Rationally inattentive exporters

Juan Sebastián Fernández-Ibáñez

University of Michigan *

May 20, 2020

[click here for the most recent version]

Abstract

What role do information frictions play in trade costs? This paper develops a model of rationally inattentive exporters to answer the question. In the model, firms do not know their export revenue, but instead have a prior. The prior can be complemented with a costly signal of chosen precision. The theory reveals (1) information costs affect medium-sized firms the most, (2) the effect on trade volume is ambiguous, (3) information frictions undermine productivity as the criterion to sort firms into exporting. I take the model to the data using the Colombian Manufacturing Survey of 2017. Being fully informed costs the firm as much as the fixed cost of exporting. The estimates of fixed costs are in the range observed in the literature.

*I would like to thank Andrei Levchenko and Sebastian Sotelo for their guidance. Zach Brown, John Leahy, Ying Fan, Dominick Bartelme, Jagadeesh Sivadasan, Brian Cevallos Fujiy, Nadim Elayan Balagué, José Ramón Morán Van Gelderen, Barthélemy Bonadio, Luis Baldomero Quintana, Luis Espinoza Bardales, Leticia Juárez, Agostina Brinatti, Ting Lan, and members of the third year seminar for helpful discussions. All errors are mine.

1 Introduction

The field of International Trade has made significant progress in reconciling the theory to the data in the last decades. Current models make use of both iceberg costs [Samuelson, 1954] and fixed costs [Melitz, 2003] to rationalize the variation of trade. But, when estimated under standard values of trade elasticity [Simonovska and Waugh, 2014], the costs are puzzlingly large. There is something the theory is missing.

This paper explores the role of information as a trade cost. I layer rational information acquisition [Matějka and McKay, 2015] on top of a canonical model of trade with heterogeneous firms [Melitz, 2003]. Using the Colombian Manufacturing Survey, I take the model to the data. I find that information frictions are in the same order of magnitude as fixed costs.

In canonical models, fixed costs prevent firms below a certain export profitability from exporting. Constant marginal cost and elasticity of substitution make export variable profits proportional to domestic sales. We can then think of domestic sales as the criteria used to choose whether to export or not. If a firm has sales one cent above the threshold for exporting, it exports with probability one. If a firm has sales one cent below the threshold for exporting, it exports with probability zero.

Uncertainty is incorporated by assuming firms know their domestic sales, but do not know the scaling factor between their domestic sales and export variable profits. They have a prior regarding the value of the scaling factor, and can acquire a signal to complement the prior. Signals are costly, and more accurate signals are more expensive. The firm faces a trade-off between reducing the probability of entering an unprofitable market and incurring in information costs. Such trade-off might make firms decide to acquire a noisy signal, a random variable. Since firms use the signal to decide whether to export or not, selection into exporting becomes stochastic.

The model implies information frictions affect firms of different sizes differently. Iceberg costs affect firms proportionally to their sales. Fixed costs affect firms too small to pay them. Information costs affect medium-sized firms the most.

Medium-sized firms are the most affected by information costs because they are the least certain regarding their profitability. Variable export profit is proportional to domestic sales. A firm with little domestic sales will have little export profits and knows it will not be able to pay the fixed cost. A firm with large domestic sales will have large export profits and knows it will be able to pay the fixed cost. A firm with medium domestic sales does not know and needs to acquire information.

Information costs affect productive efficiency. In canonical trade models selection follows an efficiency criterion: large firms are more productive, are able to pay the fixed costs, and export. With information costs, firms select into exporting randomly. A firm might receive a good signal and export, while a larger, more productive firm might get a bad signal and not export. Lowering information costs makes firms buy more accurate signals, making exporting more correlated with productivity, increasing efficiency.

The aggregate effect on trade volumes is theoretically indeterminate. Trade with information frictions might be larger than without, if firms are overall optimistic regarding their suitability to export. They will think they should export when they should actually not. The opposite is true for pessimistic firms.

To take the model to the data I use the Colombian Manufacturing Survey of 2017. The data corroborates on average the story of selection based on domestic sales, both in the aggregate and by sectors. Firms vary greatly in size, which is in line with the literature [Okuyama et al., 1999], [Axtell, 2001], [Luttmer, 2007]. Because of a statistical power restriction and to prevent arbitrariness in the sector codification, I restrict the analysis to four sectors.

I estimate the model for the four sectors. The model interprets firms selecting into exporting noisily as evidence of information costs. Given the high degree of noisiness of the data, the cost of information is estimated to be high; being fully informed costs the firm as much as the fixed cost of exporting. The estimate of the fixed cost of exporting is in line with results from a paper looking at Colombian exporters for a previous time period [Das et al., 2007].

The model performs best in settings where the prior of the firms differs greatly from reality. If the firms are a priori wrong, their mistakes can be explained both by the incorrect priors and noisiness of the signals. Conversely, the model has problems when priors are correct. If firms are a prior right, noisy signals need to do all the work rationalizing the mistakes.

The model tends to overpredict entry for large firms. Since export profits are proportional to domestic sales, a firm can be large enough to afford the fixed cost both under its prior and under the true state. The model assigns very high probability of exporting to such a firm. The data says otherwise, several firms in the right tail of the distribution have zero exports. This issue could be fixed by choosing different parametrical assumptions of the prior.

This paper contributes to a growing literature on trade frictions. Iceberg trade costs date from [Samuelson, 1954], and are now present in all modern canonical models [Anderson, 1979], [Eaton and Kortum, 2002], [Melitz, 2003]. [Melitz, 2003] added fixed cost for exporting to rationalize the large asymmetry in size of exporters and non exporters emerging from the growing empirical literature [Bernard and Jensen, 1999]. As expressed above, traditional frictions are insufficient to rationalize the data and yield unrealistically large values when estimated. The anomaly persists even in empirical-centric extension of the models as in [Eaton et al., 2011].

As a result, there are many recent papers adding frictions to the standard models. [Arkolakis and Muendler, 2010] adds fixed-cost heterogeneity motivating it from the need of firms to advertise in each destination. Building on the recent network theory, [Chaney, 2014] develops a network-based model of international trade, where the frictions arise from firms searching in their network. Given the mixed results, [Eaton et al., 2019] are working on extending [Chaney, 2014] much like how they did with [Melitz, 2003] in [Eaton et al., 2011], to take it to the data in its best possible version. This paper could be understood as using a similar motivation, albeit a substantially different theoretical apparatus, to explore the same problem.

This paper contributes to the growing literature of rational inattention. [Matějka and McKay, 2015] solves the problem of the rationally inattentive agent. [Brown and Jeon, 2019] provides closed-form solutions, making the model much more amenable for estimation, and test the presence of information frictions on the health insurance market. [Porcher, 2020] employs a dynamic rational-inattention setting to explore the effects of information frictions in migration in Brazil. This paper will apply the theoretical framework in [Matějka and McKay, 2015] to an empirical setting in trade.

Finally, this paper also contributes to the literature of information frictions in trade. [Allen, 2014] explores the effect of information frictions on rice traders in the Philippines. [Jensen, 2007] the effect of cellphones in South Indian fisheries. [Steinwender, 2018] the telegraph between the UK and the US in the 19th century. Closer to us, [Dickstein and Morales, 2018] tests which signals are in the information set of exporters in Chile. Compared to the first three, this paper has a broader setting: all firms in current economy. Compared to [Dickstein and Morales, 2018] the framework allows us to have a deeper theoretical understanding of the consequence of information frictions.

The rest of the paper is organized as follows. In Section 2, I present a canonical model of trade, add a layer of rational inattention and comment on the implications of information frictions. In Section 3, I explore the data from Colombia, and show how the main features of the model are present in it. In Section 4, I take the model to the data. Finally, Section 5 concludes.

2 Model

This section develops a model of trade with rationally-inattentive and heterogeneous firms. I present first a canonical trade model, and extract the elements which will be relevant to introduce rational inattention. I then explore the behavior of two extreme firms: the perfectly uninformed and the perfectly informed. I introduce the information acquisition problem, derive choice probabilities of exporting and explore the consequences of information costs on the decision of the firm.

I start from a basic [Melitz, 2003] model of firms deciding whether to export in a monopolistic competition setting. Firm i from sector s faces CES demand for its good in every market j:

$$x_{ijs} = \zeta_{ijs}^{\eta-1} p_{ijs}^{-\eta} P_{js}^{\eta-1} Y_{js} \tag{1}$$

Where

- x_{ijs} is the quantity demanded
- ζ_{ijs} a demand shifter
- η the elasticity of substitution
- p_{ijs} the price
- P_{js} the price level
- Y_{js} total expenditure of the sector in which i operates

Firm i produces one unit of output with constant marginal cost c_{is}. There are both iceberg trade costs τ_{ijs} and fixed costs f_{js}. Revenue equals $r_{ijs} = p_{ijs} x_{ijs}$. Profit maximization yields

$$r_{ijs} = \left[\frac{\eta}{\eta-1}\frac{\tau_{ijs} c_{is}}{\zeta_{ijs} P_{js}}\right]^{1-\eta} Y_{js} \tag{2}$$

Call the home market $j = h$. We can write

$$r_{ijs} = \alpha_{ijs} r_{ihs} \tag{3}$$

Where

$$\alpha_{ijs} = \left(\frac{\zeta_{ihs}\tau_{ijs}P_{hs}}{\zeta_{ijs}\tau_{ihs}P_{js}}\right)^{1-\eta} \frac{Y_{js}}{Y_{hs}} \tag{4}$$

Profits can be expressed as the difference between variable profits and fixed costs. Constant marginal cost and constant elasticity of substitution buys us the ability to write variable profit as revenue multiplied by $\frac{1}{\eta}$. As a result, export profits are:

$$\pi_{ijs} = \frac{1}{\eta} r_{ijs} - f_{js} \tag{5}$$

A firm decides to export iff

$$\mathbb{E}(\pi_{ijs}) > 0 \tag{6}$$

In order to define lack of information, we need to define what we mean by fully informed. Since we assume firms enter a market if their expected profit is positive, defining full information is the same as defining the expected profit of a fully-informed firm. A natural definition is ex-post profits. We cannot define it that way, however, because we do not observe ex-post profits for firms that do not export. Therefore we will define the expected profit of the fully-informed firm as

$$\mathbb{E}(\pi_{ijs}) = \frac{1}{\eta} \alpha_{js} r_{ihs} - f_{js} \tag{7}$$

Where

$$\alpha_{js} = \mathbb{E}_s(\alpha_{ijs} | \pi_{ijs} > 0) \tag{8}$$

In words, α_{js} is the sector-average of the ratio between domestic and export sales of exporting firms. This implies that when forming expectations, firms can at most predict the sector component of revenue perfectly (conditional on knowing their marginal cost) but not idiosyncratic shocks.

We now move to the other extreme, and define the fully uninformed firm. A fully uninformed firm does not condition its decision on its true profitability, but decides based on a prior belief regarding their profitability. Defining a fully uninformed firm is the same as defining the prior.

Inspection of (5) shows that this amounts to defining priors regarding the elasticity of substitution η, export revenue r_{ijs}, and the fixed cost f_{js}. We assume firms know η, f_{js} but do not know r_{ijs}. We assume, however, that firms know domestic sales r_{ihs}, then by (3) we are loading all uncertainty on the parameter α_{js}.

The firm assumes all sectors are equally adapt at generating revenue relative to their domestic revenue. That is, the firm ignores the true sector shifter α_{js}, but knows its distribution across sectors for a given export market j $\alpha_j = \mathbb{E}_s(\alpha_{js})$. Assume $\alpha_{js} \sim N(\alpha_j, \sigma_{\alpha_j})$.

Given our distributional assumption on α_{js} and since r_{ihs}, f_{js} are just data to the firm, the prior of the profit of the firm is just

$$\pi_{ijs}^{prior} \sim N\left(\frac{\alpha_j r_{ihs}}{\eta} - f_{js}, \frac{\sigma_{\alpha_j}^2 r_{ihs}^2}{\eta^2}\right) \tag{9}$$

To describe firms in between the extremes, and where on this continuum of information the firm chooses to place itself, we employ the framework of [Matějka and McKay, 2015]. The assumptions we have made so far incorporate all requisites to use the framework.

Our decision maker is a firm. The firm can choose either one of two discrete actions from the set $A = \{I, O\}$, where I represents entry (In) and O not entry (Out). The state of nature is one scalar α_{js}.

The information acquisition problem can be summarized as follows. The firm has imperfect information about the state of nature and so is unsure of the payoff that results from each action. Having more information is beneficial to the firm, since it will be able to choose the "right" action, in our setting, choosing to export iff profit is positive. The firm can acquire information regarding the state, in the form of a signal. More precisely, the firm can choose an information-processing strategy that determines the joint distribution of the signal and the state. Signals are costly and more informative signals are more costly. After observing the signal, the firm can infer the state from the signal choosing an action as a Bayesian expected utility maximizer.

The decision problem has two stages. In the first stage the firm chooses an information strategy to refine its prior about the state. In the second stage it acts on the information and maximizes expected profit.

In the second stage, the decision maker acts on a belief $B \in \Delta(\mathbb{R})$ where $\Delta(\mathbb{R})$ is the set of all probability distributions on $\mathbb{R}$. Based on B, the decision maker chooses the action with the highest expected payoff.

$$\max_{a=I,O} E_B(\pi_{ijsa}) \tag{10}$$

In the first stage the decision maker selects an information strategy. The decision maker has a prior belief $G \in \Delta(\mathbb{R})$ about the state and can receive signals $s_{sj} \in \mathbb{R}$ on the state to update its beliefs. The information strategy is a joint distribution $F(s_{js}, \alpha_{js}) \in \Delta(\mathbb{R}^2)$ of signals and states such that the marginal distribution over states equals the decision maker's prior G, ensuring the decision maker's posterior beliefs are consistent with its prior. Given this restriction, the decision maker is only free to select the conditional distribution $F(s_{js}|\alpha_{js})$ The other conditional distribution $F(\alpha_{js}|s)$ corresponds to the posterior belief after observing the signal s_{js}, and is obtained through Bayes' rule.

The problem of the firm in the first stage amounts to maximizing the ex-ante expected payoff from the second stage less the cost of information $c(F)$ which reflects the cost of generating signals of different degrees of precision. Mathematically:

$$\max_{F\in\Delta(\mathbb{R}^2)} \int_{\alpha_{js}} \int_{s_{js}} \left[\max_{a=I,O} E_{F(.|s_{js})}(\pi_{ijsa})\right] F(ds_{js}|\alpha_{js})G(d\alpha_{js}) - c(F) \tag{11}$$

$$\text{s.t} \int_{s_{js}} F(ds_{js}, \alpha_{js}) = G(\alpha_{js}) \quad \forall \alpha_{js} \in \mathbb{R} \tag{12}$$

We assume the entropy-based cost function used in the rational inattention literature

$$c(F) \equiv \lambda_{js} \left(H(G) - E_{s_{js}}\left[H(F(.|s_{js}))\right]\right) \tag{13}$$

Where the parameter $\lambda_{js} \geq 0$ is the unit cost of information and $H(B)$ denotes the uncertainty of belief B measured by its entropy. Intuitively, entropy measures the dispersion of a distribution. Therefore, if the difference between the entropy of the prior G and the expected entropy of the posterior $F(.|s_{js})$ is positive, the posterior has less dispersion than the prior. The firm is more certain under the posterior then under the prior. (13) says the firm pays in proportion to this reduction in uncertainty.

We jump to the result of the model. [Matějka and McKay, 2015] show that the solution to the problem implies choice probabilities of the form:

$$P_I = \frac{P_I^0 \exp\left(\frac{\frac{1}{\eta}\alpha_{js}r_{ihs} - f_{js}}{\lambda_{js}}\right)}{P_I^0 \exp\left(\frac{\frac{1}{\eta}\alpha_{js}r_{ihs} - f_{js}}{\lambda_{js}}\right) + P_O^0} \tag{14}$$

$$P_O = \frac{P_O^0}{P_I^0 \exp\left(\frac{\frac{1}{\eta}\alpha_{js}r_{ihs} - f_{js}}{\lambda_{js}}\right) + P_O^0} \tag{15}$$

Where

$$P_I^0 = \int_{\alpha_{js}} P_I(\alpha_{js}) dF(\alpha_{js}) \tag{16}$$

$$P_O^0 = \int_{\alpha_{js}} P_O(\alpha_{js}) dF(\alpha_{js}) \tag{17}$$

We can obtain P_a^0 as a solution to the following problem:

$$\max_{P_I^0, P_O^0} \int_{\alpha_{js}} \lambda_{js} \log \left[P_I^0 \exp \left(\frac{\frac{1}{\eta}\alpha_{js} r_{ih} - f_{js}}{\lambda_{js}} \right) + P_O^0 \right] dF(\alpha_{js}) \tag{18}$$

$$\text{s.t } P_I^0 + P_O^0 = 1 \tag{19}$$

$$P_I^0, P_O^0 \geq 0 \tag{20}$$

If there is an interior solution, P_I^0 is implicitly defined by:

$$\int \frac{\lambda_{js} \left[\exp\left(\frac{\frac{1}{\eta}\alpha_{js} r_{ihs} - f_{js}}{\lambda_{js}} \right) - 1 \right]}{P_I^0 \left(\exp\left(\frac{\frac{1}{\eta}\alpha_{js} r_{ihs} - f_{js}}{\lambda_{js}} \right) - 1 \right) + 1} dF(\alpha_{js}) = 0 \tag{21}$$

Where we can write $P_I^0 = h(r_{ihs}, \eta; f_{js}, \lambda_{js})$

Given that (14), (15) are logit choice probabilities we can also write the problem of the firm as choosing between two random profits defined by:

$$\hat{\pi}_{ijsI} = \frac{\frac{1}{\eta}\alpha_{js} r_{ihs} - f_{js}}{\lambda_{js}} + \log P_I^0(\eta, r_{ihs}; \lambda_S, f_{js}) + \varepsilon_{ijsI} \tag{22}$$

$$\hat{\pi}_{ijsO} = \log P_O^0(\eta, r_{ihs}; \lambda, f_{js}) + \varepsilon_{ijsO} \tag{23}$$

Where ε_{ija} are iid errors distributed extreme value type I. Or equivalently,

$$\tilde{\pi}_{ijsI} = \frac{\frac{1}{\eta}\alpha_{js} r_{ihs} - f_{js}}{\lambda_{js}} + \log \frac{P_I^0(\eta, r_{ihs}; \lambda_S, f_{js})}{P_O^0(\eta, r_{ihs}; \lambda, f_{js})} + \nu \tag{24}$$

$$\tilde{\pi}_{ijsO} = 0 \tag{25}$$

Where ν is a logistically distributed random variable.

We can think about the $\tilde{\pi}$ as posterior beliefs. The firm compares $\tilde{\pi}_I$ with $\tilde{\pi_O}$. It chooses to enter iff $\tilde{\pi}_I > \tilde{\pi}_O = 0$.

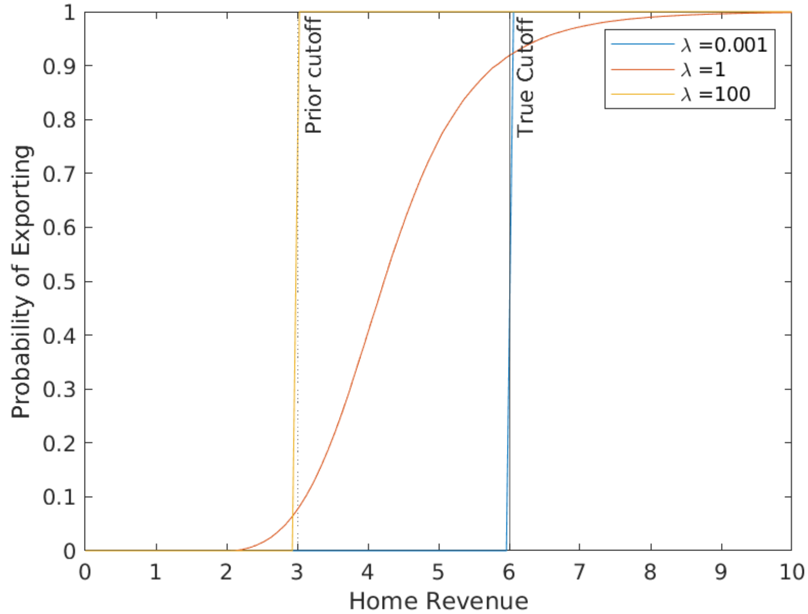
For the case of not exporting, the posterior belief $\tilde{\pi}_O$ is always zero, because the firm neither gets revenue nor has to pay a fixed cost.

For the case of exporting, the posterior $\tilde{\pi}_I$ is based on three terms corresponding to three factors: the information of the signal, the prior, and the noise in the signal.

The first term shows the information of the signal. We have the true profit π_I modulated by λ_{js}. λ_{js} modulates the effect of true profitability because the higher the λ_{js}, the more noisy the signal the firm buys. The higher the noise, the less the decision of the firm will be based on true profitability.

The second term corresponds to priors. Remember P_a^0 is the probability a firm

Figure 1: Probability of exporting, for different values of λ_{js}



chooses action a after deciding on its information strategy but before observing the signal. Therefore priors have a large role to play in P_a^0. Note that P_a^0 work as expected: the higher P_I^0, the higher $\tilde{\pi}_I$; the higher P_O^0, the lower $\tilde{\pi}_I$.

The third term corresponds to the noise in the signal. ν is an independent random variable responsible for making the decision stochastic. Its logistic distribution is a consequence of the assumptions on the information cost function.

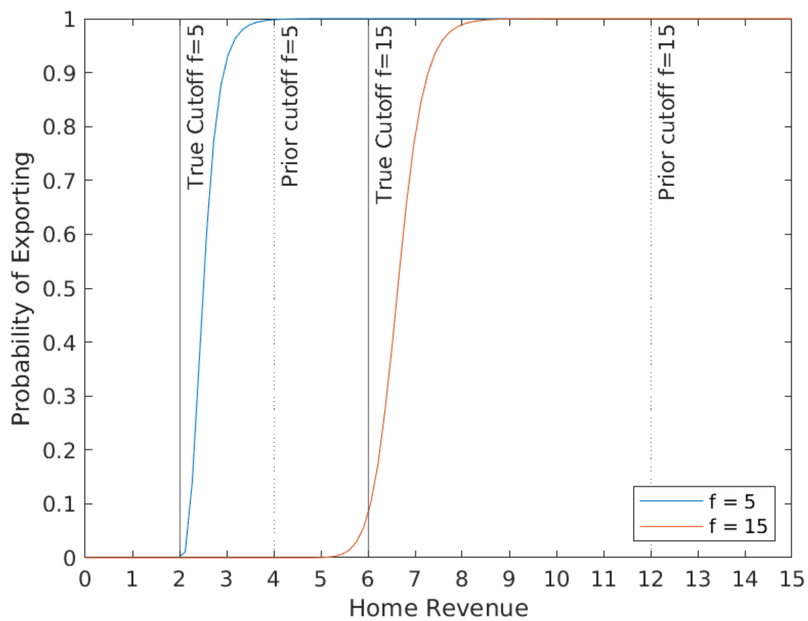
To shed light on the workings of the model, I plot the probability of exporting as a function of r_{ihs} in Figures (1) and (2)

Note in both figures that the probability of exporting is increasing in r_{ihs}. Since in the model exports are multiple of domestic sales, firms with large domestic sales will be more likely to have exports large enough to compensate for the fixed cost of exporting.

Figure (1) shows the effect of the unit cost of information λ_{js}. λ_{js} affects the behavior of the firms by changing the source of information they use. If the cost of information is high, they optimally decide to buy a noisier signal. Given Bayesian updating, in turn they will give a lower weight to the signal, relative to the knowledge of α_{js} they have from their prior.

In the limit, when $\lambda_{js} \to \infty$, firms choose based solely on prior and get a signal uncorrelated with the state. They decide based on a cutoff informed by the prior, namely $\frac{f_{js}\eta}{\bar{\alpha}_j}$. Firms export with probability one if their sales are above such cutoff, and export with probability zero otherwise. Figure (1) exemplifies this cases with

Figure 2: Probability of exporting, for different values of f_{js}



$\lambda_{js} = 100$.

In the other extreme case, when $\lambda_{js} = 0$ firms buy a signal that reveals the true state. They select into exporting iff $\pi_{ijs} > 0$, or equivalenty, if they have domestic sales above $\frac{f_{js}\eta}{\alpha_{js}}$. Figure (1) exemplifies this cases with $\lambda_{js} = 0.001$.

In intermediate cases, the firm buys a signal with some noise, but informative of the state. Its posterior cutoff will be somewhere in between the prior and true cutoffs. Since the signal used to decide is noisy, the firm's decision is also noisy. Selection becomes fuzzy around the posterior cutoff. Figure (1) exemplifies this case with λ_{js} =1.

Figure (2) shows the effect of the fixed costs f_{js}. The role of f_{js} is to rescale the cutoffs. As the fixed cost triples, so do the cutoffs and the choice probability also shifts to the right proportionally.

Figure (1) shows information frictions are not equally relevant for all firms. Changes in λ do not affect the probability of exporting firms of two particular sizes: the small and the large. Equivalently, information frictions are only relevant for medium-sized firms.

Large firms need α_{js} to be several standard deviations below their prior mean for exporting to be unprofitable. Such an event has almost zero probability. They are almost certain that exporting is the right choice, therefore they do not have incentives to acquire any extra information.

Small firms need an α_{js} to be several standard deviations above their prior mean for exporting to be profitable. Such an event has almost zero probability. They are almost certain that not exporting is the right choice, and therefore they do not have incentives to acquire any extra information.

Consider a medium-sized firm. If a firm has domestic sales equal to its prior cutoff, the probability that exporting is profitable is exactly 50%. Its uncertainty is total. It has large incentives to acquire information. Changes in lambda will affect how much information it will acquire, and greatly affect its ultimate decision.

This result changes completely the implications that arise from empirical work such as [Dickstein and Morales, 2018]. If we observe large firms correctly deciding to export to a market, we should not conclude that large firms are well informed of the state. Theory says the opposite is true: they do not need to know the state because in almost every state the correct decision will be to export. Their decision is based not on their knowledge of the export market, but on their own high productivity.

The model also sheds light on the nature of information costs. One might be inclined to ascribe them to one of the traditional categories: iceberg costs or fixed costs, but they are a different category. Information costs neither make exporting less profitable for every firm proportional to their sales, nor prevent firms smaller than certain size from exporting. They make medium-sized firms select into exporting inefficiently. When information costs are present, firms close to the cutoff export not because they are the more productive firms, but because they had a good signal.

If information costs generate inefficiency, lowering information costs increases efficiency, much like a reduction of regular trade costs. But in a striking difference, lowering information costs does not necessarily lead to a higher volume of trade. If information costs are reduced, firms with optimistic priors will, on average, export less because they will realize the baselessness of their optimism. The opposite is true for firms with pessimistic priors. The theory cannot determine which effect is larger on the aggregate.

Table 1: Summary Statistics

Variable	Mean	Sd	Min	Median	Max	Obs
exporter	0.291	0.454	0.000	0.000	1.000	8170
domestic_sales	24.482	127.046	0.001	2.937	5220.827	8170
exports	4.325	31.541	0.000	0.000	1014.830	8170

Note: I present both domestic_sales and exports in billions of Colombian Pesos per year (in 2017, 1 USD was approximately 3,000 Colombian Pesos). exporter is a dummy variable that equals one if the firm has positive exports.

3 Data

This section describes the data and how it correlates to the model. Larger firms are more likely to be exporters, on average. This is true both in the aggregate economy and by sectors. Exporting probability is confirmed to be monotonically increasing in domestic sales. I restrict the analysis to four sectors and verify they are similar to the aggregate economy.

The data source is the Colombian Annual Manufacturing Survey (EAM in Spanish) from the year 2017. The survey samples the universe of Colombian manufacturing firms. There are a total of 8,170 firms split into 120 4-digit ISIC Rev 4 sectors. For each firm I observe annual exports and total sales. I construct the dummy variable exporter to be equal to one if annual exports are positive. I also construct the variable domestic_sales to be equal to the difference between total sales and exports.

The summary statistics of domestic_sales, exports, and exporter are presented in Table (1). There are no missing observations for the three variables. Domestic sales average 25 billion Colombian Pesos, or 80 million USD. Average exports are lower, around 4.5 billion Colombian Pesos, or 1.5 million USD. The relative size of exports compared to sales is consistent with macroeconomic data; Colombia had only 39 billion USD of exports in 2017 vs a GDP of over 300 billion USD. Almost 30% of the firms export.

The mean of both domestic sales and exports is above the median. This implies that there is a fat right tail. In Figure (3) I present the density of the domestic_sales and exports variables, in logs (therefore, excluding zeros). The graph shows both distributions are skewed, because they appear bell-shaped in logs. The skewness in the sample is consistent with the literature for firm size in general ([Okuyama et al., 1999], [Axtell, 2001], [Luttmer, 2007]) and exports in particular ([Helpman et al., 2004], [Arkolakis and Muendler, 2010], [Eaton et al., 2011]) .

The first way we can test the model is to see whether the overall patterns the model describes are true. Do firms with high domestic sales select into exporting? Does the probability of exporting increase monotonically like in Figure (1)? The answer to both

Figure 3: Empirical Distribution of Domestic Sales and Exports

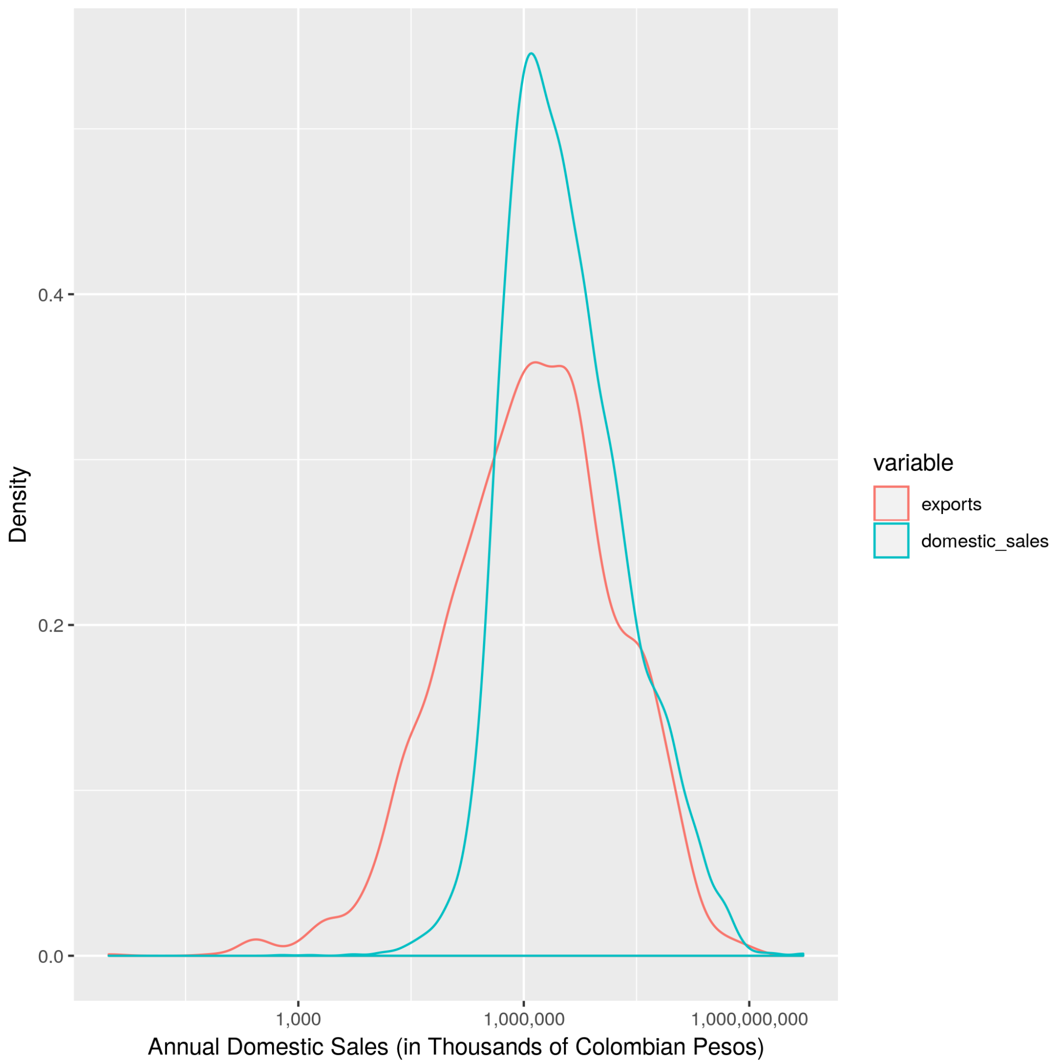


Table 2: Evidence of Selection into Exporting based on Domestic Sales

Dependent Variable: log_domestic_sales	(1)	(2)
Exporter	1.636	1.723
	(0.040)	(0.037)
Sector FE	NO	YES

Note: Standard errors in parentheses.

questions is yes.

To answer the first question, we present regressions results of (log) domestic sales on exporting status in Table (2). The coefficient is positive, statistically different from zero. The result persists almost unchanged when we add sector fixed effects.

To answer the second question turn to Figure (4). Each dot represents a firm, in the (log) domestic sales and exporter space. The non-parametric regression on top of the data has a striking similarity to the curve in the model with imperfect information, Figure (1).

In this paper I decide to present results for four sectors only. The sector numbers are 1521 (Leather/Fur Footwear), 1811 (Printing), 2221 (Plastics) and 3110 (Furniture)[1] There two reasons for this.

The first reason is statistical power. Under simulation, I found I needed upwards of 100 observations to obtain accurate estimates. Otherwise, the log-likelihood becomes extremely flat in the λ dimension, the dimension of interest. Power requirements alone reduce the number of viable sectors dramatically.

The second reason is many of the sectors in the database are catch-all sectors such as "other manufacturing". Such internally heterogeneous sectors go against the spirit of the model of firms competing monopolistically manufacturing similar products.

I present the summary statistics for four chosen sectors in Tables (3)-(6).

In general, both the mean and the standard deviation of domestic sales and exports are lower than in the complete sample. This pattern can either be caused by a rightward shift of the distribution or simply by the skewness of the data. The whole sample includes outliers in the right tail inflating the mean and standard deviation. If the sectors presented here are identical on average to the whole sample but do not include the outliers, the mean and standard deviation will be much smaller.

Inspecting the medians reveals that both factors are at play. Sectors 1521, 1811, 3110 have lower median sales than the whole sample, but less than what the average

[1]The database uses the Colombian version of the UN ISIC Rev. 4. The translation is mine based on the UN classification when possible. The original names in Spanish were 1521: *Fabricación calzado de cuero y piel, cualquier tipo de suela*, 1811 *Actividades de Impresión*, 2221 *Fabriación de formas básicas de plástico*, 3110 *Fabricación de Muebles*.

Figure 4: Empirical Probabilities of Exporting Conditional on Domestic Sales

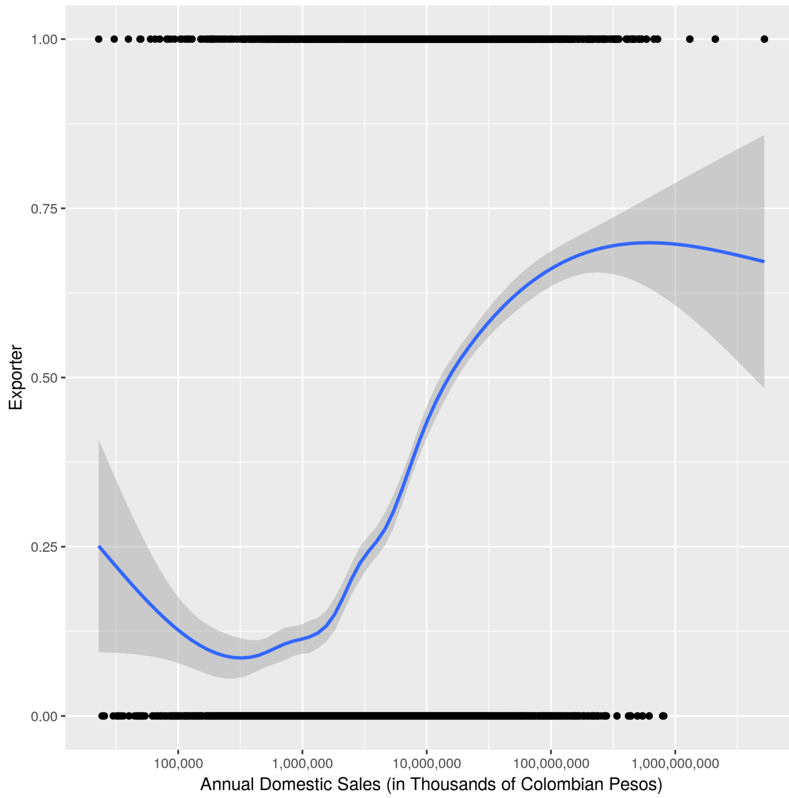


Table 3: Summary Statistics Sector 1521 (Leather/Fur Footwear)

Variable	Mean	Sd	Min	Median	Max	Obs
exporter	0.319	0.468	0.000	0.000	1.000	144
domestic_sales	4.727	12.550	0.033	1.156	80.131	144
exports	0.179	0.590	0.000	0.000	3.806	144

Note: I present both domestic_sales and exports in billions of Colombian Pesos per year (in 2017, 1 USD was approximately 3,000 Colombian Pesos). exporter is a dummy variable equal to one if the firm has positive exports.

Table 4: Summary Statistics Sector 1811 (Printing)

Variable	Mean	Sd	Min	Median	Max	Obs
exporter	0.181	0.385	0.000	0.000	1.000	382
domestic_sales	6.998	15.871	0.001	2.176	129.018	382
exports	0.623	3.486	0.000	0.000	41.514	382

Note: I present both domestic_sales and exports in billions of Colombian Pesos per year (in 2017, 1 USD was approximately 3,000 Colombian Pesos). exporter is a dummy variable equal to one if the firm has positive exports.

Table 5: Summary Statistics Sector 2221 (Plastics)

Variable	Mean	Sd	Min	Median	Max	Obs
exporter	0.377	0.486	0.000	0.000	1.000	130
domestic_sales	22.949	41.077	0.130	5.204	193.234	130
exports	2.777	7.258	0.000	0.000	36.479	130

Note: I present both domestic_sales and exports in billions of Colombian Pesos per year (in 2017, 1 USD was approximately 3,000 Colombian Pesos). exporter is a dummy variable equal to one if the firm has positive exports.

Table 6: Summary Statistics Sector 3110 (Furniture)

Variable	Mean	Sd	Min	Median	Max	Obs
exporter	0.211	0.409	0.000	0.000	1.000	284
domestic_sales	4.638	8.740	0.032	1.426	57.774	284
exports	0.663	3.502	0.000	0.000	32.342	284

Note: I present both domestic_sales and exports in billions of Colombian Pesos per year (in 2017, 1 USD was approximately 3,000 Colombian Pesos). exporter is a dummy variable equal to one if the firm has positive exports.

predicts. Sector 2221 has actually a higher median than the whole sample, even though it has a lower mean.

Regarding the decision to export, sector 2221 has a higher share of exporters than the whole sample (0.38 vs 0.29), sector 1521 is similar (0.31) and Sectors 3110 and 1811 lower (0.21 and 0.18).

Although no sector can fully replicate the aggregate economy, the analysis above shows that the four sectors chosen are not particularly anomalous in any sense. This leads me to the conclusion that they are a reasonable choice for estimation.

4 Empirical Exercise

This section discusses how the model was taken to the data and the limitations, the model fit, and the estimation results. The estimation algorithm is complex and computationally expensive. The model fits best when the prior is very different from reality. Information costs are in the same order of magnitude as fixed costs.

The main restriction I found to be able to fit a canonical trade model was the lack of disaggregation of exports by destination. This forced me to collapse all possible j markets in the model into a single "Rest of the World".

I considered merging the manufacturing survey with customs data, which have destination-level disaggregation. When I inquired about this possibility to the Colombian statistical agency DANE, they noted that although the manufacturing survey is anonymous, the customs data is not. Merging those databases would be a violation of statistical secret.

Collapsing all foreign markets causes problems. For example, the fixed cost of Rest of the World is only equal to the sum of the fixed costs of all countries if all firms either export to all countries (or at least to the same countries) or to no country at all. I called the section Empirical Exercise and not Estimation because of such problems.

Aggregation issues aside, having written the model with estimation in mind makes taking it to the data straightforward. There are two sets of parameters to be estimated. First, one needs to estimate the parameters of the prior. Estimation consists of taking the mean and standard deviation across sectors of the ratio of exports to sales. Second, we need to estimate the cost of information λ_s and the fixed cost of exporting f_s. We do so by Maximum Likelihood Estimation, using (14) (15) to inform the likelihood of each data point.

We cannot separately identify the elasticity of substitution η from the fixed cost of exporting f_s. We take a value standard in the literature of $\eta = 5$ [Simonovska and Waugh, 2014] [Dickstein and Morales, 2018], and discuss the consequences of misspecification.

η only affects the markup of the firm, the relationship between revenue and variable profit. A low elasticity of substitution means the firm needs little revenue to obtain profits, and as such, little revenue to pay the fixed cost. If we assume η is large when it is small, our estimates of f will be too small. Assuming large η is assuming small variable profit, which, in turn, makes the model think that fixed costs are small. Thankfully, the estimates in the literature are within one order of magnitude, such is the upper bound of our error in f_s.

Estimation is computationally challenging. As (14), (15) show, the likelihood of each data point depends on P_I^0, P_O^0, results of a bounded optimization problem with no closed form solution. The algorithm must, for each point, solve (21) implicitly. To

Table 7: Prior Estimation Results

$\bar{\alpha}$	σ_α	Entropy
3.686	28.631	4.773

Note: $\bar{\alpha}$ is the average across sectors of α_s, the average in each sector of the ratio of exports to domestic sales for exporting firms. σ_α is the standard deviation across sectors of α_s.

make matters worse, the implicit solution is wrapped in an integral, which the algorithm solves numerically drawing a $10,000 \times 1$ vector of normal errors. Moreover, given the problem is bounded ($P_a^0 \in [0,1]$), the algorithm must first check if the solution is on the boundaries or in the interior.

The presence of corner solutions implies that, for some parameter values, the probability of the data is zero, or one. Such cases can be seen under close inspection of Figure (1); for low values of domestic sales the probability of a firm exporting is zero, and for high values the probability is one. If the model assigns to a firm that does not export, probability one of exporting, the model will be rejected. The likelihood of the data will be zero, and the log-likelihood will be undefined. Standard optimization algorithms will then crash.

To prevent such an event, the program first does an bidimensional exponential search of parameter values to obtain a valid log-likelihood. Once the program finds a valid first guess it initializes the ordinary algorithm. Such computational measures result in a cost of up to three hours of computation for the estimation of one sector.

I present the results in the estimation of the parameters of the prior in Table (7). On average, sectors seem to be good at exporting, but there is a high variability in their suitability. Outliers on the right tail drive this result. I also add the computation of the entropy of the prior. Such measure helps to interpret λ_s, measured in pesos per entropy.

Figure (5) shows the empirical equivalent of Figure (1). I show the probabilities of exporting implied by the model in red, and the true and prior cutoffs in a solid and dotted vertical lines. I also add the dots representing the raw data and a standard loess non-parametric estimation of the probability of exporting in a solid black curve. The purpose of the non-parametric line is to give a sense of how much the model deviates from the data.

An initial inspection leads to two observations. The line showing the export probabilities predicted by the model is not perfectly smooth. The irregularity arises from the

Figure 5: Estimation Results

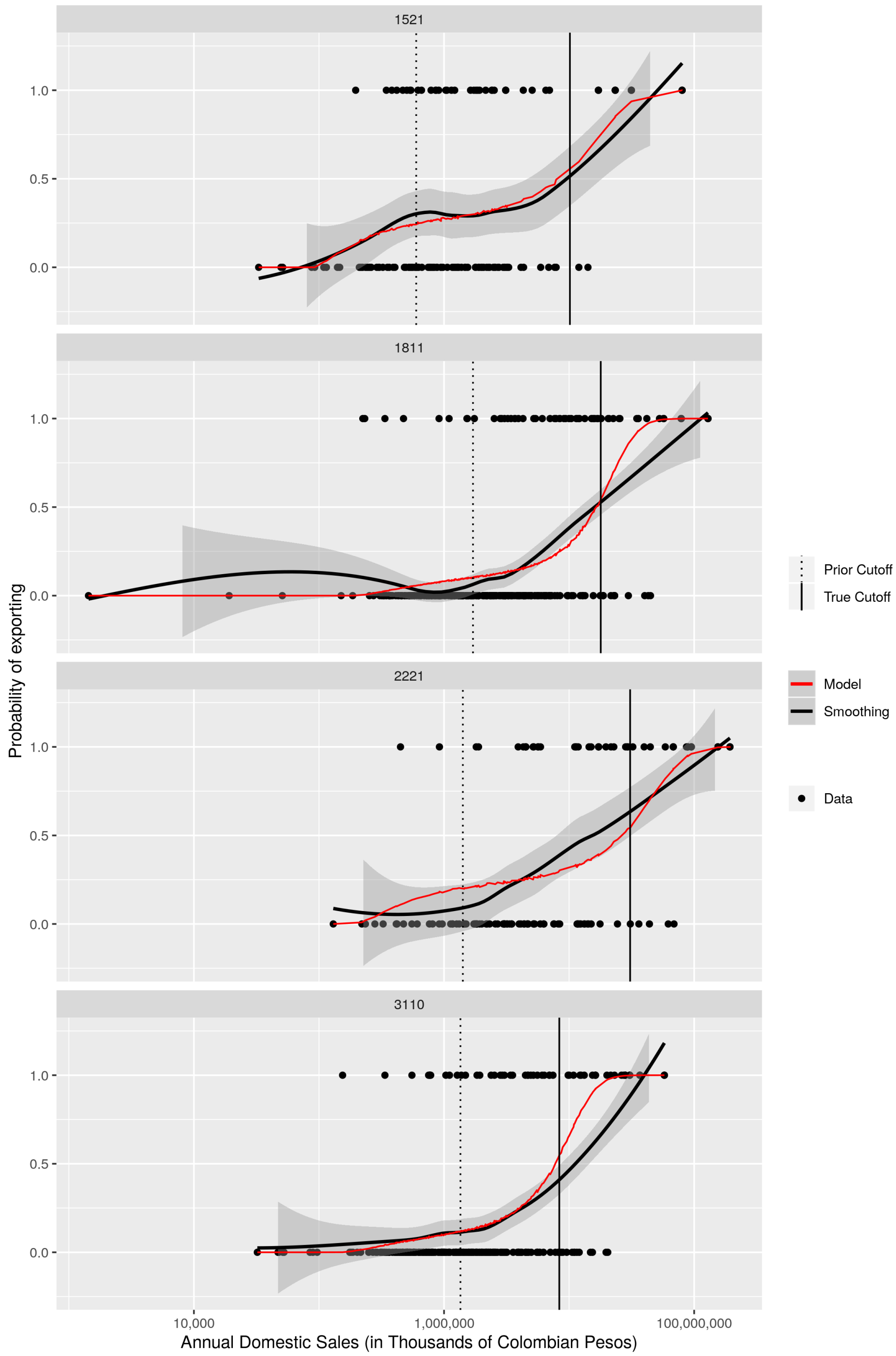


Table 8: Estimation Results

Sector	λ_s	f_s	α_s	True Cutoff	Prior Cutoff	Cost of Full Information
1521	0.372	0.442	0.218	10.149	0.599	1.775
1811	0.523	1.256	0.351	17.911	1.704	2.497
2221	0.724	1.043	0.169	30.761	1.414	3.458
3110	0.414	0.997	0.597	8.348	1.353	1.977

Note: f_s is the fixed cost of exporting and is measured in billions of Colombian Pesos (in 2017, 1 USD was approximately 3,000 Colombian Pesos). λ_s is the unit cost of information and is measured in billions of Colombian pesos per unit of entropy. α_s is the average of the ratio between exports and domestic sales for exporters in each sector. The True Cutoff, Prior Cutoff and the Cost of Full information are measured in billions of Colombian Pesos.

simulation error involved in computing the integral in (21) numerically. Additionally, the shape of these lines appears distorted, it does not correspond to the shape in Figure (1). The distortion arises from the lensing caused by the logarithmic scale in the x axis.

All sectors correspond broadly with monotonic selection into exporting based on domestic sales. There is no value for sales, however, above which firms always export and below which firms never export. On the contrary, selection is noisy. This presages a major role for information costs.

Cutoffs to exporting seem to be high, upwards of 10 billion Colombian pesos or 3 million USD. In all sectors the firms believe the cutoffs are lower than they truly are, α_s in each sector being below the average α_s across sectors.

In sectors 1521 and 2221 the model and the non-parametric estimation seem to be close. Here the gap between the true α_s and the prior is large. Firms are going to be a priori very wrong regarding their suitability of exporting. This allows the model to rationalize the noisiness in the data by attributing it to noisy signals the firms bought.

Sector 1521 has the extra benefit that large firms select cleanly, with only two firms with domestic sales above the true cutoff choosing not to export. This helps the model, because large firms should be unlikely to make a mistake regarding selection. Both their prior and reality push them to the right decision: exporting.

In sectors 1811 and 3110 the model overpredicts exporting for large firms. Here the prior and reality are more similar. Since even uninformed firms are right, it is hard for the model to rationalize the noise, specially for firms significantly above the cutoffs.

The estimation results for each sector are presented in Table (8). The table also includes the cutoffs for exporting, and how much a firm needs to pay to become fully informed. The cost of full information is simply the product of the prior entropy and the cost of information.

We can verify what we inferred from the picture, both the cost of information λ_s

and the fixed cost f_s are high. The fixed cost ranges from 0.442 billion to 1.26 billion Colombian pesos, or 150 to 400 thousand USD. A firm needs to have yearly domestic sales ranging from 8.4 billion to 30.8 billion Colombian pesos to find it profitable to export, or 3 to 10 million USD. The value of the fixed cost is similar to the results of a previous study on Colombian firms [Das et al., 2007].

Information costs seem to be as important as fixed costs, if not more important. A unit of entropy costs from 370 to 720 million Colombian pesos, or 120 to 240 thousand USD. A firm has to pay from 1.7 billion to almost 3.5 billion Colombian Pesos to become fully informed, or 0.5 to 1 million USD. In the model, most firms cannot even afford to pay full information forcing them to base their decision on a noisy signal, which explains the dispersion in the data. Even for firms with large enough profits, they also stay optimally uniformed. As stated in the model section, large firms need no information; their size makes it almost certain that they will find exporting profitable.

5 Conclusions

The field of International Trade has made significant progress reconciling the theory to the data in the last decades. Current models make use of trade costs to rationalize the observed variation of trade. When estimated, trade costs are puzzlingly large. Seeking to further enrich the theory, this paper formally incorporates information to the problem of the firm. I layer rational information acquisition [Matějka and McKay, 2015] to a canonical trade model [Melitz, 2003]. Using the Colombian Manufacturing Survey, I take the model to the data.

The canonical side of the model expects firms to select into exporting based on their domestic sales. The information side of the model interprets violations to this rule as evidence of information costs. Given the frequent violations to the rule, the cost of information is estimated to be high; full information costs the firm as much as the traditional fixed cost of exporting, if not more.

The model performs best in settings where the prior of the firms differs greatly from reality. If firms are a priori wrong, the mistakes can be explained by both incorrect priors and noisy signals. Conversely, the model has problems when priors are correct. If nothing is to be learned, noisy signals need to do all the work to rationalize the mistakes. The model also tends to overpredict entry for large firms, which could potentially be fixed by choosing different parametrical assumptions of the prior.

Given it is a third year paper, there are many extensions possible to improve it. On the computational side, different assumptions could be explored to provide closed form solutions for choice probabilities as in [Brown and Jeon, 2019]. That would aid both computation and interpretation. On the modelling part, different layers of variability could be added similar to [Eaton et al., 2011] to have firm-level heterogeneity in suitability to export and fixed costs as in [Arkolakis and Muendler, 2010]. Finally, on the data side, using detailed census data will both increase statistical power and allow for destination-specific estimation. Such a level of disaggregation can rule out variation in destination-specific fixed costs from being the true source of variation in firm entry. Having a panel would also allow to build a dynamic model and use previous firm performance to inform priors.

References

[Allen, 2014] Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

[Anderson, 1979] Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American economic review*, 69(1):106–116.

[Arkolakis and Muendler, 2010] Arkolakis, C. and Muendler, M.-A. (2010). The extensive margin of exporting products: A firm-level analysis.

[Axtell, 2001] Axtell, R. L. (2001). Zipf distribution of us firm sizes. *science*, 293(5536):1818–1820.

[Bernard and Jensen, 1999] Bernard, A. B. and Jensen, J. B. (1999). Exceptional exporter performance: cause, effect, or both? *Journal of international economics*, 47(1):1–25.

[Brown and Jeon, 2019] Brown, Z. Y. and Jeon, J. (2019). Endogenous information acquisition and insurance choice.

[Chaney, 2014] Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104(11):3600–3634.

[Das et al., 2007] Das, S., Roberts, M. J., and Tybout, J. R. (2007). Market entry costs, producer heterogeneity, and export dynamics. *Econometrica*, 75(3):837–873.

[Dickstein and Morales, 2018] Dickstein, M. J. and Morales, E. (2018). What do exporters know? *The Quarterly Journal of Economics*, 133(4):1753–1801.

[Eaton and Kortum, 2002] Eaton, J. and Kortum, S. (2002). Technology, geography and trade. *Econometrica*.

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

[Eaton et al., 2019] Eaton, J., Kramarz, F., Kortum, S., et al. (2019). Firm-to-firm trade: Exports, imports, and the labor market. (702).

[Helpman et al., 2004] Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus fdi with heterogeneous firms. *American economic review*, 94(1):300–316.

[Jensen, 2007] Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *The quarterly journal of economics*, 122(3):879–924.

[Luttmer, 2007] Luttmer, E. G. (2007). Selection, growth, and the size distribution of firms. *The Quarterly Journal of Economics*, 122(3):1103–1144.

[Matějka and McKay, 2015] Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

[Melitz, 2003] Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

[Okuyama et al., 1999] Okuyama, K., Takayasu, M., and Takayasu, H. (1999). Zipf's law in income distribution of companies. *Physica A: Statistical Mechanics and its Applications*, 269(1):125–131.

[Porcher, 2020] Porcher, C. (2020). Migration with costly information. *Job Market Paper*.

[Samuelson, 1954] Samuelson, P. A. (1954). The transfer problem and transport costs, ii: Analysis of effects of trade impediments. *The Economic Journal*, 64(254):264–289.

[Simonovska and Waugh, 2014] Simonovska, I. and Waugh, M. E. (2014). The elasticity of trade: Estimates and evidence. *Journal of international Economics*, 92(1):34–50.

[Steinwender, 2018] Steinwender, C. (2018). Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–96.