Appendix B explains these checks.

[9]We thus exclude oil and coffee, which constitute the bulk of Colombian exports to the U.S. The National Federation of Coffee Growers centralizes coffee exports. A few players also dominate oil exports.

[10]Because of disclosure restrictions, as well as our exclusion of non-manufactures and trade between affiliated parties, we cover only a fraction of the total value of Colombian exports to the U.S. Table 18 in Appendix B compares patterns in our sample to patterns in aggregates from both U.S. and Colombian official sources.

[11]Since the Colombian data have been used widely in other studies, we don't provide further description here. We do not have data on the buyers other than the international transactions data.

the US, and, among exporting plants, the relationship between exports and domestic sales. Confidentiality restrictions keep us from merging the data from Colombian sources with the transactions data from the US side. We therefore use the two sources separately.

As is usual in the literature, the tables we report in this section treat a buyer- seller pair as “matched” during a particular calendar year if it executes at least one transaction. When estimating our model we will distinguish between all matches and “successful matches” or “established relationships”, which are those that go beyond a single shipment.

We now turn to some key patterns in the data that we seek to capture in our modeling and estimation. The features described in sections 2.2 and 2.3.1 show that our data reproduces patterns now standard in the literature, while sections 2.3.2 and 2.3.3 introduce more novel patterns.

2.2 Cohort maturation

Following Brooks (2006), Table 1 reports average patterns of maturation across cohorts of Colombian exporters of manufactured goods to the United States. Since maturation patterns vary little across individual cohorts, we’ve averaged across the seven cohorts entering each year between 1993 and 1999.

The second row of the Table implies that, on average, only 29 percent of the firms that entered initially (year one) continue exporting the next year (column 1), yet these survivors generated 11 percent more export revenue in year two than the entire cohort did in year one (column 2), because sales per survivor were 3.77 times as large in year two as sales per cohort member in year one (column 3).

Table 1: Average aggregates by cohort age

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: Based on LFTTD customs records, U.S. imports of manufactured goods from Colombia, 1992-2009. Figures for cohorts aged 2-10 are relative to the corresponding figure for one-year-old cohorts.

Subsequent rows apply to subsequent years of exporting by members of that cohort, all relative to the cohort's entry year.[12] After the first year, survival rates are much higher, with the average attrition rate falling below 40 percent between the second and third year and continuing to drop thereafter. Aggregate exports (column 2) decline more gradually than cohort membership in the typical new cohort after the second year, reflecting a combination of selection effects—larger exporters survive with higher probability—and growth in the sales of surviving firms. So after 10 years, the typical cohort has lost 94 percent of its initial exporters but still delivers 40 percent of its initial sales, with sales per 10-year survivor reaching 6.6 times sales per exporter in the cohort's first year (column 3). Taken together, these patterns conform to those that have been reported in a number of other studies.[13]

2.3 Patterns of buyer-seller matches

We now characterize buyer-seller matches during 1992-2009.

2.3.1 Number of clients distribution

Mirroring a now standard finding in the literature, the distribution of buyers per exporter in our data is heavily right-skewed.[14] The first row of Table 2 presents that distribution in our data (averaged over 1992-2009). 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 in a given year. However, since the remainder of the matches are polygamous, the average Colombian exporter sold to around 1.3 U.S. firms per year while the average U.S. buyer bought from around 2.3 Colombian firms per year. Only 3% of exporters had 4 or more buyers, and the fraction of those with over 6 buyers is negligible. [15]

[12] Appendix tables A.1-A.3 provide a breakdown of the numbers appearing in Table 1 by year of entry in the 1992-2009 period.

[13] Also studying Colombia, but using different time periods and data sets, Brooks (2006) and Eaton, et al. (2008) report similar tables with similar patterns. Papers that report relatively high exit rates for new exporters include Arkolakis (2016) for Brazilian exporters, Araujo et al. (2016) for Belgian exporters, Kohn et al. (2016) for Chilean exporters, Fitzgerald et al. (forthcoming) for Irish exporters, and Piveteau (2021) for French exporters. Papers that report growth in firm-level exports with market tenure include Araujo et al. (2016) for Belgian exporters, Ruhl and Willis (2017) for Colombian exporters, Fitzgerald et al. (forthcoming) for Irish exporters, Alessandria et al. (2021) for Colombian and U.S. exporters, and both Berman et al. (2019) and Piveteau (2021) for French exporters.

[14] Right-skewed buyers-per-exporter distributions are reported by Bernard et al. (2018) for Norway, Carballo et al. (2018) for Costa Rica, Ecuador and Uruguay, Benguria (2021) for Colombia, Eaton et al. (2022a) for the United States, and Eaton et al. (2022b) for France. The concentration of export sales among a handful of firms has been noted in many papers. Early references include Bernard et al. (2007) for the U.S. and Mayer and Ottaviano (2007) for France, Germany, Italy, Hungary and the UK.

[15] It is also worth noting that the number of Colombian exporters in our sample grew at roughly 2 percent per year, from 2,232 in 1992 to 3,300 in 2009, while the number of U.S. importing firms grew by 3 percent per year, from 1,190 to 2,079 (Appendix A, Table 17. 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.

Table 2: Client Distribution

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

Notes: Second line based on transition probabilities reported in Table 3

2.3.2 Transition probabilities

We now exploit our detailed data on matches to characterize the evolution of an exporter's relationships. We begin with the question of how stable are exporters' match counts. Except for the case of market exit, firm-level transitions across buyer counts have not been widely reported.[16] Table 3 gives the probability with which a Colombian exporter (with the number of clients in the first column) transitions to the indicated number of clients (along the rest of the corresponding row) the following year. We classify a firm that stops exporting but re-appears as an exporter sometime later in our sample period as "dormant", in contrast with a firm that doesn't appear again in our sample, which we classify as "out". We treat the pool of potential entrants as firms that ever appear as exporters in our sample.

Table 3: Transition Probabilities, Number of Clients

t *n*t+1	Out	Dormant	1	2	3	4	5	6-10	11+
Entrant	.	.	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

Notes: Based on LFTTD customs records, U.S. imports of manufactured goods from Colombia, 1992-2009. Figures are cross-year averages of annual transition rates. Confidentiality restrictions prevent us from reporting numbers for cells that are too sparsely populated.

[16] Eaton et al. (2022a) is an exception.

Like sellers' exporting stints (Table 1), most buyer-seller matches are short-lived. Even among those matches involving more than one shipment, the overall year-to-year death rate is roughly 40 percent, as we show later.

Among first-time exporters, roughly 93 percent sell to only one firm in their first year (first row).[17] Of single-buyer exporters, 62 percent don't export the next year, while only 6 percent go on to establish a larger number of relationships. For firms with 3 relationships in a year, 12 percent move up to a larger number the next year, but 67 percent lose clients on net. Firms starting with other client counts also, on average, move to a smaller number the following year. Hence, in addition to an enormous amount of churning among smaller exporters, we see a general tendency for firms to lose clients on net from one year to the next.

What does this pattern of entry and growth imply about the ergodic distribution of relationships? The second row of Table 2 gives the ergodic distribution implied by the transition matrix in Table 3 under the assumption that the number of new entrants equals the number that exit. The ergodic and actual (first row) distributions are very close, suggesting that over our period the transition process has been quite stationary. Both distributions are very nearly Pareto, reflecting the coexistence of many small scale exporters with a few "super-exporters."

2.3.3 Match maturation

We now turn to the issue of within-match dynamics. Table 4 sorts matches into quartiles according to first-year sales, reporting year-to-year separation rates. In addition to the very low overall survival rates, two patterns stand out. First, the higher the quartile of initial sales, the lower the annual separation rate for the next four years. Second, survival probabilities rise year after year across the four quartiles.[18]

[17] Many of these matches involve a single shipment. As we will show later, the overall match success rate (i.e., shipping to that buyer more than once) is roughly 41 percent.

[18] Monarch and Schmidt-Eisenlohr (forthcoming) also report very high separation rates for matches, especially in their fist year. Araujo et al. (2016) report a positive association between Belgian exporters' survival in a new market and their initial market-wide sales, though they do not use match-level data.

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

Notes: Based on LFTTD customs records, U.S. imports of manufactured goods from Colombia, 1992-2009.

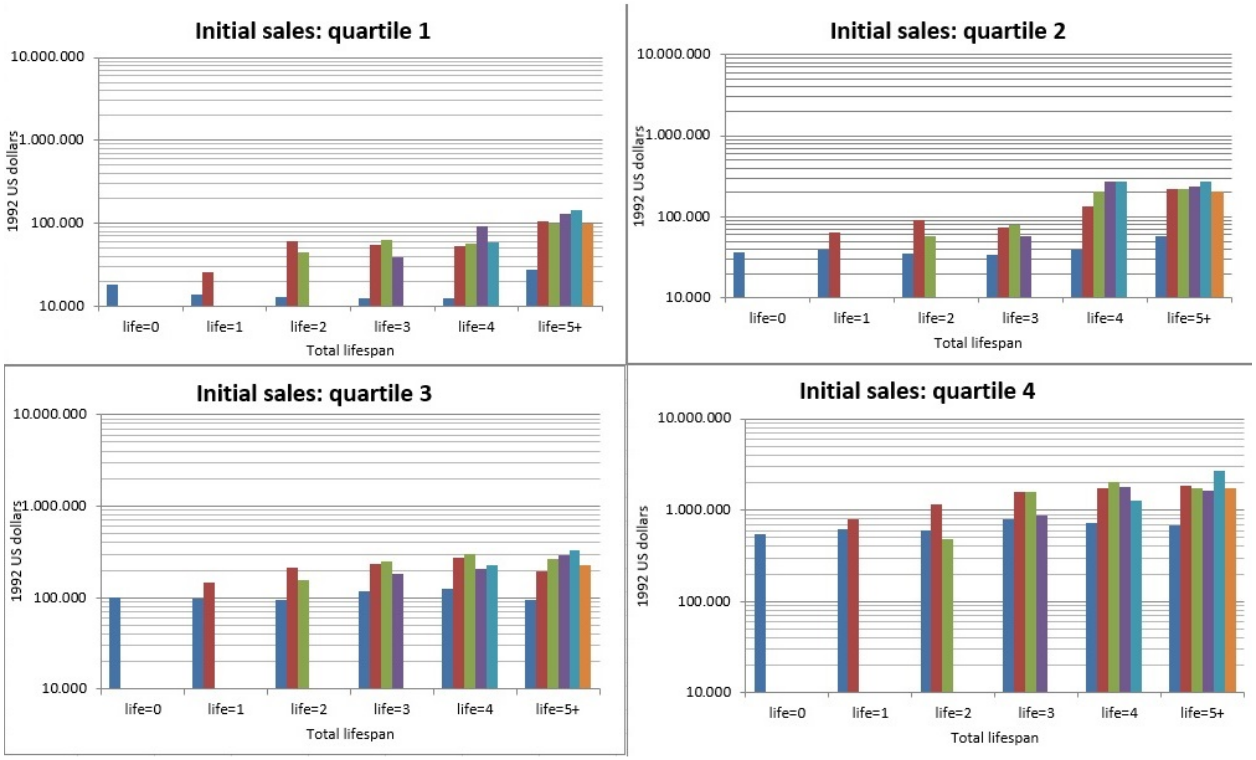
Figure 1 shows average annual sales per match, broken down by initial sales quartile, match age and total life span: less than one year (life=0), 1 to 2 years (life=1), and so forth. For each cluster of bars, the left-most bar corresponds to sales in the initial year of the match, the next bar corresponds to sales during the second, and so forth. The vertical axis is in log scale to facilitate comparisons across panels, but be mindful that this tends to mask growth within panels for the bare eye.

The first message is that initial sales are a good predictor of sales in subsequent years, conditioning on survival. Annual sales in later years rise monotonically with sales in the first year across quartiles. 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 (Bernard et al. (2017)). An analogous effect is 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 not so much for 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 of initial sales.

Within-match growth, although present for longer-lived matches, is not spectacular. It tends to remain below 25% annually for the average match.[19] It thus seems that the significant growth in firms' total exports is mainly due to changes in their portfolio of buyers.

[19]Consistent with this finding, Monarch and Schmidt-Eisenlohr (forthcoming) show that the relationship-specific sales of exporters to the U.S. grow only modestly after their first year (Table 7 and Figure 4). Also, using Norwegian customs records, Bernard et al. (2018) find that "[t]he buyer margin explains a large fraction of the variation" in destination-specific aggregate exports. Further, looking across Norwegian exporters in a given destination market, they find that "better connected sellers are not selling more to their median buyer . . . [than] less well-connected sellers."

Figure 1: Average sales per match, by initial size quartile, age, and total lifespan of match,



Notes: Based on LFTTD customs records for manufactured goods imported from Colombia, 1992-2009. Each group of bars represents a total match lifespan. Within a group, the left-most (second from left) bar represents per match average annual sales of matches in the initial(second) year, and so forth.

3 A Model of Exporting at the Transactions Level

We now develop a dynamic empirical model that is motivated by the data patterns presented in the previous section. Costly search will account for the tendency of firms to only gradually expand in export markets. Learning about product appeal abroad will help explain the shakedown process that each new cohort experiences. The dependence of search costs on successful export experience will help account for the right-skewed tail of client distributions in the data.

Our primary focus is understanding the dynamics of buyer-seller relationships between exporters from one country (in our case Colombia) and importers in a single foreign market (in our case the United States). Hence the model developed in sections 3.1, 3.2, and 3.3 applies to firms from a single source selling in a single foreign destination. We show in section 3.5 how to generalize the model to accommodate multiple foreign destinations, though we don't have the data to pursue this extension here.

With a single source and destination, our model provides a means of dissecting the dynamics of aggregate bilateral exports into, first, the sales of individual exporters and, then, into exporters' sales to individual clients.

In presenting the model we first consider the relationship between a seller and an individual buyer. Having derived the seller's return from a relationship with an individual buyer, we turn to its learning about the popularity of its product in that market, i.e., the chance that a potential buyer there likes its product. Finally, we characterize its search for buyers.

3.1 A seller-buyer relationship

A relationship is a sequence of shipments from a seller to a buyer. We start with the seller's profit from an individual shipment, and then show how the dynamics of these shipments determine the overall value of the relationship.

3.1.1 Profit from a single shipment

Several features of our model are standard. At any time t seller j can hire workers at a wage w_t in real local currency units, each of whom can produce φ_j units of output, where φ_j is time-invariant and known by the seller. Hence seller j's unit cost in local currency is w_t/φ_j. Selling at price p_{jt} in foreign currency units, profit in local currency is

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

where e_t is the indirect-quote exchange rate.

Goods markets are monopolistically competitive with each producer supplying a unique product. Once buyer i has matched with seller j, the buyer periodically buys from j. Each shipment generates revenue:

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

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} is a time-varying component of demand idiosyncratic to the ij relationship.[20]

We assume that the seller posts a non-negotiable price, charging the optimal markup over unit cost:[21]

[20] Since not all buyers necessarily face the same range of goods and hence the same aggregate price index P, we can treat i-specific components of the price index as P as embodied in y_{ijt}.

[21] Alternative specifications include bilateral bargaining between buyer and seller, as in Eaton et al. (2022), and pricing rules that recognize a link between current sales volume and future growth in customer base, as in Fitzgerald et al. (forthcoming) and Piveteau (2021). To keep our model tractable, and in view of Fitzgerald et al.'s (forthcoming) finding that exporters' prices don't covary with market tenure, we opt for constant mark-up pricing.

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

From (1), (2), and (3), the profit for seller j generated by an order from buyer i at time t is:

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

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

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

where x_t is common across all potential buyers in the foreign market. We can then write (dropping subscripts) the profit from a shipment as:

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

where the scalar Π allows us to normalize the means of $\ln x$ and $\ln y$ to zero.

Equation (4) is all we take from our specification of preferences and pricing behavior into the dynamics that follow. Any set of assumptions that deliver this simple multiplicative expression for a firm's profit from a shipment would serve us equally well.

3.1.2 Relationship dynamics

A match can dissolve for two reasons. First, it can simply end exogenously with a constant hazard δ (due, say, to the demise of the buyer or the buyer's finding a more suitable or cheaper substitute). Second, immediately after each sale to a particular buyer, the seller evaluates whether it's worth continuing the relationship. Doing so keeps the possibility of future sales to that buyer alive, but requires paying a fixed cost F.[22]

When deciding whether to maintain a match, the seller knows its own efficiency φ, the macro state x, and profit from the current sale, $\pi_\varphi(x, y)$ to the buyer in question. It can thus infer this buyer's current y and calculate the value of the match as:

$$\widetilde{\pi}_\varphi(x, y) = \pi_\varphi(x, y) + \max \{\widehat{\pi}_\varphi(x, y) - F, 0\} \tag{5}$$

where $\widehat{\pi}_\varphi(x, y)$ is the expected value of continuing a match that's currently in state (x, y).

[22]Fixed exporting costs are standard in the trade literature. Firms typically pay them per unit time to maintain their presence in a foreign market. Our model departs from this convention. Because it works in continuous time, and because it characterizes behavior at the match level, our firms incur fixed costs after each shipment if they wish to keep the associated match active. These costs can be interpreted to reflect maintenance of the account, technical support, or client-specific product adjustments. Colombian producers of construction materials interviewed for a related project (Domínguez et al, 2023) mentioned that a foreign buyer may request costly adjustments to a product or require special packaging.

The seller terminates this match if $\widehat{\pi}_{\varphi}(x,y) < F$.[23]

If the seller pays F to keep a match active, one of several events will next affect the match: with hazard δ the pair is exogenously dissolved; 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 $y' \neq y$.[24]

Let τ_r be the random time that elapses until one of these (match-specific) events occurs. Given that x and y are independent Markov jump processes, τ_r is distributed exponentially with parameter $\lambda^b + \lambda^X_x + \lambda^Y_y$, where

$$\lambda^X_x = \sum_{x' \neq x} q^X_{xx'} \tag{6}$$

and

$$\lambda^Y_y = \sum_{y' \neq y} q^Y_{yy'}, \tag{7}$$

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:

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

Before meeting a new buyer, the seller expects that the buyer will be in state y^s with probability $\Pr(y^s)$. The expected pay-off to forming a new match for a type-φ seller in market state x is thus:[25]

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

The function $\widetilde{\pi}_{\varphi}(x)$, which is identified by data on match-specific revenue streams, determines a seller's search intensity.

[23] Buyers' needs evolve, as do aggregate shocks in the respective market, so if the seller were to meet the same buyer after the current match dissolved, a new match with that buyer will ignore previous information. Such encounter, in any case, is a near-zero probability event.

[24] 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.

[25] In our numerical analysis 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.

3.2 Beliefs about product appeal and learning

We will consider two characterizations of firms' beliefs about their product's appeal in foreign markets. In one, our "benchmark" case, they update their beliefs as they acquire foreign market experiences. Thus, there is learning. In the other, our "known-θ^f" case, firms know the appeal of their products with certainty, even before they have met foreign buyers. The contrast between the two cases allows us to evaluate the importance of learning. Each is described below.

Beliefs in the benchmark model: A seller searches for buyers in the market anticipating that that some fraction $\theta \in [0,1]$ of them will be willing to do business with it. Given market state x, an encounter with a willing buyer generates the expected profit stream worth $\widetilde{\pi}_{\varphi}(x)$, while an encounter with an unwilling buyer generates a small sample shipment and nothing more.

Each seller enters the market with an unknown θ drawn from the (common knowledge) beta distribution:

$$b(\theta_0|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1},$$

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). Given its θ, the probability that a random sample of n potential buyers will yield a seller a interested customers is binomially distributed:

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

Hence, after meeting n potential buyers, a of whom were interested in its product, a seller's posterior beliefs about its θ are distributed:

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

where the factor of proportionality is the inverse of the integral of the right-hand side over the support of θ. A firm's expected success rate after a successes in n trials has the convenient closed-form representation:

$$\overline{\theta}_{a,n} = E\left[\theta|a,n\right] = \int_0^1 \theta p(\theta|a,n)d\theta = \frac{a+\alpha}{n+\alpha+\beta}. \tag{8}$$

As the beta distribution is the conjugate prior of the binomial, this posterior mean converges to

$$plim\left(\frac{a}{n}\right) = \theta$$

as n gets large.

Our formulation makes it crucial to distinguish between encounters or matches, of which a firm has had n, and the subset of those encounters that have succeeded in becoming established relationships, of which there are a. Section 4.2 explains how we make the distinction in the data.

In our formulation a firm learns something about its demand in a market with each encounter with a new potential buyer, successful or otherwise. We thus depart from other models with learning in which there is only zero or one signal per period, depending upon the firm's market participation (Timoshenko, 2015; Arkolakis et al., 2018; Fitzgerald et al., forthcoming). Our formulation creates an extra incentive for new entrants to search intensively, which we quantify in Section 6 below.

Beliefs in the known-θ^f model: The known-θ^f model simply amounts to replacing equation (8) with $\overline{\theta}_{a_j,n_j} = \theta_j$, where θ_j is firm j's true draw from $b(\theta \boldsymbol{j} \alpha, \beta)$. It is not nested by the benchmark model described above, although it involves the same parameters.

3.3 Searching for buyers

A seller continuously chooses a market-specific hazard s with which she encounters a potential buyer, incurring the instantaneous flow cost $c(s, a)$, which is increasing and convex in s.[26] In standard models of endogenous search, the cost of search depends only on the search intensity, s. But we follow Arkolakis (2010) in allowing the cost of search to also depend on the accumulated number of established relationships, adapting his formulation to a dynamic environment. How $c(s, a)$ varies with the number of successful matches a depends on the relative strength of different forces. The cost might fall with a as successful matches increase the seller's visibility with additional potential buyers. The cost might rise if the pool of easy-to-reach buyers becomes "fished out," as in Arkolakis' (2010) original formulation. We leave it to the data to decide the direction and magnitude of the effect, to which we refer as a "visibility" effect in search.

To derive the return to search, recall that when the foreign market state is x, a type-φ seller expects the value of a new successful match to be $\widetilde{\pi}_\varphi(x)$, and the seller believes the next encounter will be successful with probability $\overline{\theta}_{a,n}$. Hence the expected value of an encounter is $\overline{\theta}_{a,n}\widetilde{\pi}_\varphi(x)$

Let τ_s be the random time until the next search event, which could be either an encounter with a potential buyer or a change in the marketwide state x^f. Then the optimal search intensity s for a type-φ firm with foreign market search history (a, n) solves the Bellman

[26] Interviews conducted with Colombian exporters revealed a variety of activities firms pursue to first engage with potential buyers in a foreign market (Domínguez, et al, 2023). Activities included maintaining a foreign sales office; paying the exports promotion office to organize visits with prospective clients, 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 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.

equation:

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

(Recall that λ_x^X is given by (6).) Taking expectations over τ_s yields:

$$V_{\varphi}(a,n,x) = \max_{s} \frac{1}{\rho+s+\lambda_x^X} \Bigg[-c(s,a) + \sum_{x' \neq x} q^X_{xx'} V_{\varphi,}(a,n,x') + s \left\{ \bar{\theta}_{a,n} [\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1,n+1,x)] + (1-\bar{\theta}_{a,n}) V_{\varphi}(a,n+1,x) \right\} \Bigg] \tag{9}$$

Applying the multiplication rule for differentiation and using expression (9) for $V_{\varphi}(a,n,x)$, the optimal search intensity s^* satisfies:

$$\frac{\partial c(s^*,a)}{\partial s} = \bar{\theta}_{a,n} [\widetilde{\pi}_{\varphi}(x) + V_{\varphi}(a+1,n+1,x)] + (1-\bar{\theta}_{a,n}) V_{\varphi}(a,n+1,x) - V_{\varphi}(a,n,x) \tag{10}$$

That is, the marginal cost of search equals the expected benefit of a match $\bar{\theta}_{a,n}\widetilde{\pi}_{\varphi}(x)$ plus the expected value of the information and visibility it generates.

3.4 Mechanisms

Our model brings together several features found in the single-agent literature on exporting costs. Learning is captured by uncertainty about θ draws, costly matching is captured by the dependence of search costs on search intensity (s), and visibility effects are captured by the dependence of search costs on previous match successes (a). Combined, these features mean that the expected payoff to an additional match exceeds the potential earning stream it generates for two reasons: Each match informs the seller about the popularity of its product and, if successful, reduces the cost of finding additional buyers.

The accrual of experience in the model, through both learning and visibility, generates firm dynamics. All firms know their productivity draws, φ_j, but before acquiring foreign market experience, they share the same prior beliefs about their success rates there, θ_j. Therefore, upon entering the foreign market, high productivity firms expect high returns from successful matches, and they all search relatively intensively for clients. However, the link between productivity and search intensity weakens as firms acquire experience. Some will experience mostly failed encounters and lower their beliefs about θ_j, scaling back their search efforts in consequence. Others will enjoy a string of successes and revise their beliefs about θ_j upward, intensifying their search for new customers. Finally, since early signals are the most informative, the learning incentive to search falls off as firms meet more potential buyers.

3.5 Entering multiple markets

So far we've focused on firm entry into a single foreign market. We can accommodate firms' activity across multiple markets, designating a particular market by m. We treat seller j from a particular source as having an efficiency φ_j that applies across markets. In contrast, we assume seller j draws its product appeal measures, θ_j^m, independently, market by market.[27] Further, its search efforts, learning, and visibility effects are all market-specific.

Since a seller's' profits per match depend on its efficiency, φ_j, firm sales and profits tend to be positively correlated across markets. But the randomness of matches and separations, and the consequences of this randomness in search, makes this correlation imperfect, as the strings of successes and failures will be different.

Unlike foreign market entry and exit, rm entry and exit are not the focus of our analysis. Nonetheless, they merit brief discussion. There is a fixed population of potential firms in our model—each distinguished by a particular combination of productivity and product appeal. Each "enters" by making its first successful match in either the home or the foreign market, and "exits" when it goes without matches for at least 12 months. At this point, another firm steps up with the same productivity-appeal combination and begins searching for clients. One could call this re-entry, except the new firm does not inherit the knowledge or visibility of the firm it replaces.

4 Specification for Estimation

To adapt our theoretical framework to the data at hand we make a number of assumptions about destination markets, search costs, and the stochastic processes that generate exogenous state variables.

4.1 Destinations

Our source country is Colombia and our destination country is the United States. We use U.S. customs records to observe the individual shipments of each Colombian manufacturer to unaffiliated U.S. buyers, and we use Colombia's Annual Manufacturing Survey to observe each Colombian manufacturer's total sales to Colombian buyers. In what follows we indicate magnitudes specific to the foreign (U.S.) market with $m = f$ and specific to the home (Colombian) market with $m = h$.

[27]Eaton et al. (2011)'s static model also treats firm efficiency as common across markets but demand shocks as market-specific. While our dynamic model implies positive correlation across destinations in the cross-section, since we treat φ_j as time invariant, our model doesn't deliver any ergodic correlation in sales across countries over time unless we introduce demand shocks y that are temporally correlated across destinations.

Colombian firms typically don't export to the United States until they've sold in the home market for several years. Since our model implies that learning effects would be largely exhausted by then, we treat firms as aware of their product appeal in the home market θ^h by the time they enter our window of reference.

4.2 Matches and relationships

We treat a firm's first sale to a new buyer as a "match," which becomes an "established relationship" or "successful match" if and only if the firm sells to the buyer again. A one-off shipment is a "failed match."[28] For each Colombian firm that ever sells in the U.S. market in our period of observation we can thus keep track of its number n^f of encounters with U.S. buyers, and the number $a^f \quad n^f$ that succeed in generating additional transactions with the buyer after the first encounter.

4.3 Search costs

We generalize Arkolakis's (2010) formulation of search costs to allow for visibility effects, specifying the cost of searching with intensity s^m in market m as:

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

Here a^m is the number of previous successful matches the seller has had in market m, κ_0^m is a market-specific cost parameter, while the parameters κ_1 and γ are common across markets.

Several properties of this function merit note. First, the parameter γ governs how the number of previous successes affects the current cost of search, with $\gamma > 0$ implying a benefit (a positive "visibility effect") and $\gamma < 0$ consistent with a "fishing out" effect.[29] Second, the cost of search asymptotes to zero in market m as s^m goes to zero, so all firms have at least some visibility in each market, though it many well be too minimal to matter. Third, given the cumulative number of successful matches, a^m, the marginal cost of search increases with s at a rate determined by κ_1. Finally, since a^m is the cumulative number of successes in market m, visibility effects endure, even after a particular match is severed or while a firm isn't actively searching.

[28] This interpretation of the trade data is motivated by Rauch and Watson (2003) and Besedes (2008), who view small initial shipments as samples that potential buyers use to test potential suppliers.

[29] To limit the dimensionality of our computational problem, we assume that firms with more than $\bar{a}$ buyers have both (i) exhausted their learning effects and (ii) reap no additional visibility effects from further matches. We choose $\bar{a}$ to exceed the observed maximum a for 99 percent of sellers in the U.S. market, and we assume learning effects have played out after the first 20 encounters.

4.4 Exogenous Markov processes

Our model incorporates four exogenous Markov jump processes: x^h, x^f, y^h and y^f. We assume that the logs of these variables follow independent mean-zero Ehrenfest diffusion processes.[30] Because we do not observe match-level data on the domestic sales of Colombian firms, we impose that the idiosyncratic match shocks y^f and y^h follow the same process in both markets. However, we do not impose symmetry on the marketwide demand shocks x^f and x^h. This assumption accommodates the market-specific effects of exchange rate shocks, among other things.

Following Shimer (2005), we facilitate numerical simulation by discretizing the log support of each jump process $z \in \{x^h, x^f y^h, y^f\}$ into $2g+1$ possible values a distance Δ_z apart: $z \in \{-g\Delta_z, -(g-1)\Delta_z, .., 0, .., (g-1)\Delta_z, \Delta_z g\}$, where $g \in I^+$ and $2g+1$ is the number of mass points. We assume that, beginning from any state z, the hazard that a jump occurs is λ_z. And conditional on a jump occurring, we impose that the probabilities of possible outcomes are given by:

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

This implies that for any chosen grid size, the intensity matrices that characterize these processes are block diagonal with two unknown parameters: λ_z and Δ_z.

5 Estimation

We are now ready to explain how we bring our model to the data.

5.1 Constraints

We impose several constraints that reduce the number of parameters to estimate, to facilitate identification. First, because revenues at the match level are only observed for matches in the foreign market, we impose that the product appeal distribution parameters (α and β) and the fixed costs of sustaining a match (F) are common to both the domestic and foreign markets. For the same reason, we assume that the home shipment arrival hazard (λ_b^h) is twice as large as the foreign shipment arrival hazard (λ_b^f), and the foreign shipment-level profit scalar (Π^f) is twice as large as the domestic scalar (Π^h).[31] Next, we fix the

[30] The scalar Π in the profit function (4) absorbs the net effect of our mean-zero normalizations.

[31] We base this figure on Alessandria et al. (2010), who find that "for the typical product, international orders to tend to be about 50 percent larger and occur nearly half as frequently as domestic orders" (p. 2310). These figures are unfortunately specific to a U.S. steel wholesaler. Ideally we would have calculated this ratio from Colombian value-added tax records to track domestic transactions, but it was not possible to access these data.

jump hazard for idiosyncratic match shocks (λ_y) exogenously, given that it proved difficult to separately identify this parameter and the jump size (Δ_y).[32] In particular, we impose λ_y=0.33 per month. Third, we set the elasticity of demand η to 5 in both markets, implying that operating profits amount to 20 percent of sales revenue. Fourth, we assume that the cost function is quadratic ($\kappa_1 = 2$), since this parameter proved to be poorly identified. Finally, since endogenous and exogenous match deaths are difficult to distinguish in the data, we assume that the former are concentrated among young matches, and we set the exogenous match death hazard $\delta = 0.326$ to be the observed death hazard among matches more than three years old.

Together, these constraints force all cross-market differences in firm-level sales patterns to be absorbed through differences in the parameters we allow to be market-specific, namely, the parameters that govern market-specific aggregate demand shocks, the shipment arrival hazards, and the cost function scale parameters, κ_0^f and κ_0^h.

5.2 Estimated parameters and identification

Given the constraints reviewed above, the parameters we need to estimate can be collected into two groups: those that characterize exogenous random variables, and those that characterize technologies. In addition to the Markov processes for marketwide demands, the exogenous random variables include firm productivity draws, φ, firm appeal draws, θ^m, the Poisson process that generates shipment arrival rates for active matches, and the Poisson process that generates exogenous match separations. The technologies include the search cost function, $c^m(s^m, a^m)$, the profit-per-shipment function $\pi_\varphi^m(x^m, y)$, and the fixed costs of match maintenance, F.

[32]This was probably a consequence of the fact that most matches don't live for more than a few years.

Table 5: Parameters and Identification Strategy

	Parameters	Key targeted data features
Exogenous distributions		
marketwide shocks (x_t^h, x_t^f)	$\lambda_x^h, \lambda_x^f, \Delta_x^h, \Delta_x^f$	Autoregressions, marketwide expenditures, Table 6
idiosyncratic match shocks (y_{it})	Δ^y	Autoregression, match-level sales, Table 8, column (i)
firm productivity draws (φ_i)	σ_φ	Autoregression, match-level sales, Table 8 column (i), Cross-market sales correlation, Table 7 column (iii) Match exit regression, Table 7 column (iv)
Product appeal draws (θ_i^h, θ_i^f)	α, β	Success rate equations, Table 7 columns (i) and (ii), Cross-market sales correlation, Table 7 column (iii) Match-level sales dispersion, Table 8 column (i)
Match shipment process	λ_b^f	Shipment rates, Table 7 column (v)
Technologies		
Cost functions	$\kappa_0^h, \kappa_0^f, \gamma$	Match hazard equation, Table 7 column (iii); Export rate, Table 8 column (iv) Export sales share, Table 8 column (v)
Profit-per-shipment function	Π	Revenue per match, Table 8 column (i)
Match continuation costs	F	Match exit regression, Table 7 column (iv)

Together with the complete set of econometrically estimated parameters, Table 5 presents the key targeted moments for each parameter. To estimate these parameters we proceed in two stages. First, we fit the exogenous jump processes for the market-wide variables x^f and x^h using aggregate time series data. Then, exploiting our shipment-level data, we use indirect inference to obtain the remaining parameters.

Table 6: Market-wide expenditure dynamics

foreign market			home market		
AR1 estimates					
		$ln(x_t^f)$			$ln(x_t^h)$
1	μ^f	0.639	1	μ^h	0.875
		(0.239)			(0.188)
σ_{x^f}		0.1101	σ_{x^h}		0.0469
Jump process parameters					
λ_x^f		2.527	λ_x^h		0.875
Δ^{x_f}		0.069	Δ^{x_h}		0.050

Notes: Estimates for both countries are based on detrended annual data, 1991-2007. The Colombian series was deflated using the implicit price deflator constructed from DANE industrial survey data. The U.S. series was converted to nominal Colombian pesos using the prevailing exchange rate and deflated by the same price deflator. Parameter estimates in the lower panel are imputed from the AR1 estimates in the upper panel. Details are provided in footnote 33.

5.2.1 Estimating observable jump processes

To estimate the jump process parameters that govern the evolution of x_t^h and x_t^f, we use aggregate manufacturing expenditure data from Colombia and the U.S. Both are expressed in real pesos, so x_t^f moves partly in response to real exchange rate shocks. These series are, of course, observed in discrete time. To relate them to the continuous time jump processes in our model, we assume that each is generated by an independent Ornstein-Uhlenbeck process of the form:

$$dz = \quad \mu z dt + \sigma dW \tag{12}$$

where μ and σ are parameters and W follows a Weiner process. Then we exploit Shimer's (2005) result that (12) is the limit of the stationary jump process described above in Section 4.4 with hazard $\lambda = \mu g$ and jump size $\Delta = \sigma/\sqrt{\bar{\lambda}}$, given a proper scaling of parameters.[33]

The upper panel of Table 6 reports estimates of the AR1 processes for market-wide manufacturing expenditures in both countries; the lower panel converts these estimates to values for λ_x^m and Δ_x^m as discussed in footnote 33. The results imply that x^f jumps more frequently and by larger amounts than x^h. This is because innovations in x^f reflect movements in the real exchange rate as well as movements in dollar-denominated expenditures.

5.2.2 Estimating remaining parameters

To estimate the vector of remaining parameters, $\Lambda = \{\Pi^f, F, \alpha, \beta, \sigma_\varphi, \Delta_y, \lambda_b^f, \gamma, \kappa_0^f, \kappa_0^h\}$, we use the method of indirect inference (Gouriéroux and Monfort, 1996). For each

[33] The limit is constructed for $(\Delta\sqrt{\epsilon}, \lambda/\epsilon, g/\epsilon)$ as $\epsilon \to 0$. We approximate μ as 1 minus the root of the AR1, and we approximate σ as the root MSE of the AR1 residual variation. This establishes an approximate mapping from OLS estimates of the $AR1$ parameters for x_t^m, $m \in \{h, f\}$, to the corresponding jump process parameters, λ^{x_m} and Δ^{x_m}.

candidate Λ value, we first use the model to simulate the foreign and domestic transactions of an artificial sample of producers. Then, using these simulated transactions, we construct a collection of statistics that summarize the relationships we want our model to capture. Finally, searching the support of Λ, we choose the vector that makes the statistics based on our simulated data match the corresponding vector of statistics based on the sample data as closely as possible.

Algebraically, our estimator is:

$$\hat{\Lambda} = \arg\min \left[\bar{m} \quad m(\Lambda)\right]^{0} W \left[\bar{m} \quad m(\Lambda)\right],$$

where $\bar{m}$ is a column vector of statistics based on sample data, $m(\Lambda)$ is the analogous vector of statistics for a model-based simulated dataset, and W is a block-diagonal version of $var(\bar{m} \quad m(\Lambda))^{\quad 1}$, with each block corresponding to the moments from a particular regression or descriptive statistic.

Our $\bar{m}$ estimates and their standard errors can be found in Tables 7 and 8. Most are regression coefficients, but some are means, shares, or standard deviations. Note that for all regressions, we target mean values of the dependent variables rather than intercepts.

We now review our reasons for choosing the elements of $\bar{m}$ listed in Table 5 and the parameters they help most to identify, organizing our discussion by parameter. We are using 21 sample statistics to estimate a total of 10 parameters, and there is no simple mapping from the former to the latter. This is especially true in our setting, both because the model is highly nonlinear and because selection into export markets plays an important role in the determination of many relevant moments. Accordingly, our discussion serves only to give a general sense for the logic of our approach. Appendix C provides further details, including the "sensitivity matrix" (Andrews, et al., 2017) relating moment perturbations to associated adjustments in parameter estimates.

Distributions of time-invariant firm effects $(\alpha, \beta, \sigma_\varphi)$ There are three mutually independent firm effects in our model: θ^f, θ^h and φ. The first two are firm- and market-specific success rates drawn from a Beta distribution with mean $\alpha/(\alpha+\beta)$ and variance $\alpha\beta/[(\alpha+\beta)^2\ (\alpha+\beta+1)]$. They determine the fraction of potential buyers who would form business relationships with the seller, should they meet. The other firm effect, φ, is common across markets. Its log is drawn from a normal distribution with mean 0 and variance σ_φ^2. It controls for firms' productivity and any aspects of their product appeal that are common to both markets.

What features of the data identify α and β? Under our assumption that all unsuccessful meetings generate a single shipment, the success rate, a/n, is observable after each meeting for each firm in our sample. So, up to selection effects (to be discussed), the mean value of a/n is directly informative about $\alpha/(\alpha+\beta)$, and the variance of a/n is directly informative about $\alpha\beta/[(\alpha+\beta)^2\ (\alpha+\beta+1)]$. Further, the gradient of these statistics with respect to cumulative number of meetings, n, is informative about the extent to which cohorts skew toward high-appeal firms as they mature. That is, it helps us distinguish the benchmark model—which presumes Bayesian learning—from the known-θ^f model.

Estimates of all of these moments appear in Table 7. They imply that among firms with more cumulative experience (larger n), the average success rate is higher (column i) and the dispersion in success rates is lower (column ii). So in addition to providing evidence on the distribution of success rates across the population of firms, columns (i) and (ii) suggest that firms learn about their types as they acquire experience and they adjust their search efforts accordingly.

For identification of σ_φ, a large set of moments is relevant. First, the projection of firms' log exports on their log domestic sales (Table 8, column iii) is informative about the variance of φ relative to the variance of θ^f. The reason is that, unlike θ^f, φ draws are common to both markets.[34] Second, σ_φ plays a key role in determining patterns of export market participation and learning. Inexperienced firms self-select into exporting on the basis of their productivity alone, and large values of σ_φ put more of these firms above their participation threshold, reducing the fraction of firms that are near their entry-exit margin. So large σ_φ values affect the match sales distribution, and they make match deaths relatively insensitive to the signals firms accumulate about their product appeal. For these reasons, moments that characterize the size distribution and evolution of matches are all helpful in identifying productivity dispersion.

We summarize these features of the data with a match-level autoregression (Table 8, column iv) and a regression predicting match deaths (Table 7, column iv). Both control for partial year effects by including a dummy for new matches and both condition on firms' market tenure. Finally, because we want our model to capture the dependence of exit rates on match age that we documented earlier in Table 4, we include match age as an explanatory variable in the match exit regression. Coefficients on all of these control variables conform to expectations.

Shipment arrival process (λ_b^f) Aside from partial-year effects due to match births and deaths, the mean number of shipments per match-year is determined by the monthly shipment hazard, λ_b^f. This is directly related to the cross-match average log shipment count per calendar year, 1.176, which we report in column (v) of Table 7.

Shipment sales scalar (Π^f, Δ_y) Since our log productivity effects (φ) and macro shocks are normalized to mean 0, we require an additional parameter to allow the model to match the average log match sales level in the data. Π^f plays this role. And for the same reason that our match autogression (Table 8, column iv) and match death regression (Table 7, column iv) help to identify productivity dispersion, they help to pin down this parameter.

Match-specific shocks (Δ_y) Using only cross-match variation in sales, we would have difficulty distinguishing variation due to match-specific shocks from variation due to firm

[34] This figure is smaller than typcially reported in the literature. The reason is that it describes exports to the U.S. market alone. The correlation of firms' *total* exports (to all destinations) with their domestic sales is roughly 0.77.

effects φ and random shipment arrivals. But within-match time series variation in sales allows us to do so, and thus to identify Δ_y. This variation is captured by the coefficient on lagged sales and the mean-squared error estimates for our match autoregression (Table 8, column i).

Match continuation costs (F) Equation (5) implies that matches terminate endogenously whenever the expected future profit stream they generate is exceeded by the fixed costs F of maintaining them. This is most likely to occur when match sales are small and/or matches are new. And the nature and strength of this relationship depends on F. Accordingly, to help identify F we target the regression of match death rates on match sales and match age in equation (iv) of Table 7. Match age and a "new to market" dummy variable are also included in the regression to control for the fact that firms do not incur fixed continuation costs until they have made their first shipment, and thus tend to exit relatively frequently in their first year. Our estimates in Table 7 column (iv) essentially summarize the information in Table 4: death rates are high on average (0.395), fall with match revenues, and fall with match age–especially after the first year.

Search costs ($\gamma, \kappa_0^f, \kappa_0^h$) Search costs do not affect the likelihood that any given buyer-seller meeting will lead to a successful match, nor do they affect the earnings stream that successful matches generate, once formed. So γ, κ_0^f, and κ_0^h impact the moments we have discussed thus far only through selection effects, that is, by affecting the distribution of meetings across the various seller types in their various states.

The these selection effects are quite important, as can be seen in the Andrews et al. (2017) "sensitivity matrix" we report in Appendix C. But the influence of the search cost parameters is more directly reflected in firms' search intensities, s, which are governed by the search policy function (10). To exploit this source of identification, we proxy s with the inverse of the time elapsed until the next match, top-coded at 36 months.[35] And we regress the log of this measure on the log of the firm's cumulative number of meetings, $\log(1+n)$. (Refer to column (iii) of Table 7.)

Clearly, the average level of foreign market search, $\overline{\ln s}$, is informative about the search cost parameters, γ and κ_0^f. But the association between search intensity and experience (n) should also help with identification. The reason is that sample selection effects induce a positive correlation between experience and success rates, as discussed earlier, so the correlation between and s and n should be informative about the extent to which firms adjust their search in response to learning (inter alia).

We cannot use the same strategy to identify the home search cost parameter κ_0^h because we don't observe individual matches in the home market. Nonetheless, κ_0^h is identified separately from κ_0^f because cross-market differences in the search cost function are the only

[35]We observe this variable once after each meeting for each firm in the sample, excluding the last meeting in the last sample year. By assigning a 3 year interval to firms with no further in-sample matches, we effectively treat firms that appear to exit the export market as firms with very low search intensity. (We of course apply the same convention in our simulated data set when implementing our indirect inference estimator.)

means through which the model can replicate foreign market participation rates and the average share of output that exporters sell to foreign buyers. These are generated by equations (iv) and (v) of Table 8, respectively. They indicate that about 10 percent of Colombian manufacturing firms exported to the U.S. during our sample period, and among those that did so, their U.S. exports amounted to about 16 percent of their domestic sales.

Table 7: Success rates, match hazards, match endurance, and shipment rates

	(i) $\frac{a_{ij}}{n_{ij}}$	(ii) $u^2_{a_{ij}/n_{ij}}$	(iii) $\ln(s_{ij})$	(iv) $D_{ijt}^{exit\ match}$	(v) $\ln(m_{ijt})$
mean, dep. variable	0.413 (1.53e-3)	0.091 (2.65e-4)	-3.051 (5.30e-3)	0.395 (2.07e-2)	1.176 (1.42e-3)
$\ln(1 + n_{ij})$	0.093 (2.61e-3)	-0.056 (3.46e-4)	0.837 (8.20e-3)	–	
$D_{ijt}^{new\ to\ mkt}$	–	–	–	0.034 (1.17e-2)	
$\ln X_{ijt}^{f}$	–	–	–	-0.031 (1.58e-3)	
$\ln A_{ijt}^{match}$	–	–	–	-0.054 (9.02e-3)	
$\ln A_{jt}^{firm}$	–	–	–	-0.028 (6.53e-3)	
sample restrictions	$n_{ij} > 0$	$n_{ij} > 0$	$t < T$ 12	$t < T$ 12	$m_{ijt} > 0$
obs. (rounded)	35,800	35,800	38,500	23,500	87,000

Notes: All estimates are based on U.S. customs records. The unit of observation for all columns is seller j's i^{th} match. Variable definitions are as follows: s_{ij} = inverse of time interval between commencement of match i and the next meeting for exporter j; $D_{ijt}^{exit\ match} = 1$ if exporter $j^0 s$ i^{th} match dies in the current year; a_{ij} is the cumulative number of successes for exporter j at the time of match i; $D_{ijt}^{new\ to\ mkt} = 1$ if exporter $j^0 s$ i^{th} match is in its first year; A_{ijt}^{match} is the age of exporter $j^0 s$ i^{th} match; A_{jt}^{firm} is the market tenure of exporter j; X_{ijt}^{f} is sales generated by exporter $j^0 s$ i^{th} match; and m_{ijt} is the number of shipments per year to client i by exporter j.

Table 8: Home and foreign sales

	(i) $\ln X^f_{ijt}$	*(ii)* $\ln X^h_{jt}$	*(iii)* $\ln X^f_{jt}$	*(iv)* D^f_{jt}	*(v)* $\frac{X^f_{jt}}{X^f_{jt}+X^h_{jt}}$
Mean, dep. variable	10.665 (2.36e-3)	–	–	0.095 (8.71e-4)	0.162 (2.27e-3)
$D^{new\ match}_{ijt\ 1}$	0.328 (3.84e-3)	-	-	-	-
$\ln X^f_{ijt\ 1}$	0.826 (1.82e-2)	-	-	-	-
$\ln X^h_{jt\ 1}$	-	0.979 (2.94e-2)	-	-	-
$\ln X^h_{jt}$	-	-	0.344 (1.10e-1)	-	-
$\ln A^{firm}_{jt}$	0.063 (1.39e-2)	-	-	-	-
root mse	1.2079	0.447	2.477	0.294	0.277
sample restrictions	$X^f_{ijt}, X^f_{ijt\ 1} > 0$	$X^h_{jt}, X^h_{jt\ 1} > 0$	$X^f_{jt}, X^h_{jt} > 0$	$X^h_{jt} > 0$	$X^f_{jt}, X^h_{jt} > 0$
observations	25,400	93,729	10,325	113,656	10,838

Notes: Column *i* is based on U.S. customs records. Column *ii* is based on data from Colombia's Annual Manufacturing Survey (EAM). Columns *iii-v* are based on the EAM merged with Colombian customs records. Variable definitions are as follows: $D^{new\ match}_{ijt\ 1} = 1$ if exporter j^0s i^{th} match is in its first year; A^{firm}_{jt} = foreign market tenure of exporter j; X^f_{ijt} = foreign sales volume generated by exporter j^0s i^{th} match; X^f_{jt} = total foreign sales to the U.S. generated by exporter j; X^h_{jt} = total home sales volume generated by firm j; and $D^f_{jt} = 1$ if firm j is an exporter to the U.S.

5.3 Structural parameter estimates and fit

Table 9 reports estimates of the parameter vector Λ for the benchmark model (columns 1-2) and the known-θ^f model (columns 3-4). Both models fit well overall, as can be seen in Figure 2, which plots the simulated moments, $m(\Lambda)$, against their data-based counterparts, $\overline{m}$. (For ease of visualization we have also plotted the 45° line).[36] However, both overestimate the average log match hazard of -3.05 and underestimate the gradient of the log match hazard with respect to cumulative matches of 0.84, implying a flatter (albeit positive) relationship between meeting frequencies and previous successes than we observe in the data. More strikingly, unlike the benchmark model, the known-θ^f model overestimates average log match revenues (14.00 versus 10.67), underestimates the share of firms that export (0.04 versus 0.10), and underestimates the fraction of output exported among these firms (0.10 versus 0.16). Accordingly, we easily reject it in favor of the benchmark model using the Rivers and

[36] Appendix C reports the data points behind these plots.

Table 9: Structural parameter estimates

		Benchmark model		Known-θ^f model	
	Parameter	*value*	*std. error*	*value*	*std. error*
log of profit scalar	$\ln \Pi$	-19.635	(1.51e+00)	-23.092	(1.094e-00)
fixed cost per shipment	$\ln F$	-3.784	(4.36e+00)	-7.004	(1.43e+00)
First θ distribution parameter	α	0.032	(5.02e-03)	0.019	(3.09e-03)
Second θ distribution parameter	β	0.192	(5.42e-02)	0.128	(1.77e-02)
demand shock jump size	Δ^y	0.044	(3.10e-01)	0.053	(5.81e-03)
shipment order arrival hazard	λ_b	1.014	(3.70e-02)	1.042	(1.24e-01)
std. deviation, log firm produc.	σ_φ	2.384	(1.72e-01)	2.942	(7.84e-02)
visibility effect parameter	γ	0.046	(1.29e-02)	0.061	(8.45e-03)
home search cost scalar	$\ln \kappa_0^h$	5.132	(5.35e-04)	4.331	(1.70e-05)
foreign search cost scalar	$\ln \kappa_0^f$	15.161	(2.36e-08)	18.020	(1.93e-11)
log of loss function	$\hat{Q}$	11.691		12.087	

Notes: Both models were fit using indirect inference, targeting the statistics in Tables 7-8 and using a block-diagonal weighting matrix based on their regression-specific covariance matrices. Standard errors were constructed using the Delta method.

Voung (2002) test statistic.[37]

The known-θ^f model has trouble fitting these moments because it has no good mechanism with which to explain the prevalence of transitory, small scale exporters. In the benchmark model, all inexperienced young firms above a productivity threshold actively search for clients abroad. Most learn after a few meetings that they are low-θ^f types and stop their searching.[38] But until they do, they constitute a large fraction of the population of exporting firms. In contrast, in the known-θ^f model, such low-θ^f firms never bother to participate in foreign markets.

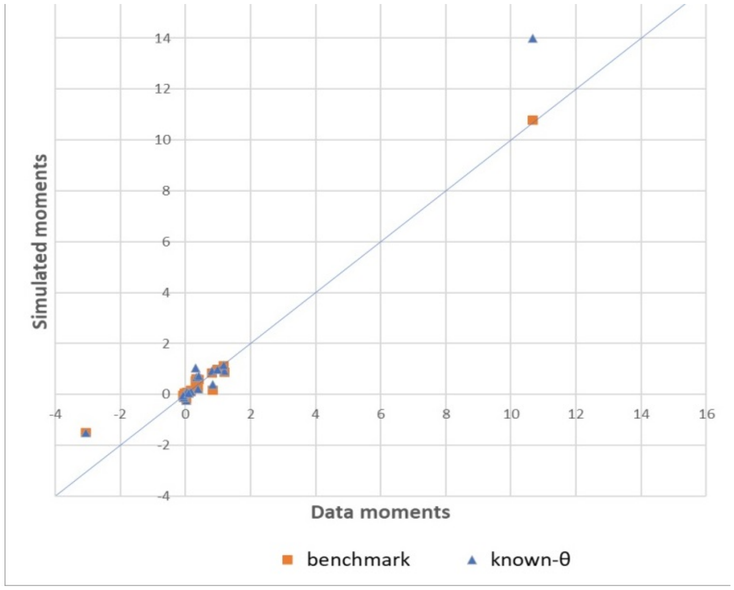
Since the export market participation rate is a heavily weighted target, the known-θ^f model uses other (non-learning) mechanisms to generate export market participation.[39] Most importantly, it chooses a higher value for σ_φ than the benchmark model, making high productivity firms relatively common. Similarly, by deviating from the benchmark values of α and β, the known-θ^f model put relatively more potential firms in the right-tail spike of the θ_f

[37]This statistic is for comparing non-nested models. It takes the form $T_n = (\sqrt{n}/\hat{\sigma}_n)\left[\hat{Q}^1 \quad \hat{Q}^2\right]$, where $\hat{Q}^1$ and $\hat{Q}^2$ are the MSM fit metrics for the two models, and $\hat{\sigma}_n^2$ approximates $var\left[\hat{Q}^1 \quad \hat{Q}^2\right]$. This statistic has a standard normal distribution under the null $E(\hat{Q}^1) = E(\hat{Q}^2)$. With model 1 the benchmark and model 2 the known-θ^f variant, we get T_n =-1,736.6 (treating the weighting matrix W as nonstochastic). Two caveats apply. First, since the targeted regression coefficients are based on a variety of samples, it's not obvious what sample size n we should use for this statistic. We use a very conservative approximation to the number of firms we base our inferences on ($n = 1000$), Second, this test statistic doesn't recognize randomness in the fit statistics due to simulation.

[38]Our estimates of α and β imply a bimodal θ^f distribution, with 75-80 percent of potential firms having virtually no chance of success, and 7-8 percent having success probabilities near 1.

[39]The heavy weight reflects this statistic's small standard error—refer to Table 8, column *iv*.

Figure 2: Simulated versus data-based moments



Notes: Simulated figures generated using 50,000 potential firms over a 50 year period.

distribution.[40] But by increasing the share of high-φ, high-θ^f firms among active exporters, these parameter adjustments compromise the model's ability to match other moments in Tables 7 and 8. In particular, average log match revenues become too large, and average log meeting hazards become too high.[41] To partly offset the former effect, the known-θ^f model chooses a much smaller profit-per-shipment scalar, Π. And to better match search intensities among active exporters, it chooses a higher search cost parameter, κ_0^f. These adjustments improve the model's fit, but it still overstates mean log match sales and the mean meeting hazard (Figure 2 and Table 21, Appendix C).

Despite their differences, both models deliver similar estimates of many parameters and thus agree on some basic messages. Note first that the visibility effect parameter γ is positive and significant in both models, albeit fairly small. This implies firms that have accumulated successful relationships find it relatively cheap to maintain any particular matching hazard. We will quantify the importance of this effect on exporter dynamics in section 6 below.

Next consider the monthly shipment hazard, λ_b. For both models, estimates of this parameter imply that matches will generate about 12 shipments per calendar year, provided they are active for the entire period. This seems high relative to the average log annual shipment count of only 1.18 (Table 7, column v), which implies about 3 shipments per match year. Yet both models are able to match this figure almost exactly because matches

[40]Relative to the benchmark model, the known-θ^f model puts an extra one percent above the θ_f=0.95 threshold.

[41]Evaluating the simulated moments at the known-θ^f model estimates of σ_φ, α, and β, but otherwise using the benchmark parameter values, the mean log match sales is 17.77 and the mean log search hazard is 2.97. In contrast, as reported in Tables 7 and 8, the data-based estimates of these moments are 10.67 and -3.05, respectively.

Table 10: Cohort evolution: simulated data

Cohort age	Firm count	Total exports	Avg. exports
1-yr old	1.000	1.000	1.000
2-yr old	0.540	0.817	1.511
3-yr old	0.382	0.660	1.728
4-yr old	0.298	0.469	1.576
5-yr old	0.175	0.299	1.711
6-yr old	0.168	0.241	1.431
7-yr old	0.133	0.244	1.836
8-yr old	0.123	0.231	1.879
9-yr old	0.104	0.231	2.231
10-yr old	0.087	0.188	2.156

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

turn over rapidly: within any calendar year, most matches are either beginning or dying.

Finally, note that the fixed cost of sustaining a match, F, are very small for both models. This implies that exporters rarely cut off buyers because their orders are too small. [42]

5.4 Explaining the untargeted stylized facts

It is instructive to ask how well the benchmark model explains the data-based patterns reported above in Section 2.1, none of which have been directly targeted. Tables 10, 11, and 12 repeat the statistics reported in Tables 1, 2, and 4, respectively, using data simulated for 50,000 potential exporters with the benchmark model. Here, to be consistent with those tables, we follow all matches, regardless of whether they are unsuccessful (single shipment) or successful.

Consider Table 10 first. Qualitatively, the firm count and average export patterns match up. The largest drops in the number of exporters occur during a cohort's first two years, with cohort size dropping gradually thereafter. Also, exports per firm rise over the entire time horizon, mainly reflecting selection effects. However, this effect is not as strong in our simulated data as it is in the sample data. Thus, contrary to the data-based pattern in Table 1, total cohort exports fall in the first year rather than rising. (Thereafter they exhibit the same gradual decline we observe in the data.)

Table 11 reports the distribution of client counts across exporters implied by our model. Qualitatively, this distribution exhibits the same shape as in Table 2, inasmuch as the most common type of exporter has a single client and exporter frequencies decline monotonically with match counts. And the model replicates the data by giving roughly 99 percent of all exporters 10 clients or fewer. However, the fraction of firms with a single client is substantially

[42] As noted earlier, F is conceptually different from the fixed exporting costs that appear in most previous studies. (The standard assumption is that they are incurred to maintain foreign market presence.) Nonetheless, our estimate is consistent with the fixed cost estimates reported by Das et al. (2007), which are also small and insignificant.

Table 11: Exporter distribution: simulated data

Number of buyers	share of exporters	cumulative share
1 match	0.462	0.462
2 match	0.234	0.696
3 matches	0.131	0.827
4 matches	0.070	0.898
5 matches	0.043	0.941
6-10 matches	0.050	0.991
>10 matches	0.009	1.000

Notes: Figures give the ergodic distribution of current buyer counts across exporting firms.

Table 12: Match separation rates: simulated data

Match age	quantile 1	quantile 2	quantile 3	quantile 4
1-yr old	0.849	0.938	0.929	0.386
2-yr old	0.210	0.402	0.453	0.339
3-yr old	0.375	0.237	0.388	0.348
4-yr old	0.612	0.077	0.306	0.380
5+ yr old	0.425	0.175	0.556	0.317

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

smaller in Table 11 than in Table 2, and the fraction of firms with 2-4 clients is substantially larger.

Finally, Table 12 reports match exit rates by cohort age and initial size, redoing Table 4 with simulated data. The model replicates the higher failure rates among first-year matches, and the tendency for matches that begin in the largest sales quartile to fail less frequently than others. However, Table 4 shows that the match failures drop off smoothly as one moves from the smallest size quartile to the largest, while our simulated pattern in Table 12 implies that the dropoff occurs between the third (next to largest) size quartile and the fourth (largest).

6 Quantification of trade frictions

We now use our estimates to quantify trade frictions and illustrate other model implications. We model three types of trade friction: fixed costs of sustaining matches, endogenous search costs, and imperfect knowledge about foreign market conditions. We find that fixed costs are not an important barrier to trade. It costs a firm less than one dollar to maintain a match after a shipment. We therefore focus on information and search frictions in this section.

6.1 Search costs, learning, and visibility

For a Colombian firm with no prior success in the foreign market, our results imply that a search intensity sufficient to yield an average of one new match per year has an annualized cost of $c^f(1,0) =$ \$30,029.[43] Because search costs are convex, an expected yield of one new match every other year costs only $c^f(0.5,0) =$ \$ 7,507. To put these results in context, the simulated average value of a shipment for an exporter in steady state is \$ 45,871, and an active match results in one shipment per month.[44]

Search costs vary as the firm meets buyers and establishes successful relationships with them (the visibility effect). Our estimated visibility parameter $\gamma = 0.04$ implies that the costs of search fall somewhat as a firm racks up successes. A firm with ten successful foreign matches pays $c^f(1,10) =$ \$ 28,378 annually for an expected match every year, roughly five percent less than the cost before its first match. Accumulating ten successful foreign matches is not easy. The probability of a firm's establishing a successful relationship with any given buyer in the foreign market is drawn from a beta distribution, which we estimate to have mean $\alpha/(\alpha+\beta) = 0.14$ and variance $\alpha\beta/\left[(\alpha+\beta)^2(\alpha+\beta+1)\right] = 0.09$. Hence, before acquiring export market experience, a firm expects that roughly 1 in 7 encounters with a potential buyer will lead to a successful business relationship. All of this implies that, although there is indeed a visibility effect (i.e. a reduction of searching costs as the firm accumulates exporting experience), its magnitude is modest.

Besides providing visibility, exporting experience also allows the firm to learn about the appeal of its products. The left panel of Figure 3 assesses the combined importance of the learning and visibility effects by evaluating the continuation value as the firm adds new matches to its exporting history, for different success histories of those matches. The panel shows the perceived continuation value after each additional meeting for firms with our highest discretized productivity type.[45] These values depend on the firm's belief at each moment about its success probability $\bar{\theta}^f$, which in turn depends on the number of matches (n) it's had already. We show three extreme histories: an unbroken string of successes (blue line, with $a = n$), an unbroken string of failures (orange line, where $a = 0$), and alternating success and failure (yellow line, $n \quad 2a$).

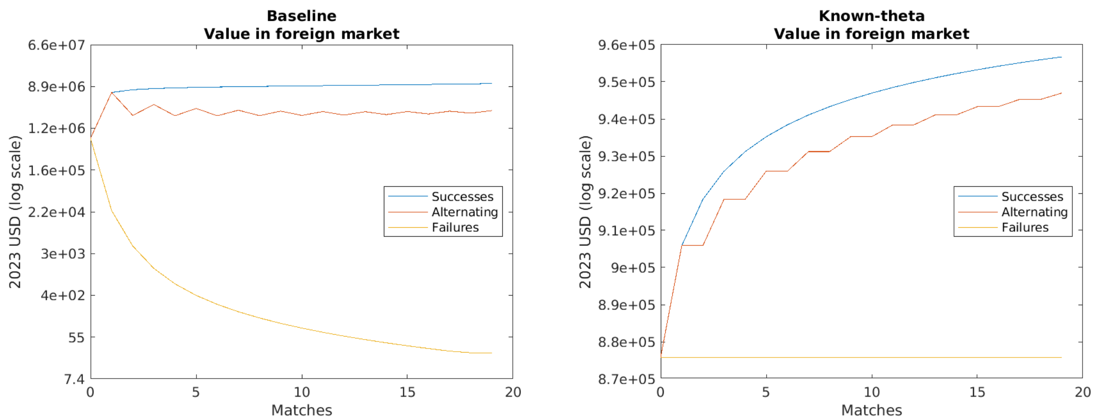
A new match increases or decreases perceived value depending on whether the match is successful. Perceived value is quite sensitive to signals from encounters with foreign buyers, especially for neophyte exporters who haven't yet formed networks or learned anything about their appeal. As all firms have the same prior before their first match, they all have the same perceived value, in this case \$ 739,784. The first match has the biggest impact on continuation values, and most of the impact of additional information is gone by the twentieth match. For example, if its first match is a success, the firm's value jumps to \$ 6,668,762.

[43] All figures in this section are in 2023 US dollars.

[44] In this section, all simulated results are from 2000 independent runs of 50 years at our best fit parameterization. We call results from the last ten years of the simulation our steady state.

[45] In our baseline simulation, all exporters are from the highest discretized productivity type.

Figure 3: Log continuation values conditioned on match history



Notes: Continuation value trajectories for high productivity firms, separately for baseline and known-theta models. We plot values for only successful matches, alternating success and failure, and only failures. Continuation values exclude expected profit streams from current relationships.

Failures quickly erase value. After two failed matches, a firm perceives its value to be only \$ 4,273. As discussed above, the estimated distribution of success probability is bimodel with mass close to zero and close to one, so only a few signals are required for a firm to be confident it is either high or low appeal.

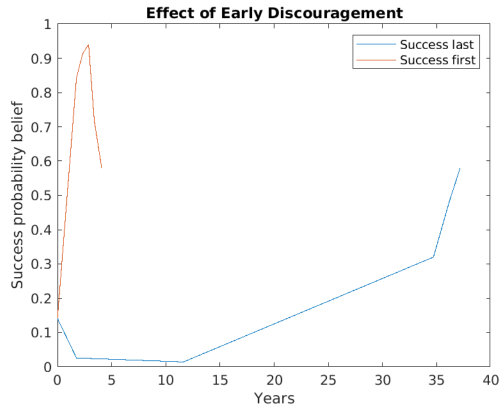
6.2 Foreign-market amnesia

Another way to quantify the value of exporting experience is to consider the value lost to Colombian exporters if they were to suddenly forget their foreign-market experience – to come down with foreign-market amnesia. To do this, we measure the value of a firm's experience in the export market, which can be separated into the expected future profits generated by current business relationships and the knowledge and visibility a firm has acquired for its future relationships.

Among active exporters, the simulated average discounted value of foreign market operations is \$ 6.2 million. But among firms with the same productivity and market appeal distribution as simulated exporters, but which have no foreign market experience, we get an average value of foreign market access of US\$ 654 thousand. The difference between these figures of US\$ 5.6 million is the average discounted value of foreign market experience, including both future sales from current relationships and the expected value of future relationships. Thus, the value of foreign market experience is around nine times larger than the initial value of the firm. This large difference is consistent with research that finds high costs of entering the exporting market for the first time (e.g. Das et al, 2007).

The value of foreign market experience can be split into the expected discounted future sales from existing clients and the combined value generated by visibility and accumulated knowledge about its foreign market appeal θ^f. Expected future sales to existing customers

Figure 4: Evolution of success probability belief



Notes: Beliefs of a high-productivity exporter over success probability, and expected learning time. Top line is three successes followed by two failures. Bottom line is two failures followed by three successes.

comprise US$ 1.3 million of the difference in values. The remaining US$ 4.3 million reflects the value of visibility and knowledge about foreign market appeal. Multiplying this value by the 3300 Colombian exporting firms observed in 2009, the last year of our sample, implies a total value of US$ 14.2 billion, or about 242% percent of annual manufacturing export revenues from the United States.[46] In other words, accumulated visibility and knowledge of product appeal in the Colombian non-affiliated manufacturing sector has a value more than twice as high as annual Colombian manufacturing revenues from the U.S. market.

6.3 The role of luck

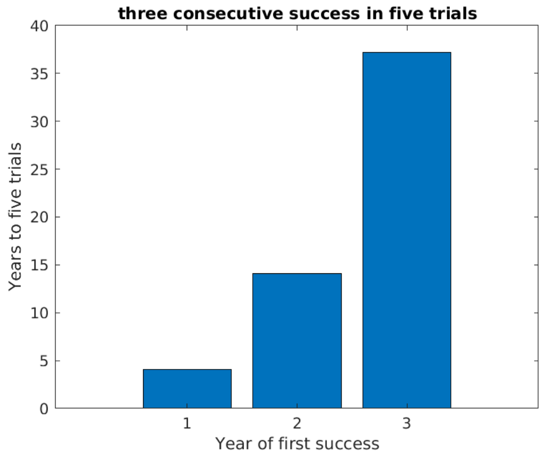
Learning and visibility effects can cause two *ex ante* identical firms to have different long-term experiences in an export market, depending on whether their early matches succeed. Because match histories affect continuation values and search costs, luck also affects the intensity with which a firm searches for new clients.

Figure 4 plots the evolution of beliefs about θ^f for two high-productivity firms (where, for simplicity, searching with intensity s results in waiting exactly $1/s$ years for the next match). The top line plots perceived appeal from a sequence of three successes followed by two failures, and the bottom line perceived appeal from two failures followed by three successes. These sequences of successes and failures have the same probability, and they end (and start) with the same belief.

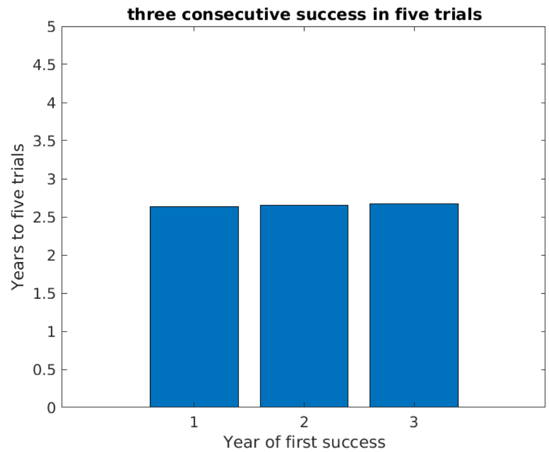
The two trajectories have markedly different durations. The time it takes a firm to reach five meetings depends heavily on luck, in particular the placement of the successes. It takes four years if the successes come first, but 37 years if the failures come first. This is also

[46] According to our data from the Colombian manufacturing survey combined with customs records, Colombian manufacturing export revenues were US$ 5.87 billion in 2007 (in 2023 dollars). This figure includes both affiliated and unaffiliated manufacturing exports.

Figure 5: Time to five meetings by placement of three consecutive successes



(a) Baseline



(b) known-θ^f

shown in Figure 5 , which plots the expected time to five meetings when three consecutive meetings succeed and the other two fail. The x-axis is the meeting at which the firm gets its first success. For example, if it's one, the first three meetings are successful, and the next two fail. In the baseline model (panel a), the discouraged firm takes almost ten times as long to get to five meetings (37 rather than four years) because it reduces its search intensity dramatically after its first negative signal. It takes the firm ten years to receive its second match.

How does luck affect outcomes when there is no learning? Since firms know their success probabilities, the known-θ^f version of Figure 4 (not pictured) is simply two horizontal lines with height θ^f. But the lengths of these lines still depend on match histories through the visibility effect on search. This is captured in panel b of Figure 5, for a firm with known $\theta^f = 0.64$. In this known-theta case, it takes 2.63 years to reach five meetings if the successes come first and 2.67 years if they come last. The very modest difference between the good- and bad-luck cases is due to our small visibility effect.

Comparing the baseline and known-theta cases, luck is more important if firms must learn about their product appeal abroad. That is, it is the presence of learning that makes exporting performance so heavily dependent on past luck.

6.4 Comparing learning and visibility

The patterns we've depicted so far in this section reflect both visibility and learning effects. Because our estimated visibility effect is modest, however, most of that combined value is attributable to learning. To see this, compare the left and right panels of Figure 3, which depict the exporting continuation value for different match histories. The right panel is constructed in the same manner as the left panel, but for an economy in which firms know their true foreign product appeal from the start. (Note that the units on the vertical axis differ in the two panels of this figure.) We consider firms with $\theta^f = 0.64$, corresponding to the 88th percentile of active exporters in our simulated data. Using the estimated "known-θ^f" policy function in Table 9, we simulate the continuation values of of such a firm in the highest discretized productivity type. As in the left panel, the right panel shows the histories of only successes (blue line), only failures (yellow line), and alternating success and failure (orange line) for such a firm. In contrast to our baseline model, failure does not affect the perceived value of a firm with a known θ, since it learns nothing from its experience. Successes still increase the value of the firm, but only because of the visibility effect. The change in value as a firm garners experience is much smaller in the known-theta case than what we observe in the baseline model, which corroborates that the estimated visibility effect is small.

7 Exchange rate shocks

The trade frictions we measure slow the economy's reaction to a shock. To explore the magnitude of this effect, we characterize the dynamic response of exports to an exchange rate shock. We consider a 20 percent devaluation of the Colombian Peso against the US Dollar. We first look at what happens to an average firm in our simulation. We then turn to aggregate trade dynamics, breaking down the response into different margins of adjustment. We conclude with a discussion of the implied short and long-run trade elasticities.

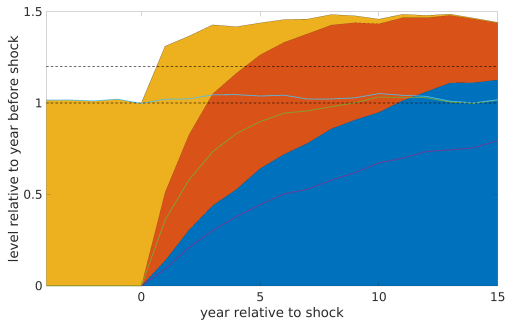
Our results for individual exporters are averages of 2000 50-year simulations in which the permanent devaluation shock unexpectedly hits at the end of the 25^{th} year. In this section, we consider the total value of the firm including both the value of current and expected future relationships. Calculated this way, the mean value of foreign market access of an active exporter in the year before the shock hits is 5.9 million dollars. The value of the mean exporter increases 37 percent in the year after the shock. Average exporter value jumps more than the mechanical 20 percent from the devaluation, because after the shock exporters search harder and expect to both learn and become visible more quickly.

How do these changes in value translate into aggregate export dynamics? Figure 6 summarizes the results of simulating the aggregate export trajectories associated with the exchange rate devaluation.[47] The permanent 20 percent peso devaluation occurs at the end of the 25^{th} year (marked period 0 in the figure). Panels a and b break down aggregate export sales and total matches into three segments: contributions from matches created before period 0 (yellow area), contributions from matches created after period 0 with exporters that existed in period 0 (red area), and matches formed after period 0 with exporters that entered after period 0 (blue area). Panel c breaks down the number of active exporters into those active before period 0 (red area) and those who entered after period 0 (blue area). The solid lines show how the boundaries between the shaded areas would have evolved, had

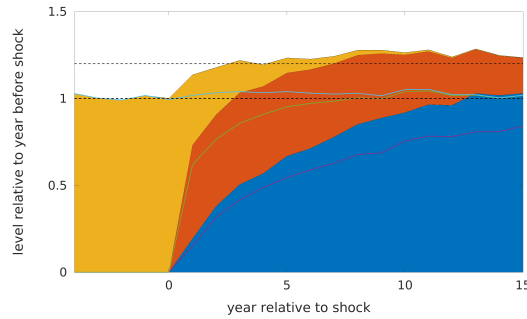
[47] Our single-agent model abstracts from interactions between exporters in the foreign market. But since Colombia constitutes a small share of the U.S. spending, such general equilibrium effects are likely negligible.

there been no permanent devaluation. The dotted lines are fixed at 1.0 and 1.2 to reflect the mechanical increase in sales from the exchange rate shock. All figures are normalized by levels in the year before the shock.[48]

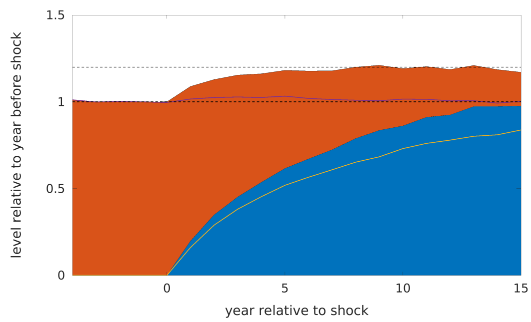
Figure 6: Baseline response to a permanent devaluation: export aggregates



(a) Total sales



(b) Total active matches



(c) Total active exporters

Notes: Figures depict aggregate responses to a permanent 20 percent real devaluation at time 0. Shaded areas in panels a and b reflect contributions of matches that existed at time 0 (yellow), matches formed after time 0 by exporters that were active at time 0 (red), and matches formed after time 0 by exporters that entered the foreign market after time 0 (blue). Panel c depicts incumbent exporters active before time zero (red), and exporters entering after time zero (blue). Thin solid lines show patterns in the absence of the shock. Dotted lines reflect levels the year before the shock and 20 percent growth relative to those levels. All series are averages across 2000 simulations.

Panel *a* describes total export sales. The rapid turnover is striking, both with and without the shock. In the devaluation scenario, after seven years incumbent exporters lose more than half of their market share. This churning in our simulation is a reflection of the churning in the raw data, as documented in Tables 1 and 3. Despite rapid match turnover, adjustments to the new exchange rate take time to play out fully. In the year after the shock, total sales expand 31% relative to the year before the shock (Panels *a*). After five years, total sales are

[48] Piveteau (2021) provides similar graphs that inspired Figure 6. To highlight the role of learning and endogenous match separations, we use a decomposition that distinguishes matches to new exporters from others. Piveteau (2021) distinguishes the consumer margin, the extensive margin, and an aggregate valuation effect.

Table 13: Simulated Trade Elasticities

Time since shock	1 year	3 years	10 years
Sales	1.486	1.953	2.072
	(0.110)	(0.140)	(0.171)
Matches	0.698	1.086	1.285
	(0.135)	(0.155)	(0.178)
Exporters	0.466	0.787	0.961
	(0.073)	(0.098)	(0.119)

Notes: All elasticities are based on 2000 simulations of a 20 percent real devaluation of the Colombian peso. Bootstrapped standard errors are in parentheses.

44% higher than in the year before the shock. The reaction grows over time both because the total number of exporters adjusts, and also because the number of matches per exporter increases, as is implied by panels *b* and *c*. After 10 years, these effects add an extra 48% percent to the first-year response.

While it is not immediately apparent from Figure 6, the shares of new matches and new exporters in each aggregate are nearly invariant to the devaluation shock. Hence, if we were only interested in the rate at which new exporters displace incumbent exporters, or the rate at which new matches displace incumbent matches, it would matter very little whether we were analyzing the aftermath of a permanent devaluation or a period without any regime switching.

Table 13 reports the short, medium, and long-run trade elasticities implied by the permanent 20% real peso devaluation with standard errors based on cross-simulation standard deviations in parentheses.[49] Our long-run sales elasticities resemble Piveteau's (forthcoming) and Boehm et al.'s (2020), but they are substantially lower than the long-run elasticities typically generated by calibrated general equilibrium models (e.g., Alessandria and Choi, 2014; Alessandria, et al., 2018).[50]

8 Summary

Research exploiting customs records has generated a robust set of stylized facts regarding firm-to-firm trade dynamics: First, most exporters are inexperienced, ship small amounts, and have few foreign clients. Second, the typical buyer-seller relationship lasts only a year or two, so business connections evolve rapidly, and it's common to see firms with only a few clients cease exporting entirely, giving way to the next entering cohort of inexperienced

[49]Figure 6 shows that effects have stabilized after around ten years. Our long-run estimates are thus obtained at a 10-year horizon./medskip

[50]Alessandria and Choi (2014) use a symmetric 2-country dynamic model with endogenous firm creation, capital accumulation, fixed exporting costs, and iceberg costs. Analyzing movement from a global eight percent tariff to free trade, they find trade elasticity increases of about five in the short run and eight in the long run, which is reached in 5-8 years. In a similar model, but with firms' exporting costs depending upon their incumbency, Alessandria et al. (2018) estimate a short-run trade elasticity of four and a long-run elasticity of 11.55. Their model generates transition dynamics over 10-15 years.

exporters. Third, however, each new cohort contains a small number of firms that survive and grow many times faster than aggregate exports. They do so not by selling more to the same clients, but by finding new customers.

We document these patterns for Colombian manufacturers shipping to the United States, and develop a continuous-time model to account for them. Firms wishing to export must engage in costly search to find potential buyers, who may either reject their products or form finite-lived business relationships with them. Buyers who form business relationships with exporters send them favorable signals about the appeal of their products and, in doing so, encourage those exporters to search more intensely for additional buyers (learning effects). Successful business relationships also reduce sellers' search costs by improving their visibility (visibility effects). Finally, sellers' search intensities depend both on their permanent idiosyncratic characteristics and on market conditions.

Fit using the indirect inference, the model replicates both targeted and untargeted patterns in customs records and allows us to quantify several types of trade costs, including the cost of searching for potential clients and the cost of maintaining business relationships with existing clients. It also allows us to estimate the visibility effect of previous successes on the costs of meeting new clients, and to characterize the cumulative effects of learning on firms' search intensities and intangible capital stocks.

While our model delivers similar long-run elasticities to other one-sided search models, the presence of learning means that it takes longer to reach the long run. Aggregate export dynamics are slowed by search frictions which limit the ability of exporters to connect with new potential buyers. The reason the long-run effect of learning and visibility is modest is that they are most important among newer exporters, which account for a small share of total export volume. Overcoming these frictions is costly, and the older firms that have done so have accumulated substantial intangible capital.

References

Aeberhardt, R., I. Buono, and H. Fadinger (2014): "Learning, Incomplete Contracts and Export Dynamics: Theory and Evidence from French Firms." *European Economic Review* 68: 219–249

Albornoz, Facundo, Hector Calvo Pardo, Gregory Corcos, and Emanuel Ornelas (2012) "Sequential Exporting." *Journal of International Economics* 88: 17-31.

Alessandria, George, Costas Arkolakis, and Kim Ruhl (2021) "Firm Dynamics and Trade." *Annual Review of Economics* 13: 253-280.

Alessandria, George and Horag Choi (2007) "Do Sunk Costs of Exporting Matter for Net Export Dynamics?" *Quarterly Journal of Economics* 122(1): 289-336.

Alessandria, George and Horag Choi (2014) "Establishment Heterogeneity, Exporter Dynamics, and the Effects of Trade Liberalization." *Journal of International Economics* 94: 207-233.

Alessandria, George and Horag Choi (2019) "The Dynamics of the U.S. Trade Balance and Real Exchange Rate: The J Curve and Trade Costs?" NBER Working Paper 25563.

Alessandria, George, Horag Choi, and Kim Ruhl (2018) "Trade Adjustment Dynamics and the Welfare Gains from Trade." Working Paper, The University of Rochester.

Alessandria, George, Joseph Kaboski, and Virgiliu Midrigan (2010) "Inventories, Lumpy Trade, and Large Devaluations." *American Economic Review* 100: 2034-2339.

Alessandria, George, Sangeeta Pratap, and Vivian Yue (2014) "Export Dynamics in Large Devaluations." Working Paper, Federal Reserve Bank of Philadelphia.

Alviarez, Venessa, Michele Fioretti, Ken Kikkawa, and Monica Morlacco (2022) "Two-Sided Market Power in Firm-to-Firm Trade," Working paper, University of British Columbia, Sauder School of Business.

Andrews, Isiah, Matthew Gentzkow, and Jesse Shapiro (2017) "Measuring the Sensitivity of Estimated Parameters to Estimation Moments." *Quarterly Journal of Economics* 132(4): 1151-1199.

Araujo, Luis, Emanuel Ornelas and Giordano Mion (2016) "Institutions and Export Dynamics." *Journal of International Economics* 98: 2-20.

Arkolakis, Konstantinos (2010) "Market Access Costs and the New Consumers Margin in International Trade." *Journal of Political Economy* 118(6): 1151-1199.

Arkolakis, Konstantinos (2016) "A Unified Theory of Firm Selection and Growth" *The Quarterly Journal of Economics* 131(1): 89–155

Atkeson, Andrew and Ariel Burstein (2010) ”Innovation, Firm Dynamics, and International Trade.” *Journal of Political Economy* 118(3): 433-484.

Atkeson, Andrew and Patrick Kehoe (2005) “Modeling and Measuring Organization Capital.” *Journal of Political Economy* 113(5): 1026-1053.

Baldwin, Richard and Paul Krugman (1989) ”Persistent Exchange Rate Effects of Large Exchange Rate Develulations.” *Quarterly Journal of Economics* 104: 635-654.

Békés, Gábor, Lionel Fontagné, Balázs Murakozy, and Vincent Vicard (2017). ”Shipment Frequency of Exporters and Demand Uncertainty.” *Review of World Economics* 153(4): 779-807.

Benguria, Felipe (2021) “The matching and sorting of exporting and importing firms: Theory and evidence.” *Journal of International Economics* 131(1): 103430

Berman, N., V. Rebeyrol, and V. Vicard (2019) ”Demand learning and Firm Dynamics: Evidence from Exporters.” *Review of Economics and Statistics* 101(1): 91-106.

Bernard Andrew, Bradford Jensen, Stephen Redding and Peter K. Schott (2007) ”Firms in International Trade.” *Journal of Economic Perspectives* 21(3): 105-130.

Bernard, Andrew, J. Bradford Jensen, and Peter K. Schott (2009) ”Importers, Exporters, and Multinationals: A Portrait of Firms in the U.S. that Trade Goods,” in Timothy Dunne, J. Bradford Jensen and Mark J. Roberts eds. *Producer Dynamics*, University of Chicago Press.

Bernard, Andrew, Esther Ann Boler, Renzo Massari, Jose-Daniel Reyes, and Daria Taglioni (2017) ”Exporter Dynamics and Partial-Year Effects.” *American Economic Review* 107(10): 3211-3228.

Bernard, Andrew and Swati Dhingra (2019). ”Importers, Exporters and the Division of the Gains from Trade.” Working paper, Tuck School of Business at Dartmouth.

Bernard, Andrew and Andreas Moxnes (2018) ”Networks and Trade.” *Annual Review of Economics* 10(65): 65-85.

Bernard, Andrew, Andreas Moxnes, and Karen Helene Ulltveit-Moe (2018) ”Two-Sided Heterogeneity and Trade.” *Review of Economics and Statistics* 100(3): 424–439.

Besedes, Tibor (2008). “A Search Cost Perspective on the Formation and Duration of Trade.” *Review of International Economics* 16(5): 835-849.

Blum, Bernardo S., Sebastian Claro, Kunal Dasgupta, and Ignatius J. Horstmann (2019). ”Inventory Management, Product Quality, and Cross-country Income differences.” *American Economic Journal: Macroeconomics* 11(1): 338-388.

Boehm, Christoph, Andrei Levchenko and Nitya Pandalai-Nayar (2020). “The Long and Short (Run) of Trade Elasticities.” Working Paper, The University of Michigan.

Brooks, Eileen (2006) “Why Don’t Firms Export More? Product Quality and Colombian Plants” *Journal of Development Economics* 80: 160-178.

Burstein, Ariel and Marc Melitz (2013) “Trade Liberalization and Firm Dynamics,” in *Advances in Economics and Econometrics Tenth World Congress*. Applied Economics, Econometric Society Monographs. Vol. 2. Cambridge, UK: Cambridge University Press.

Carballo, Jeronimo, Gianmarco Ottaviano, and Christian Volpe Martincus (2018) ”The Buyer Margin of Firms’ Exports.” *Journal of International Economics* 112: 33-49.

Cebreros, Alfonso (2016) ”The Rewards of Self-discovery: Learning and Firm Exporter Dynamics.” Banco de México Working Paper 2016-08.

Chaney, Thomas (2014) ”The Network Structure of International Trade.” *American Economic Review* 104(11): 3600-3634.

Das, S., Roberts, M.J. and Tybout, J.R. (2007). ”Market Entry Costs, Producer Heterogeneity, and Export Dynamics.” *Econometrica* 75: 837-873.

Dixit, Avinash (1989) ”Entry and Exit Decisions under Uncertainty.” *Journal of Political Economy* 97: 620-638.

Dominguez, Juan Carlos, Jonathan Eaton, Marcela Eslava, and James Tybout (2023) ”Search and Learning in Export Markets: Evidence from Interviews with Colombian Exporters.” *Review of International Economics* 31:1093–1116

Drozd, Lukasz A. and Jaromir B. Nosal (2012) “Understanding International Prices: Customers as Capital.” *American Economic Review* 102(1): 364-395.

Eaton, Jonathan, Marcela Eslava, Maurice Kugler and James Tybout (2008). “Export Dynamics in Colombia: Firm-Level Evidence,” in Elhanan Helpman, Dalia Marin and Thierry Verdier, eds., *The Organization of Firms in a Global Economy*, Cambridge, MA: Harvard U. Press.

Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2011). ”An Anatomy of International Trade: Evidence From French Firms” *Econometrica* 79(5): 1453-1498.

Eaton, Jonathan, David Jinkins, James Tybout, and Daniel Xu (2022a) ”Two-sided Search in International Markets.” NBER Working paper 29684.

Eaton, Jonathan, Samuel Kortum and Francis Kramarz (2022b). ”Firm-to-Firm Trade: Imports, Exports, and the Labor Market.” NBER Working Paper 29685.

Eslava, M., J. Haltiwanger, and Nicolas Urdaneta. (2024). The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand and Wedges. *The Review of Economic Studies*, 91(1): 259–300.

Fajgelbaum, Pablo D. (2020). ”Labor Market Frictions, Firm Growth, and International Trade.” *Review of Economic Studies* 87(3): 1213-1260.

Fitzgerald, Doireann, Stefanie Hallerz, and Yaniv Yedid-Levi (forthcoming). ”How Exporters Grow.” *Review of Economic Studies*.

Gouriéroux and Monfort, 1996. *Simulation-Based Econometric Methods*. New York: Oxford U. Press.

Handley, Kyle and Nuno Limão (2017). ”Policy Uncertainty, Trade, and Welfare: Theory and Evidence for China and the United States.” *American Economic Review* 107(9): 2731-2783.

Heise, Sebastian (2019) “Firm-to-firm relationships and the pass-through of shocks: Theory and evidence,” Staff Report, No. 896, Federal Reserve Bank of New York, New York, NY

Hornok, Cecília and Miklós Koren (2015). ”Per-Shipment Costs and the Lumpiness of International Trade.” *The Review of Economics and Statistics* 97(2): 525-530.

Impullitti, Giammario, Alfonso. Irarrazabal, and Luca Opromolla (2013) ”A Theory of Entry into and Exit From Export Markets.” *Journal of International Economics* 90: 75-90.

Kohn, David, Fernando Leibovici, and Michal Szkup (2016) ”Financial Frictions and New Exporter Dynamics.” *International Economic Review* 57(2): 453-486.

Kropf, Andreas and Philip Sauré (2014). ”Fixed Costs per Shipment.” *Journal of International Economics* 92(1): 166-184.

Mayer, Thierry and Gianmarco Ottaviano (2007) “The Happy Few: the Internationalization of European Firms: New Facts Based on Firm-Level Evidence.” *Bruegel Blueprint Series, Volume III,* University of Bologna.

Monarch, Ryan (2022) “It’s Not You, It’s Me.” *Review of Economics and Statistics* 104(5): 909–928.

Monarch, Ryan and Tim Schmidt-Eisenlohr (forthcoming) “Longevity and the Value of Trade Relationships.” *Journal of International Economics*

Li, Shengyu (2018) “A Structural Model of Productivity, Uncertain Demand, and Export Dynamics.” *Journal of International Economics* 115: 1-15.

Nguyen, Daniel (2012) "Demand Uncertainty, Exporting Delays and Exporting Failures." *Journal of International Economics* 86: 336-344.

Piveteau, Paul (2021) "An Empirical Dynamic Model of Trade with Consumer Accumulation." *American Economic Journal: Macroeconomics.* 13(4): 23-63.

Poschke, Markus (2018) "The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology." *American Economic Journal: Macroeconomics* 10(3): 1-41.

Rauch, James and Joel Watson (2003) "Starting Samll in an Unfamiliar Environment." *International Journal of Industrial Organization* 21: 1021-42.

Rivers, Douglas and Quang Vuong (2002). "Model Selection for Nonlinear Dynamic Models." *Econometrics Journal* 5: 1-19.

Rodrigue, Joel and Yong Tan (2019). "Price, Product Quality, and Exporter Dynamics: Evidence from China." *International Economic Review* 60(4): 1911-1955.

Ruhl, Kim (2008) "The International Elasticity Puzzle." Working Paper, The University of Wisconsin.

Ruhl, Kim and Jonathan Willis (2017) "New Exporter Dynamics." *International Economic Review* 58(3): 703-725.

Schmeiser, Katherine N (2012). "Learning to Export: Export Growth and the Destination Decision of Firms." *Journal of International Economics* 87(1): 89-97.

Shimer, Robert (2005) "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *The American Economic Review* 95(1): 25-49.

Sugita, Yoichi, Kensuke Teshima, and Enrique Seira (2023) "Assortative Matching of Exporters and Importers." *The Review of Economics and Statistics* 105(6): 1544–1561.

Timoshenko, O. A. (2015): "Learning Versus Sunk Costs Explanations of Export Persistence." *European Economic Review* 79: 113–128

<u>**FOR ONLINE APPENDIX**</u>

A data tables

year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	total
1992	2,232																		2,232
1993	823	1,235																	2,058
1994	583	330	1,160																2,073
1995	440	213	339	953															1,945
1996	372	163	178	255	899														1,867
1997	321	128	133	170	248	877													1,877
1998	268	104	124	132	153	256	893												1,930
1999	232	85	87	114	117	187	262	1,026											2,110
2000	203	85	79	91	103	136	170	344	1,372										2,583
2001	187	70	65	79	85	109	145	229	389	1,251									2,609
2002	173	64	62	72	68	88	112	171	242	399	1,373								2,824
2003	165	51	58	62	62	77	86	140	185	301	440	1,719							3,346
2004	150	52	41	53	63	76	80	132	164	223	327	616	1,768						3,745
2005	140	52	47	39	54	77	69	115	145	196	235	398	661	1,902					4,130
2006	122	46	44	39	44	71	65	110	131	157	168	308	410	564	1,896				4,175
2007	113	37	39	31	42	55	48	91	101	132	156	240	305	365	548	1,681			3,984
2008	93	29	30	24	38	50	45	74	90	117	130	184	198	230	331	447	1,455		3,565
2009	80	25	28	24	28	40	39	60	72	88	97	145	175	157	230	248	386	1,378	3,300

Table 14: Number of Exporting Firms, by Entry Cohort

year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	total
1992	469																		469
1993	352	83																	435
1994	336	83	92																510
1995	313	75	102	58															549
1996	256	67	62	40	60														484
1997	247	84	43	41	48	119													581
1998	225	49	42	36	45	131	63												590
1999	207	51	49	41	39	197	74	81											739
2000	180	53	55	37	51	102	53	158	109										799
2001	150	22	51	41	28	57	36	80	101	111									677
2002	124	23	47	34	27	28	23	45	65	83	40								538
2003	147	42	51	31	42	24	22	37	71	107	50	78							702
2004	156	43	53	19	57	21	23	42	78	106	60	107	90						855
2005	150	22	75	17	52	18	23	43	78	80	58	81	75	84					855
2006	117	31	52	14	64	43	17	38	61	79	32	51	52	112	78				838
2007	103	7	18	11	67	58	19	30	28	64	22	35	33	66	67	62			689
2008	95	6	9	8	33	37	17	33	26	34	20	31	37	54	42	53	57		591
2009	68	22	7	6	13	24	10	23	16	16	14	22	41	25	39	37	36	64	485

Table 15: Value of Exports, by Entry Cohort (millions of $US)

year	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	pooled
1992	210																		210
1993	428	67																	211
1994	576	251	79																246
1995	712	353	300	61															282
1996	687	411	346	158	67														259
1997	771	652	321	241	192	136													310
1998	839	468	339	269	297	510	71												306
1999	893	601	561	361	336	1,054	281	79											350
2000	885	623	697	407	496	750	313	460	80										309
2001	801	316	783	519	329	521	251	350	259	89									260
2002	716	353	757	473	399	318	207	260	268	207	29								191
2003	891	827	870	493	677	315	257	260	385	355	114	46							210
2004	1,039	828	1,281	358	900	281	291	318	478	476	183	174	51						228
2005	1,071	413	1,593	444	967	231	326	375	535	408	248	204	113	44					207
2006	958	675	1,177	356	1,448	605	256	341	464	505	188	165	126	198	41				201
2007	915	175	466	357	1,606	1,048	391	327	278	481	140	145	108	181	123	37			173
2008	1,023	208	283	341	860	747	379	443	289	287	153	166	186	236	125	120	39		166
2009	855	864	262	266	478	607	255	389	221	176	143	152	235	162	169	151	93	47	147

Table 16: Exports per Firm, by Entry Cohort (thousands of $US)

Year	Colombian Sellers	U.S. Importers	Pairs
1992	2,232	1,190	3,087
1993	2,058	1,183	2,824
1994	2,073	1,212	2,810
1995	1,945	1,173	2,588
1996	1,867	1,191	2,490
1997	1,877	1,208	2,480
1998	1,930	1,191	2,495
1999	2,110	1,386	2,793
2000	2,583	1,661	3,411
2001	2,609	1,698	3,483
2002	2,824	1,826	3,733
2003	3,346	2,110	4,483
2004	3,745	2,296	5,071
2005	4,130	2,457	5,552
2006	4,175	2,471	5,607
2007	3,984	2,343	5,307
2008	3,565	2,221	4,751
2009	3,300	2,079	4,467

Table 17: Exporters and importers by year

B Data checks

To investigate the quality of the exporter id (manuf_id) in the U.S. import records, we ran a series of robustness checks. The Colombian and U.S. data overlap for the years 2000-2008 and both contain measures of the value of exports as well as the number of exporting firms. If the manuf_id variable is error-prone and noisy, we would expect the U.S. data to over-report the number of Colombian firms exporting to the U.S. That is, each time a customs broker wrongly enters the data in the field, a new firm would be created. Table 18 below summarizes the total value of exports to the U.S. and the number of Colombian firms, by year, for each data set.

The datasets align much more closely on value than they do on firm counts. The difference in value is never more than 10% (and declining) while the firm count difference ranges from 28% to 74%.

To look more closely at the cause of the differences in firm counts, we compared the number of firms across sources by HS2 categories. The counts in the LFTTD were higher than the Colombian data in only 28 of the 82 codes and by far the biggest differences are in HS codes 61 and 62: textiles. In these two product classes the U.S. data identify 4025 more firms than the Colombian data. If we remove these two sectors from the list, the difference in firm counts flips and the Colombian data contain 1001 more firms than the LFTTD.

Interestingly, Title 19 of U.S. code requires that the manuf_id variable for textile products represent the manufacturer of the textile products, not an intermediary. That is, for this sector only, the CBP 7501 form must report the manufacturer, not an intermediary. By contrast, prior work by several authors of this paper has shown that the Colombian data report the exporter, which may or may not be the manufacturer. Given that previous research has shown that developing countries tend to have a disproportionately large share of small manufacturing firms (Poschke, 2018), it may be that a large part of the reason the U.S. data report so many more firms in the textile sector is that the U.S. data count many small manufacturers while the Colombian data may more often report textile aggregators and intermediaries.

	Colombia		**United States**		**% difference**	
Year	# exporters	value	# exporters	value	# exporters	value
2000	1775	1038	2721	1140	53%	10%
2001	2026	995	2744	1019	35%	2%
2002	2230	870	2986	855	34%	-2%
2003	2800	1113	3579	1119	28%	1%
2004	3035	1379	4002	1415	32%	3%
2005	2861	1554	4288	1438	50%	-7%
2006	2689	1665	4361	1552	62%	-7%
2007	2420	1540	4175	1496	73%	-3%
2008	2161	1570	3758	1474	74%	-6%

Table 18: Colombian versus U.S. Customs Records

As a final check of the integrity of the manuf_id variable, and the robustness of our main results, we experimented with a "fuzzy" version of the manuf_id variable that did not contain any street numbers in the string (a likely source of input errors). The effect is to reduce the number of Colombian firms in the data, an approximation of fixing a portion of the noise from data entry errors. Next we re-ran Table 4 with the fuzzy data and compared the results to the original version.

Using the fuzzy version did not reduce the separation rates substantially (about 6%) and left the patterns intact. It does not appear that our results are sensitive to a modest reduction in data entry errors.

C Model Identification and Fit

Identification Table 19 reports the response of the loss function, $Q(\hat{\Lambda})$, to 5 percent perturbations of each parameter. Shocks to most of the parameters have a strong effect on this fit metric, but the effects of shocks to the jump size (Δ_y) and the home market search cost (κ_0^h) are modest, and the impact of fixed cost (F) shocks is an order of magnitude smaller still. Nonetheless, $Q(\hat{\Lambda})$ is sensitive enough to these parameters that estimated values of each is statistically significant different from 0 (Table 9).

The identification of these parameters (and others) can be better understood with reference to Table 20, which reports the sensitivity matrix proposed by Andrews et al. (2017). This matrix shows the responsiveness of the estimated parameter vector to perturbations of the moments.

Note that all of the moments have a large effect on the search cost scalars, $\hat{\kappa}_0^h$ and $\hat{\kappa}_0^f$, though in percentage terms their effects are much more modest.[51] This sensitivity reflects the importance of these parameters as determinants of firms' search efforts, which in turn influence almost all of the moments we study. Similarly, as a key scaling parameter, $\hat{\Pi}^f$ responds to most moments. Note also that our estimates of the match exit regression (Table 7, column *iv*) and the match sales autoregression (Table 8, column *iv*) strongly influence all of our parameter estimates. This is because the simulated versions of both regressions are sensitive to the set of active exporters generated by our model, which in turn depends upon all of the structural parameters. The influence of the remaining moments is more focused on subsets of structural parameters. For a detailed discussion, refer to Section 5.2.2 of the text.

Model fit Table 21 compares the data-based moments reported in Tables 7 and 8 with their simulated counterparts from the benchmark model and the known-θ^f model. Discussion of these results can be found in Section 5.3 of the text.

[51] κ_0^h is order 10^2 and κ_0^f is order 10^6.

Table 19: Loss Function Sensitivity to Parameters

	$\Delta\hat{Q}$
Π	0.5538
F	0.0007
α	0.2633
β	0.2079
Δ^y	0.0088
λ_b	0.2978
σ_φ	0.7040
γ	0.0384
κ_0^h	0.0023
κ_0^f	0.3346

Notes: Figures are loss function responses to parameter perturbations. For each parameter, we average the responses to a 5 percent upward perturbation and a 5 percent downward perturbation, starting from the benchmark estimates.

Table 20: Sensitivity matrix

	$\hat{F}$	$\hat{\Pi}^f$	$\hat{\alpha}$	$\hat{\beta}$	Δ^y	$\hat{\lambda}_b$	$\hat{\gamma}$	$\hat{\kappa}_0^h$	$\hat{\sigma}_\varphi$	$\hat{\kappa}_0^f$
Table 7, equation *iv*										
mean $D^{exit\ match}$	0.458	-9.146	-0.031	0.251	-1.413	-0.631	0.445	5.905E+01	-0.403	8.854E+05
$D^{new\ to\ mkt}$	-0.736	-5.059	0.009	-0.099	2.972	0.124	0.057	-5.548E+03	0.321	-3.574E+06
$\ln X^f$	-0.013	-0.739	-0.006	0.041	0.090	-0.008	0.005	-2.792E+02	0.094	3.195E+04
$\ln A^{match}$	-0.573	-6.584	0.030	-0.215	2.327	0.302	0.130	-4.608E+03	0.197	-5.819E+06
$\ln A^{firm}$	0.182	6.344	-0.013	0.061	-0.868	-0.245	-0.127	2.267E+03	-0.280	4.240E+06
Table 8, equation *i*										
mean $\ln X^f$	-0.022	-0.481	0.000	-0.012	0.109	0.028	-0.001	-2.267E+02	0.100	-3.171E+05
$\ln X^f_{t-1}$	0.071	1.527	0.008	-0.026	-0.338	-0.036	-0.006	7.477E+02	-0.228	2.726E+05
$D^{new\ match}_{t-1}$	-0.329	5.890	0.015	-0.291	1.000	0.105	-0.016	-9.469E+02	-0.726	2.022E+06
$\ln A^{match}$	0.386	-9.710	-0.034	0.380	-1.019	0.644	0.084	2.208E+02	1.197	-4.028E+06
$\sqrt{mse}$	-0.327	4.454	0.016	0.107	1.172	-0.212	-0.075	-1.704E+03	-0.400	1.129E+06
Table 7, equation *v*										
mean $\ln(m)$	0.015	-0.140	-0.001	0.008	-0.051	0.006	0.001	6.561E+01	0.013	-4.896E+04
Table 8, equation *ii*										
$\ln X^h_t$	-0.043	0.032	0.000	0.000	-0.021	0.000	0.000	4.333E+02	-0.003	7.821E+03
Table 8, equation *iii*										
$\ln X^h_{t-1}$	-0.013	-0.107	0.000	0.002	0.055	-0.002	0.001	-1.102E+02	0.008	-2.602E+04
Table 7, equation *iii*										
mean $\ln(s)$	-0.171	0.950	0.006	-0.060	0.626	0.026	-0.003	-9.973E+02	0.066	1.136E+05
$\ln((1+n)$	-0.036	-4.275	-0.002	-0.011	0.209	0.097	-0.006	-5.030E+02	0.260	-7.822E+05
Table 7, equation *i*										
mean$(\frac{a}{n})$	0.003	-0.808	-0.004	0.032	0.025	-0.003	0.010	-1.759E+02	0.070	-9.135E+04
$\ln(1+n)$	-0.050	-1.985	-0.008	0.063	0.282	0.016	0.020	-7.807E+02	0.201	-4.398E+05
Table 7, equation *ii*										
mean (u^2)	0.000	0.007	0.000	0.000	-0.001	0.000	0.000	2.491E+00	-0.001	2.163E+03
$\ln(1+n)$	0.000	0.011	0.000	0.000	-0.001	0.000	0.000	4.221E+00	-0.001	1.798E+03
Table 8, equation *v*										
mean $\frac{X_f}{X_f+X_d}$	0.406	2.376	0.007	-0.029	-1.664	0.029	-0.020	3.175E+03	-0.250	5.847E+05
Table 8, equation *iv*										
mean D^f	-0.004	0.260	0.002	-0.016	-0.001	0.000	-0.002	4.635E+01	-0.034	2.909E+03

Notes: This table reports the "sensitivity matrix" developed by Andrews et al. (2017). It is calculated as $-(G'WG)^{-1}G'W$ where W is the weighting matrix that appears in the loss function and $G = \frac{\partial[m - m(\hat{\theta})]}{\partial\hat{\theta}'}$ is the Jacobian of the moment vector with respect to the vector of estimated parameters. The partial derivatives it contains are calculated as the average of the moment vector's responses to a 5 percent upward perturbation and a 5 percent downward perturbation, parameter by parameter.

Table 21: Data moments versus simulated moments

	dependent variable mean or explanatory variable	data	benchmark model	known-θ^f model
Table 7, equation *i*	mean $(\frac{a}{n})$	0.413	0.580	0.722
	$\ln(1+n)$	0.093	0.057	0.036
Table 7, equation *ii*	mean (u^2)	0.091	0.091	0.052
	$\ln(1+n)$	-0.056	-0.026	-0.028
Table 7, equation *iii*	mean $\ln(s)$	-3.051	-1.515	-1.517
	$\ln((1+n)$	0.837	0.155	0.378
Table 7, equation *iv*	mean $D^{exit\ match}$	0.395	0.215	0.215
	$D^{new\ to\ mkt}$	0.034	-0.176	-0.246
	$\ln X^{match,f}$	-0.032	-0.070	-0.080
	$\ln A^{match}$	-0.054	-0.121	-0.123
	$\ln A^{firm}$	-0.028	0.049	0.021
Table 7, equation *v*	mean $\ln(m)$	1.176	1.136	1.165
Table 8, equation *i*	mean $\ln X^{match,f}$	10.665	10.778	13.995
	$\ln X_{t\ 1}^{match,f}$	0.826	0.851	0.927
	$D_{t\ 1}^{new\ match}$	0.328	0.621	0.663
	$\ln A^{firm}$	0.063	0.078	0.095
	$\sqrt{MSE}$	1.208	0.877	0.924
Table 8, equation *ii*	$\ln X_t^{firm,h}$	0.979	0.989	0.995
Table 8, equation *iii*	$\ln X_{t\ 1}^{firm,h}$	0.344	0.522	1.047
Table 8, equation *iv*	mean D^f	0.095	0.082	0.038
Table 8, equation *v*	mean $\frac{X^{firm,f}}{X^{firm,f}+X^{firm,d}}$	0.162	0.165	0.099

Table 7 variables: a is cumulative number of successes, n is cumulative number of matches, u is the residual from equation *i* regression, s is the inverse of the time interval between commencement of the current match and the next meeting, $D^{exit\ match} = 1$ if match dies in current year, $D^{new\ to\ mkt} = 1$ if exporter market tenure is less than 1 year, A^{match} is age of match. A^{firm} is foreign market tenure of exporting firm, $X^{match,f}$ is foreign sales volume generated by match, and m is match-specific number of shipments per year.

Table 8 variables: $X^{match,f}$ is match-specific sales in U.S. market, $D_t^{new\ match} = 1$ if match is in its first year, A^{firm} is firm export market tenure, MSE is the mean square from equation i, $X^{firm,h}$ is firm sales in the home market, and $D^f = 1$ if the firm is an exporter to the U.S.