An empirical dynamic model of trade with consumer accumulation*

Paul Piveteau†

August 2018

Abstract

Sunk entry costs have been identified as the main export barrier by standard dynamic models of trade. However, these large entry costs are inconsistent with the existence of many small new exporters with low survival rates in foreign markets. To reconcile these patterns, this paper develops a dynamic structural model of trade in which firms slowly accumulate consumers in foreign markets. Estimating the model using export data from individual firms and a particle Markov chain Monte Carlo estimator, the model correctly predicts lower survival rates for new exporters and estimates much lower entry costs of exporting - half of those estimated in the absence of consumer accumulation. Using simulations and out-of-sample predictions, I show that these results have important implications regarding the aggregate response of international trade to external shocks.

*First version: October 2015. I am grateful to Amit Khandelwal, Eric Verhoogen, Jonathan Vogel and David Weinstein for their advice and guidance. I also would like to thank Costas Arkolakis, Matthieu Bellon, Chris Conlon, Donald Davis, Jean Jacques Forneron, Juan Carlos Hallak, Colin Hottman, Thierry Mayer, Antonio Miscio, Jean-Marc Robin, Bernard Salanie, Gabriel Smagghue, Ilton Soares, Daniel Xu and seminar audiences at Columbia University and Sciences Po Paris for comments and suggestions. I also would like to thank the Alliance program and Columbia University CIBER for financial support, the CNIS and French customs for data access, and the MARCC for computing resources. All remaining errors are mine.

†School of Advanced International Studies, Johns Hopkins University. 1717 Massachusetts Avenue NW. Washington DC, 20036. Email: ppiveteau@jhu.edu

1 Introduction

The decision by individual firms to enter into an export market is responsible for most of the variations in aggregate trade flow across destinations and time. For instance, Bernard et al. (2007) estimate that around 80 percent of the decline of international trade with geographical distance is due to reductions in the number of exporting firms (extensive margin) rather than changes in exports within the firm (intensive margin). Therefore, understanding the determinants of export decisions and the barriers that firms face in foreign markets is critical.

Standard dynamic models of trade that quantify the nature of these trade costs, such as Das, Roberts, and Tybout (2007), highlight the prevalence of large sunk entry costs as barriers to trade. These large entry costs are necessary to explain the persistence in export decisions, the so-called hysteresis of exporters. However, the prevalence of these entry costs is incompatible with important characteristics of new exporter dynamics that have been recently documented in the literature: most new exporters start small and only a small fraction survive and expand in these foreign markets.

This paper introduces inertia in consumers' choices into a dynamic empirical model of trade to reconcile the observed hysteresis in exporting decisions and the dynamic features of new exporters. I introduce this inertia through the existence of a stock of consumers that firms accumulate throughout their experience in foreign markets. To assess the importance of this accumulation of consumers on exporter dynamics, I develop a particle Markov Chain Monte Carlo (pMCMC) estimator that allows me to include other sources of persistent heterogeneity at the firm level such as productivity and product appeal, and estimate the model using export data from individual French firms. The estimated model correctly predicts lower survival rates for new exporters, but also estimates low sunk entry costs of exporting. On average, entry costs are half the value of those estimated in a model without consumer accumulation. These results have important implications regarding the aggregate predictions of the model: aggregate trade responds slowly to shocks and the contribution of the extensive margin is larger in the long run than in the short run. Both of these patterns have been recently documented in the literature; however, they are inconsistent with the standard model.

I start by presenting three stylized facts about exporters that highlight the importance of growth in demand in these exporter dynamics. Consistent with recent studies, sales and survival rates of young exporters are low upon entry, but grow at a fast rate during the first years of exporting. Moreover, this growth is not due to variations in prices during the life of an exporter, but instead, prices tend to also increase on average with export experience. This result suggests that the growth in sales observed in the years following entry into a foreign market is mainly driven by an increase in the demand shifts received by exporters.[1]

Based on these findings, I develop an empirical dynamic model of trade in which consumers only buy from a limited set of firms, which generates inertia in their consumption choice.[2] Therefore, each firm has a different stock of consumers, depending on its history in the foreign

[1]This finding is consistent with recent papers that show the importance of demand characteristics as source firm heterogeneity (Hottman et al., 2016; Roberts et al., 2012).

[2]This extends to a dynamic setting the consumer margin first introduced in international trade by Arkolakis (2010). This inertia could be alternatively modeled with habits formation or other sources of state-dependence in demand.

market, which shapes its profit, expectations, and decisions in each market. This addition to the model has two important consequences on the dynamics of exporters: first, it implies that new exporters start with low levels of sales and profits when entering a new destination. As they survive and accumulate consumers their sales and profits increase, inducing increasing survival rates with their experience in a destination. Second, because current sales are a source of customer acquisition, firms have incentives to reduce their price to foster the accumulation of new consumers.[3]

In order to study the importance of this mechanism on exporter dynamics, I structurally estimate this model using customs data from France. I perform this estimation on the wine industry, which has the double advantage of being an important exporting industry in France, while also being composed of single-good producers. The dataset provides sales and quantities exported by individual firms on each destination market, which allows me to account for several sources of persistent heterogeneity across firms and destinations. In addition to heterogeneity in demand across destinations, the model identifies three types of heterogeneity at the firm-level: product appeal, defined as a demand shifter that is common across destinations;[4] productivity, acting as a cost shifter; and the firm's consumer base, which is identified from within-firm demand variations across destinations. Because this large number of persistent unobservables complicates the estimation, I take advantage of recent results from the statistical literature to develop a particle Markov Chain Monte Carlo (pMCMC) estimator that accounts for this unobserved heterogeneity through particle filtering. To my knowledge, this is the first paper to apply particle filtering within a MCMC algorithm to account for persistent unobservable heterogeneity at a microeconomic level. This method could be applied to many contexts where researchers have to handle persistent unobservables in non-linear models. Moreover, I combine this estimator with recent method developed in Imai et al. (2009) and Norets (2009) to handle dynamic discrete choice problems within a MCMC estimator. Therefore, I am able to obtain value estimates of the entry and per-period fixed costs of exporting, which are identified by rationalizing the actual entry and exit patterns of exporters on the different export markets.

The estimation results demonstrate the importance of consumer accumulation to replicate exporter dynamics. The introduction of state dependence in demand improves the model's ability to fit the dynamics of young exporters: the model can rationalize lower survival rates for young exporters, as well as the growth of sales and survival rates as exporters become more experienced. Moreover, estimated entry costs of exporting are small relative to existing estimates. The average cost to start exporting to a foreign European destination for a wine exporters is around 35 000 euros, less than the average revenue in these destinations.[5] To confirm this result, I estimate a version of the model without consumer accumulation and obtain an estimate of the average entry cost to European destinations of 60 000 euros, twice the estimates of the full model. The reason for this finding is simple: as the model accounts for the fact that it takes years for firms to grow in foreign markets and become successful, large entry costs become unnecessary to rationalize the small fraction of exporters in the data.

[3]Recent empirical evidence for this type of mechanism on domestic market was found by Foster et al. (2016) who studied the behavior of new firms producing homogeneous goods.

[4]Khandelwal (2010) at the product level or Hottman et al. (2016) at the micro level, also define appeal or quality as the demand shifter after controlling for prices.

[5]Or equivalently, 3 times the median yearly revenue on these destinations.

These results have important implications at the aggregate level. In particular, the model generates trade responses to trade shocks that are consistent with patterns documented in the literature. First, the model predicts a slow increase in trade as a response to a permanent positive trade shock: because of the slow accumulation of consumers, it takes time for existing and new exporters to expand and reach their new optimal stock of consumers. As a consequence of these adjustment frictions, the trade response is larger in the long-run than the short-run. In my simulations, the trade elasticity is around 1.3 after a year, but equals 2.8 in the long-run, consistent with the discrepancies documented in the international trade and international macroeconomics literatures. Second, the model can predict the increasing contribution of the extensive margin throughout this trade expansion, as recently documented by Kehoe and Ruhl (2013) and Alessandria et al. (2013). While they enter small in the foreign markets after the shock, new exporters record larger growth than established exporters in the following years, hence increasing their relative contribution to trade growth throughout these years.

Finally, I employ out-of-sample predictions to further confirm the importance of this consumer accumulation in explaining firms' response to shocks. During the sample period, large variations in exchange rates led to a decrease of the exported values and market shares of French wine in the Brazilian market.[6] Based on these variations in exchange rates that affected the relative price of French wine, I construct variations in aggregate demand for French wine from Brazilian consumers. This aggregate demand, in conjunction with outcomes from the model estimated on other destinations, allows me to generate predictions on entry, sales and prices in the Brazilian market, and compare them to the actual realizations of these variables. The model with consumer accumulation is able to better replicate the initial decrease and later rebound in total trade and in the number of exporters. The decrease in estimated entry costs between the two models, reduces the option value of exporting. Therefore, as economic conditions fluctuate, the model with consumer accumulation (and low entry costs) can predict larger inflows and outflows of exporting firms, and therefore larger variations in total trade.

This paper is closely related to the literature investigating exporter and firm dynamics. Das, Roberts, and Tybout (2007) is the first study to quantify entry and per-period fixed costs of exporting by estimating an entry model of trade. Their estimation emphasizes the importance of entry sunk costs to explain the hysteresis of export decisions.[7] My paper builds on their contribution by capturing this hysteresis through state dependence in demand rather than sunk entry costs, and demonstrating the importance of this extension for a number of micro and macro-level facts. Ruhl and Willis (2017) and many more recent studies have documented and studied the specific dynamics of new exporters. Nguyen (2012), Albornoz et al. (2012), Berman et al. (Forth.) and Timoshenko (2015) emphasize the role of demand uncertainty and experimentations to explain exporter dynamics, while Rauch and Watson (2003) and Aeberhardt et al. (2014) develop models where exporters need to match with foreign customers in order to trade. Foster et al. (2016), Fitzgerald et al. (2016) and Rodrigue and Tan (2015) also introduce consumer accumulation to explain the post-entry growth of firms in domestic and foreign markets

[6] The Brazilian devaluation in 1999 and the depreciation of the Argentinian peso in 2002, that fostered Argentina exports to Brazil, have increased the relative price of French wines.

[7] Lincoln and McCallum (2018) similarly shows the prevalence of entry costs when estimating fixed costs of exporting for US firms.

respectively. In particular, Fitzgerald et al. (2016) demonstrates that most of the growth of new Irish exporters is driven by destination-specific demand factors that are consistent with consumer accumulation. However, they do not study the entry decision in these markets. Similar to my paper, Eaton et al. (2014) also develop an entry model with accumulation of customers: they use an importer-exporter matched dataset to estimate an empirical model in which exporters grow through the search of foreign distributors and as they learn their own ability.[8] However, while they do not allow for other margins of firms' growth in foreign markets, my model features other sources of time-varying heterogeneity at the firm level, such as productivity and product appeal. Therefore, I am able to investigate the importance of this new margin on exporter dynamics, and its consequences on the estimation of trade costs and the predictions of aggregate trade movements.

This article is also related to macroeconomic papers that similarly introduce a consumer margin, or study aggregate trade dynamics. Arkolakis (2010, 2015) develops a static framework in which a consumer margin at the firm level generates convex costs of participation to foreign markets and heterogeneous elasticities of trade in the cross section of firms. I extend this consumer margin to a dynamic setting to empirically investigate its consequences on exporter dynamics. Drozd and Nosal (2012) and Gourio and Rudanko (2014) show how convex adjustment costs of market shares can explain several puzzles in international macroeconomics and adjustments along the business cycle. Moreover, several recent papers have investigated the reasons for the slow response to trade, and the discrepancy between short and long-run trade elasticities.[9] This series of papers develops macroeconomic models to explain this discrepancy between elasticities through the role of entry and exit of firms, the importance of establishment heterogeneity or the existence of export-specific investment (Alessandria and Choi, 2007, 2014; Alessandria et al., 2014). My paper also explains this discrepancy by combining the role of consumer accumulation at the firm-level, and the entry of new exporters. However, whereas I do not develop a calibrated general equilibrium model, I estimate an entry model using micro-data and a full-information estimator to discipline the role of this mechanism and investigate its consequences on aggregate trade dynamics.

Finally, this study heavily builds on the literature related to the estimation of dynamic discrete choice models (DDCM). These models display a high level of nonlinearity and therefore require the development of specific techniques to facilitate their estimation. Rust (1987) and Hotz and Miller (1993) can be cited as seminal papers in the development of these techniques. More specifically, I employ a MCMC estimator recently developed by Imai et al. (2009) and Norets (2009), to solve the full solution of the DDCM.[10] Moreover, I use particle filtering to account for unobservable heterogeneity, following recent results from Andrieu, Doucet, and Holenstein (2010).[11]

In the next section, I present stylized facts about the trajectories of exporters that emphasize

[8] See also Akhmetova and Mitaritonna (2012) and Li (2018) that show the importance of demand uncertainty, and Aw et al. (2011) about the impact of R&D activities on exporters.

[9] See Ruhl (2008) for a review on the discrepancy between trade elasticities in the international macro and international trade literature.

[10] An application of this method in Industrial Organization can be found in Osborne (2011).

[11] See Flury and Shephard (2011) for an application to economic models and Blevins (2015) using classical estimation.

the importance of demand in exporter dynamics. In section 3, I build an empirical model of export entry that is consistent with these facts. I present the estimation method in section 4, and show the results of the estimation on a set of French wine makers in 5. Finally, section 6 inspects the aggregate implications of the estimated results through simulations and out-of-sample predictions, and section 7 concludes.

2 Stylized facts about exporters dynamics

In this section, I present three important facts about exporters' dynamics using French customs data. First, new exporters have low survival rates upon entry, but survival increases quickly with experience. Second, exported values grow with age in foreign markets, even after controlling for survival. Third, prices also increase with exporters' age.

These facts motivate the empirical model I present and estimate in the next sections: first, the high level of attrition across age will require to develop an entry model in export markets to account for endogenous selection. Moreover, the rise in export values, while prices do not fall, indicates that this growth is driven by a positive shift in the demand schedule of the firm: the consumer margin introduced in the model will be able to replicate this increase as exporters start small, and accumulate consumers with experience. Finally, the low mark-up charged by young firms to foster this accumulation will explain the observed increase in prices with age.

2.1 Data

The dataset used in this paper is provided by the French customs services. It records yearly values and quantities exported by French firms from 1995 to 2010.[12] Annual trade flows are disaggregated at the firm, country and eight-digit product category of the combined nomenclature (CN). This dataset is used to present stylized facts about new exporters in this section, and a restricted sample from the wine industry will be used to conduct the structural estimation described in the next sections. I perform a number of procedures to improve the reliability of the data. In particular, I correct for the existence of a partial-year bias, which overestimates the growth rate during the first exporting year,[13] and clean the dataset to improve the reliability of unit values. Appendix A describes more precisely this cleaning procedure.

Table 1 provides some information on the distributions of the number of observations along different dimensions. Similarly to what have been documented in the literature, trade flows from France are sparse across firms and destinations. This is true for firms across destinations or product categories in a given year, since the median exporting firm records three flows per year, usually concentrated within one product category or one destination. But this sparsity also appears across time as shown in the second panel of Table 1: contrary to the idea that exporting is a long-lasting activity, we can see that the median exporting spell lasts one year.[14] This is

[12] This dataset records most of the exporting and importing flows of Metropolitan French firms: there exists thresholds under which a firm does not need to report its exporting activity (In 2001, these thresholds were 1,000 euros for exports to countries outside of the European union, and 100,000 for the total trade within the EU.)

[13] See Berthou and Vicard (2015) and Bernard, Boler, Massari, Reyes, and Taglioni (2017) for papers investigating the extent and consequences of this bias.

[14] An exporting spell is defined as a set of consecutive yearly exporting flows between a firm and a foreign destination, or a 8-digit product category - firm pair and a foreign destination.

true even when exports are aggregated across product categories and exporting flows defined at the firm-destination level.

TABLE 1: Description of the data

Statistics	mean	p5	p25	p50	p75	p95	N
# **observations**							
by firm-year	16.5	1	1	3	9	62	1 510 030
by firm-CN8-year	2.50	1	1	1	2	9	9 921 225
by firm-dest-year	3.34	1	1	1	3	10	7 433 090
Spells duration (years)							
firm-dest-CN8 level	1.92	1	1	1	2	6	12 937 524
firm-dest level	2.63	1	1	1	3	10	2 830 147

Notes: CN8 denotes an eight-digit category after normalization. An exporting spell is defined as a set of consecutive yearly export flows to the same destination.

These statistics provide an overview of the prevalence of short and frequent export flows in the the export data. In order to further investigate this aspect and understand the evolution of the other characteristics of these exporting flows, I specifically look at their trajectories across ages in the next subsections.

2.2 Specifications

To describe the trajectories of exporters upon entry, I look at the variation of their survival rates, sales and prices across different ages in foreign markets. I define the age of a firm-product-destination triplet as the number of years this firm has been successively exporting this product category to a market, a market being defined as a 8-digit product category-country pair. I regress the variables of interest (dummy for survival, logarithm of sales or prices) on a full set of age dummies. The specification is augmented with fixed effects that control for the large heterogeneity that exists across industries, destinations and years. Formally, indexing a firm by f, a destination by d, a product category by p, and a year by t, the econometric specification is the following:

$$Y_{fpdt} = \sum_{\tau=1}^{10} \delta_\tau \mathbb{1}(\text{age}_{fpdt} = \tau) + \mu_{pdt} + \varepsilon_{fdt}, \tag{1}$$

where age_{fpdt} is defined as the number of consecutive years a firm f has been selling the product p to destination d. Y_{fpdt} is the logarithm of export sales, the logarithm of prices (unit values),[15] or a dummy equal to one if the firm is still exporting to the market the following year. μ_{pdt} is a market×year-specific fixed effect such that the variation that identifies the coefficients δ_τ comes from variations across firms of different ages, within a given destination×product category×year pair.

Trade data at the firm-product level are known to have a very large level of attrition. These low levels of survival, especially in the early years of exporting, imply that firms surviving 10

[15] I use the terms unit values and prices interchangeably throughout the paper. As usual with this type of dataset, prices are obtained by dividing export values by export quantities.

years differ substantially from firms who recently started to export. Consequently, the variation captured when comparing old and new firms mostly comes from a selection effect, rather than changes across ages for a given set of firms. In order to partially account for this dynamic selection, I also present results when only looking at firm-product-destination triplets that survive 4, 6 and 8 years in their specific markets. Even though this only partially accounts for selection, since surviving firms are also firms with specific trajectories, it shows that the observed relationships are not only due to dynamic selection, but also hold when considering a constant set of firms.

In addition to comparing firms of different ages, I also run an alternative specification that identifies the effect of age on price using variations across destinations. To do so, I include a firm-product-year fixed effect in specification (1) to control for time-varying characteristics that are specific to the firm, such as quality or productivity. Because firm-product pairs have different ages in foreign markets, this specification identifies a relationship between export experience and prices by comparing prices charged by firms simultaneously on different destinations. In appendix B, I also present results from an additional specification that looks at price changes within-exporting spell. As I explain in the appendix, this specification does not allow to identify trend in prices or sales, but only deviations from a trend.

2.3 Results

Here I present three important facts about exporters, namely the growths of the survival rates, exported values, and prices with export experience in foreign markets. Regarding the growth of sales and survival rates, these facts have been extensively documented and discussed in the literature in international trade and macroeconomics.[16] I show that these facts still hold after controlling for the partial-year bias highlighted by Berthou and Vicard (2015) and Bernard et al. (2017). However, the increase in prices has not been documented using a comprehensive trade dataset,[17] even though Foster, Haltiwanger, and Syverson (2016) documents similar patterns for the domestic prices of homogeneous goods, and Macchiavello (2010) show evidence of similar trajectories for prices of Chilean wine in the UK market.

Fact 1: Survival rates are low for new exporters, and strongly increase with their age
First of all, the probability to survive in a market, i.e. to export in this market the following year, is very low for the average exporter. Figure 1 displays the average survival rate for a firm-product pair on a foreign market, for different age or experience levels. For an exporter in its first year, the probability to export the following year is roughly 35 percent. However, this survival probability rapidly increases once exporters have survived several years: this rate is larger than 50 percent at age 2, and close to 75 percent at age 6. This result confirms the conclusions reached from the summary statistics that most export spells are short lived.

These low, yet increasing, survival rates have theoretical and methodological consequences. On the theoretical side, it will be important to have a model of export entry that can replicate and explain these low survival rates: a model in which entry costs are prevalent will have difficulties explaining why so many firms exit the export market so rapidly. On the methodological side,

[16] See for instance Ruhl and Willis (2017) for a presentation of these facts and puzzles.

[17] Simultaneously to the redaction of this paper, several others have documented prices patterns using trade data: see in particular Rodrigue and Tan (2015) and Fitzgerald et al. (2016).

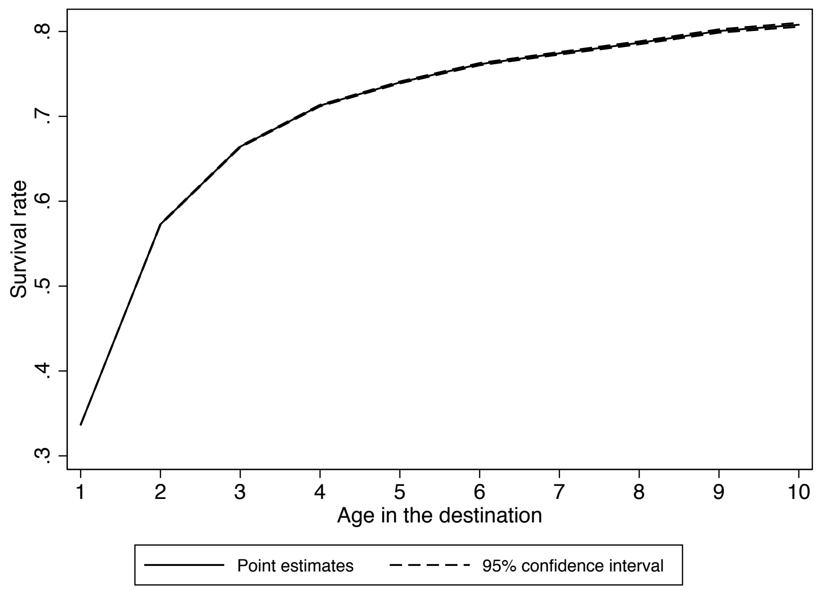

FIGURE 1: Survival rates across export ages

Notes: The figure reports the average survival rate of a firm-product category pair in a destination at different ages. Standard errors clustered at the firm-product-destination level.

these very low survival rates imply it will be necessary to account for this large attrition when interpreting differences across firms in a reduced form exercise, and to model this entry decision in the design of the structural model.

Fact 2: Exported values increase with firm age in a destination, even more so in the first years of exporting Turning to the variation of sales across ages, Figure 2 documents the large growth rates of exported values across ages. This figure is obtained using regression (1) for different sets of exports, normalizing the average log sales at age one to be zero. When comparing exported values for all products (top left panel), exporters which are in their third year of exporting export more than twice as much compared to a new exporter. This difference reaches an order of 7 when comparing an exporter with 8 years of experience to a new exporter. However, these differences are mostly due to strong selection effects: old exporters, which managed to survive in foreign markets, were initially larger than the average new exporter. The three other panels in Figure 2 emphasize this point by looking at the relationship for surviving products. Accounting for survival, the growth rate of sales with export age is strongly reduced. Nevertheless, surviving exporters still record an average growth rate of 25 percent between ages one and two. Moreover, this growth appears to continue the first six years: at this age, exporters tend to be on average twice larger compared to their first year of exporting.

In conclusion, we observe substantial growth rates of sales during the first years of exports. These growth rates are large but appear to be lower than previously described in the literature because of the correction for the partial-year effect highlighted in Berthou and Vicard (2015) and Bernard et al. (2017). Moreover, this positive relationship appears to be robust across product categories and destinations. However, it is important to emphasize that this growth could be generated by the stochastic nature of the exporting process: by focusing on surviving firms, we

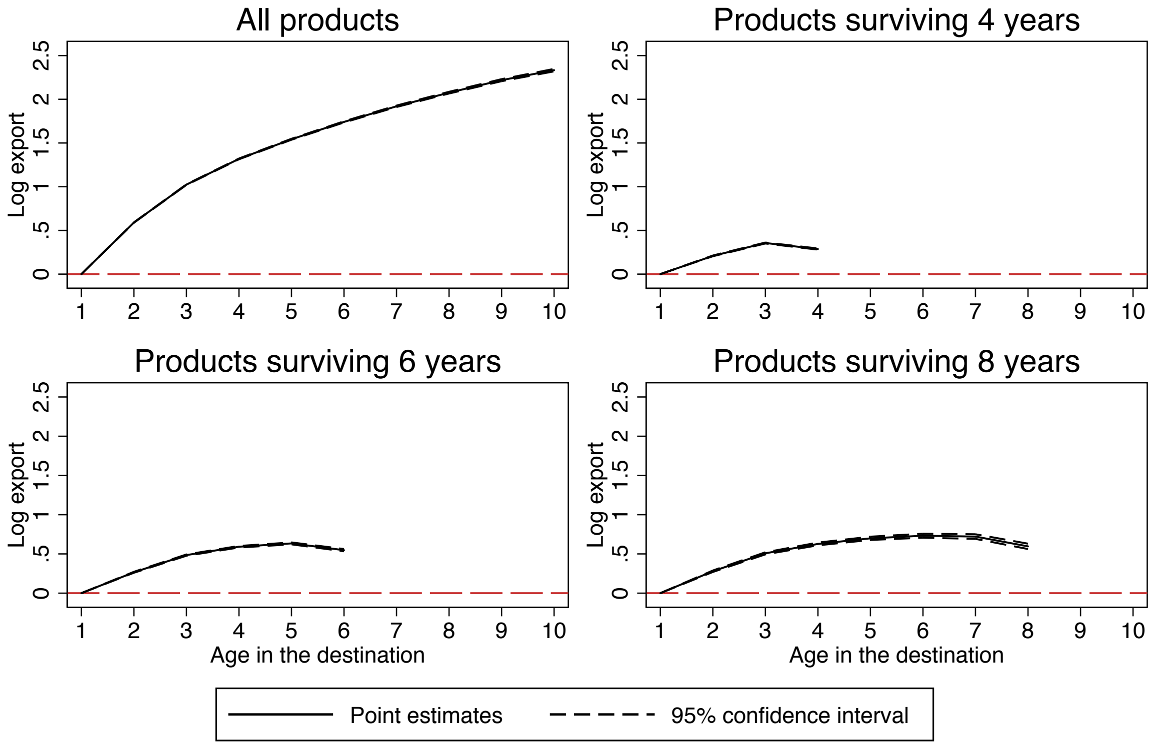

FIGURE 2: Sales across export ages

Notes: The figure reports the cumulative growth of sales of a firm-product category pair in a destination at different ages. The top left panel uses the entire sample, while the top right and bottom panels only uses products that reach ages 4, 6 and 8 respectively. Standard errors clustered at the firm-product-destination level.

are looking at the "winners" of the exporting game, which could explain unusually large growth rates. Accounting for this potential mechanism will be one of the roles of the structural model introduced in the next section.

Fact 3: Export prices increase with firm age in a destination. Given the large growth of export values in the first years of exporting, a natural question is to understand the source of this growth. One possible explanation could be productivity improvements, which would lead to a reduction in the price of exported goods and an increase in their sales. On the contrary, it appears that prices tend to also increase with firms experience in the export market.

Figure 3 reports the estimated parameters of regression (1) in which the average price at age one is normalized to zero.[18] The top left panel shows that the price of an exporter with 10 years of experience is on average 8 percent higher than the price of a new exporter. Similar to sales, this difference could be driven by the selection effects of the exporting activity: for instance, a selection process driven by product quality would imply that older firms, which managed to survive, have higher prices than young exporters. The top right and bottom panels show that most of this growth is in fact driven by selection: specifications using a constant set of firms, display much smaller growth rate of prices. However, prices are surviving firms still appears to slightly increase, mostly in the first five years of their exporting activity. Surprisingly, prices tend

[18] All regression tables are displayed in appendix B.

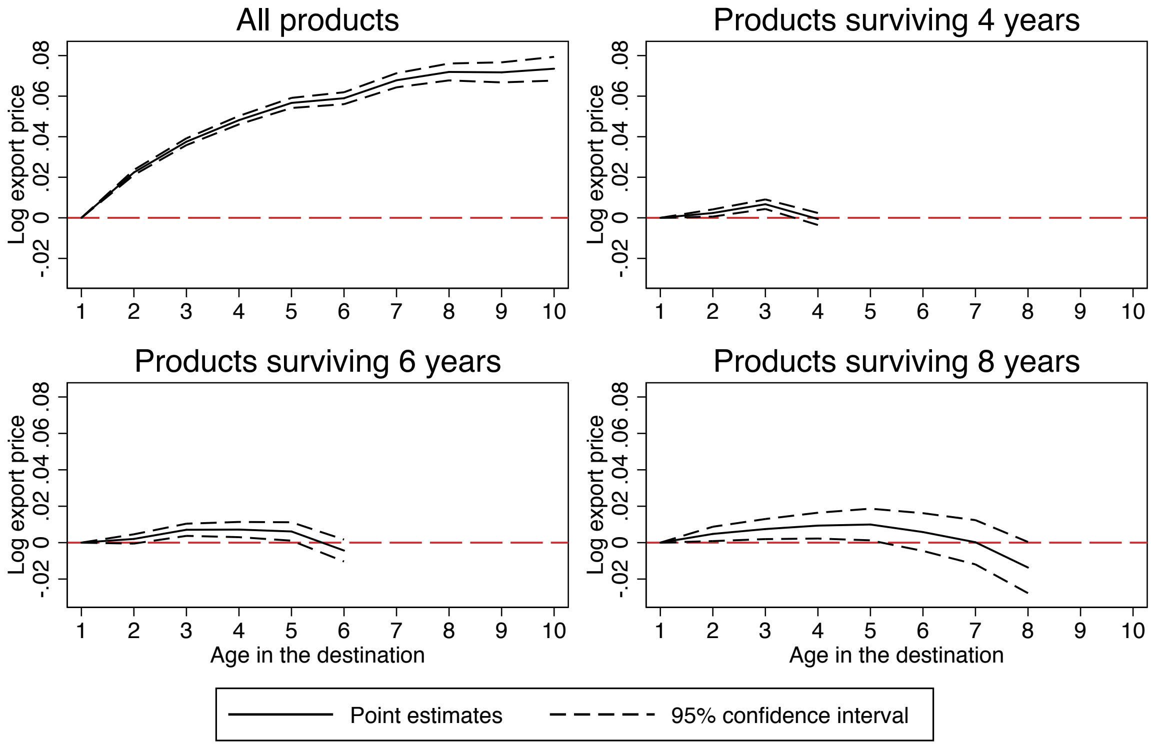

Figure 3: Prices across export ages

Notes: The figure reports the cumulative growth of prices of a firm-product category pair in a destination at different ages. The top left panel uses the entire sample, while the top right and bottom panels only uses products that reach ages 4, 6 and 8 respectively. Standard errors clustered at the firm-product-destination level.

to decrease as the product reaches a certain age. One possible explanation for this decrease would be the existence of product life cycle. After several years, new products entering the market might use a different technology or be of higher quality. As a consequence, these products are likely to be more expensive to account for these new characteristics, implying a negative relationship between ages and prices.

The main issue with the specification displayed in figure 3 is that it identifies price dynamics in the cross-section, comparing products of different ages. Therefore, prices variations could be explained by heterogeneity across firms, due to quality or productivity improvements, that are not specific to a foreign destination. In order to control for these confounding effects, I add firm-product-year fixed effects to capture all variations in prices that are common across destinations. Therefore, it is possible to identify price dynamics, by comparing prices charged simultaneously by a firm across destinations with different export experiences. Figure 4 reports the results of this specification.

The top left panel of figure 4 reveals an increasing relationship between prices and ages, where products older than 6 years are 4 to 5 percent more expensive than new products. Once again, this is likely to be driven by dynamic selection as only high quality products might survive in foreign markets. When looking at this relationship for surviving products, accounting for selection bias (top right and bottom panels), the price increase is more modest: I find that prices are 2 percent higher after 4 years relatively to the first years of exporting. Interestingly, this

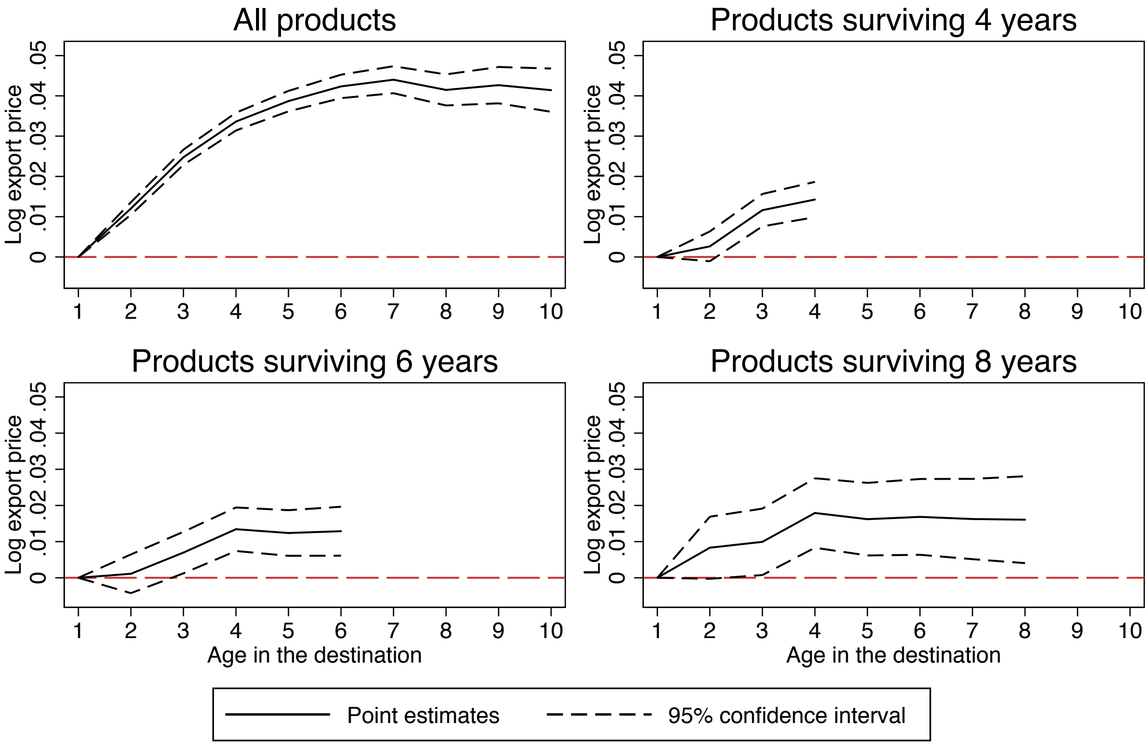

FIGURE 4: Prices across export ages, across destinations

Notes: The figure reports the cumulative growth of sales compared to age one, of a firm-product category pair in a destination at different ages. The regression uses logarithm of prices as dependent variable, and includes product category×destination×year and firm×product×year fixed effects. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product level.

increase in prices appears to only take place in the first four years of exporting. After those years, export prices flatten.

A series of recent papers have studied the relationship between prices and age in the market. Foster et al. (2016) finds similar patterns for commodity products using a specification similar to figure 3. Macchiavello (2010) also find increasing prices when looking specifically within relationships between wineries and distributors. Concurrently, Fitzgerald et al. (2016) also studies the relationship between prices and ages using Irish customs data. However, they do not find any significant dynamic in prices when using a specification similar to figure 4. Given the robustness of their findings, the difference might come from differences in the nature of exported goods between France and Ireland. Upon inspection, the positive trend in prices especially takes place in manufacturing industries such as Machinery, Textiles, Metal products, which are less prevalent in Irish exports relative to France.[19] Importantly, they also conclude to the prevalence of destination-specific demand factors as source of growth for exporters.[20]

Overall, these results highlight the predominance of selection in understanding exporters'

[19]Berman et al. (Forth.) and Timoshenko et al. (2017) also study the relationship between prices and export experience and find a decreasing trend. However, they do not account for dynamic selection in their specifications.

[20]The main result of the structural estimation, regarding the importance of demand frictions when estimating fixed costs, does not require increasing prices but non decreasing prices. However, I develop a model that can account for these observed prices dynamics.

dynamics. We observe very small survival rates for young exporters. As a consequence, it is tedious to identify the underlying processes that drive sales and prices across ages. However, we can draw three conclusions from these figures. First, given the low survival rates of young exporters, sunk entry costs cannot be the main barrier to export. Second, export values grow rapidly with experience in foreign markets, despite a wide heterogeneity across firms. Third, export prices do not explain this growth in export values. On the contrary, prices slightly increase with age in export markets.

These conclusions guide the empirical model developed below. I specifically develop an entry model to control for the endogenous sorting and attrition of firms, and recover the processes that drive the observable variables of the model. Moreover, I introduce a consumer accumulation channel that can explain the rise of sales across years, while introducing frictions to grow as an alternative barrier to export. Finally, I allow firms to use prices to foster this accumulation channel, which could explain the observed growth of export prices with age. The next section introduces this model.

3 Structural model of export entry

This section describes an empirical model of entry into foreign markets in which the accumulation of consumers creates a new source of dependence in the dynamic problem of the firm. Because this model aims at identifying the different sources of heterogeneity in firms' decisions, it is crucial to allow for unobserved and persistent heterogeneity across firms and destinations. In particular, persistent heterogeneity is the main competing hypothesis to sunk entry costs to explain persistence in export decisions. As a consequence, the model features two additional sources of persistence at the firm level - productivity and product appeal - and one persistent characteristic specific to destinations - their aggregate demand. Therefore, a potential profit for a firm-destination pair depends on four characteristics: productivity, product appeal, aggregate demand and consumer share.[21]

The introduction of consumer accumulation implies two deviations from the standard dynamic model: first, firms start small in a new market. Their sales and profit will rise in the following years as they accumulate more consumers. Second, because part of this consumer accumulation comes from sales, firms have dynamic incentives to lower their prices to foster their future demand.

I start by describing the demand schedule of the firm and how the accumulation of consumers affects the demand from foreign destinations. After introducing the costs associated with the production process, I solve the dynamic problem of the firm to study the consequences of this consumer margin on entry and pricing decisions.[22]

[21] I assume that entry decisions are independent across destinations, once controlling for firms' characteristics. McCallum (2015) provides support for this assumption. See also Morales et al. (2015) which uses moments inequalities to handle such a large state space.

[22] Note that I do not study the choices made by the firms for each product it produces. Firms are seen as single-good producers, and will be considered as such in the empirical application.

3.1 Demand

A large literature in industrial organization has found empirical evidence of inertia in consumption and state dependence in demand. This literature also points out the large number of mechanisms that can generate this dependence in demand, as well as the difficulty to empirically disentangle these different channels.[23] In order to keep the model tractable, I introduce state dependence in demand through the existence of a firm-specific customer base in each destination. This customer base, denoted n_{fdt}, describes the share of consumers, in a destination d at time t, that includes the product f in their consideration set,[24] which is consistent with the idea of customer margin introduced in the macroeconomic and international trade literature.[25]

Therefore, I assume that new exporters have an initial share of consumer n_0 when they enter a foreign destination. In the subsequent years, the consumer awareness of the products propagates through two mechanisms. First, the sales of a product increase its awareness in the next period. Specifically, a euro increase in the sales of a product increases by η_1 the potential share of consumers in the next period. This mechanism can arise in situations in which consumers have imperfect information about product characteristics, and therefore use sales as a signal for the expected utility gain from consuming a good.[26] Second, consumer accumulation also comes from word-of-mouth: I assume that each aware consumer share its awareness with η_2 consumers. Both of these mechanisms generate a potential growth in the share of consumers for the firm. However, because some of these reached consumers are already aware of the existence of the product, this acquisition of new consumers is discounted by a factor $(1-n')^{\psi}$ with $\psi > 0$, such that the marginal effect of sales s and consumer share n on the future share n' are

$$\begin{aligned} \frac{\partial n'}{\partial s} &= \eta_1 (1-n')^{\psi}, \\ \frac{\partial n'}{\partial n} &= \eta_2 (1-n')^{\psi} \end{aligned} \tag{2}$$

This specification is largely inspired from the marketing literature described in Arkolakis (2010): the accumulation of consumers has decreasing returns such that it is more difficult for an established firm to reach new consumers. For this firm, a significant share of reached consumers are already part of its consumer share, hence not contributing to its growth. Therefore, the parameter ψ describes the importance of these decreasing returns, while parameters η_1 and η_2 characterize the importance of the two sources of accumulation.

Importantly, these two margins of growth generate different optimal responses by the firm. In a world with word-of-mouth, where consumers learn from their neighbors, the growth of the consumer share can be seen as exogenous, only based on the past share of consumers. In this world, firms cannot affect this accumulation with their pricing decisions.[27] However, in a world

[23] One can cite habits in consumption, costly search, or imperfect information as mechanisms leading to state dependence in demand (see for instance Dubé, Hitsch, and Rossi (2010) for a paper distinguishing and measuring the contribution of these different mechanisms).

[24] The marketing literature defines a consideration set as the set of products that consumers consider when making purchase decisions. See for instance Shocker et al. (1991).

[25] See for instance Drozd and Nosal (2012) and Gourio and Rudanko (2014) for macroeconomic papers, and Arkolakis (2010) in international trade.

[26] With CES preferences, the amount spent for a specific good is proportional to the utility gain obtained from the consumption of this good.

[27] This model does not take into account advertising as a source of growth. Fitzgerald, Haller, and Yedid-Levi

where consumers face uncertainty regarding product characteristics and sales are seen as a signal, firms have incentives to reduce their price in order to foster the accumulation of consumers.[28]

Adding an initial condition to these differential equations, $n(0,0) = \underline{n}$, we obtain the following law of motion for the consumer share of a firm f, at date t and destination d:

$$n_{fdt} = 1 - \left[(1-\underline{n})^{1-\psi} - \eta_1(1-\psi)s_{fdt-1} - \eta_2(1-\psi)n_{fdt-1}\right]^{\frac{1}{1-\psi}} \tag{3}$$

Therefore, the share of consumers today n_{fdt} depends on the sales s_{fdt-1} and the share of consumers n_{fdt-1} in the previous period in this market.

This share of consumer acts as a demand shifter for the firm since it scales the demand firms receive from each destination. Moreover, we assume that each consumer displays CES preferences over its consideration set. Denoting Ω_i the consideration set of a consumer i, its utility function is

$$U_i = \left[\sum_{f\in\Omega_i} \exp\left(\frac{1}{\sigma}\lambda_f\right) q_{if}^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \qquad \text{with } \sigma > 1, \tag{4}$$

where q_{if} is the quantity consumed of good f and λ_f the appeal of the product. This consumer i maximizes this utility function given a budget y_i devoted to this set of goods, and prices $\tilde{p}_f$. As solution of this optimization, the quantities q_{if} demanded by consumer i for a good f are

$$q_{if} = \begin{cases} \exp(\lambda_f)\tilde{p}_f^{-\sigma}P^{\sigma-1}y_i & \text{if } f \in \Omega_i \\ 0 & \text{if } f \notin \Omega_i \end{cases}$$

where P is the CES price index faced by the representative consumer.[29] Aggregating the demand of individual consumers from each destination d and time period t, the demand received by firm f from destination d at time t is:

$$q_{fdt} = q(\lambda_{ft}, X_{dt}, n_{fdt}, p_{fdt}, \varepsilon^D_{fdt}) = n_{fdt}\exp(\lambda_{ft} + X_{dt} + \varepsilon^D_{fdt})p_{fdt}^{-\sigma} \tag{5}$$

where X_{dt} captures all the aggregate variables of the demand shifter,[30] p_{fdt} is the factory price of the good, and ε^D_{fdt} is a random demand shock.

It is important to note that the appeal of the product λ_{ft} does not vary across destinations. Given the existence of an aggregate demand shifter, this implies that firms cannot vary the relative quality or appeal of their good across destinations. Therefore, this specification can still explain that firms provide different product appeal in different destinations, as long as these differences are common across firms. This assumption is fundamental to explain the identification

(2016) provides a model in which firms invest in their consumer base.

[28] This distinction echoes differences between structural and spurious structural dependences (Heckman, 1981), that generate different optimal responses by firms.

[29] With different sets of goods, each consumer has a different price index. However, I follow Arkolakis (2010) by assuming that each consumer has probabilistically an equivalent set of goods, such that all consumers face the same price index $P = \left[\sum_{f\in\Omega} n_f \exp(\lambda_f)\tilde{p}_f^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$.

[30] $X_{dt} \equiv \log Y_{dt} - (1-\sigma)\log P_{dt} + (1-\sigma)\log(\tau_{dt}e_{dt})$ where $Y_{dt} \equiv yN_{dt}$ are total expenditures from a number of consumers N_{dt}, and τ_{dt} and e_{dt} are respectively iceberg transportation costs and exchange rates that converts the factory price to the consumer price.

assumption of the model: while λ_{ft} and X_{dt} are respectively firm and destination specific, the customer share n_{fdt} will be identified through the sales of a firm in a specific destination. After describing the demand faced by firms, I now turn to the costs associated with production and international trade.

3.2 Technology and costs

The costs associated with production and international trade are similar to those traditionally assumed in the literature. I first describe the constant marginal costs of production, then the fixed costs associated with the exporting activity.

First, I assume constant marginal costs of production. These marginal costs are a decreasing function of the firm productivity ϕ_{ft}, and vary with the appeal of the product λ_{ft}. Moreover, I assume the existence of non-persistent productivity shocks ε^S_{fdt}, and I allow costs to vary with destination markets by including a set of coefficients γ_g.[31] Formally, the marginal cost function is

$$c_{fdt} = c(\phi_{ft}, \lambda_{ft}, \varepsilon^S_{fdt}) = \exp(-\phi_{ft} + \alpha\lambda_{ft} + \gamma_g + \varepsilon^S_{fdt}) \tag{6}$$

In addition to these production costs, I assume that firms need to pay entry and per-period fixed cost for each destination they respectively enter or export to. These fixed costs are defined as follows

$$FC(\mathcal{I}_{fdt-1}, \nu_{fdt}) = \begin{cases} f^c_g + \nu^c_{fdt} & \text{if} \quad \mathcal{I}_{fdt-1} = 1 \\ f^e_g + \nu^e_{fdt} & \text{if} \quad \mathcal{I}_{fdt-1} = 0 \end{cases}$$

where $\mathcal{I}_{fdt}$ is a dummy that equals one if the firm f is active (records positive sales) in destination d at time t, and ν_{fdt} are random shocks on fixed costs. Note that these fixed costs will vary across groups g of destinations.[32] Moreover, I assume that shocks ν^c_{fdt} and ν^e_{fdt} follow a logistic distribution with respective variance parameters σ^c_ν and σ^e_ν. These shocks allow the model to rationalize all observed decisions made by the firms.

3.3 Profit and value function

From the demand received by the firm, and the costs of production, I derive the potential profit of the firm for each destination market. After describing the timing of a typical period, I define the entry problem of the firm, and the associated value functions. This dynamic problem depends on five variables that define the state space of the problem: the exogenous variables - product appeal λ, productivity ϕ and aggregate demand X - the share of consumer n, and the presence in the market in the previous year $\mathcal{I}_{-1}$.

In this model, firms decisions are limited: they decide whether to be active on the market, and the price they charge if they decide to export. Consequently, the appeal of the product, the productivity and the aggregate demand from each destination are exogenous but persistent variables that potentially capture the hysteresis of the exporting decisions. For ease of exposition,

[31] Destination markets will be geographically divided in three groups indexed by g.

[32] For instance I assume that fixed costs are equal for all European destinations.

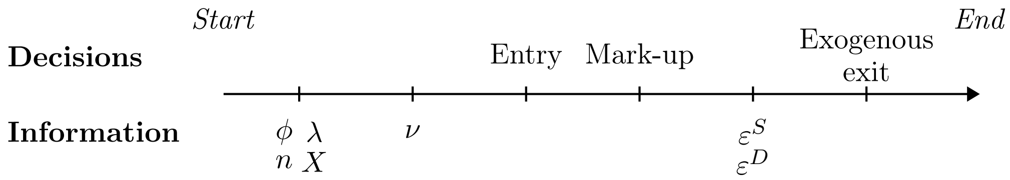

Figure 5: Timeline of a period

I denote these variables $\xi \equiv (\lambda, \phi, X)$ such that, ignoring the subscripts and the parameters of the model, the profit function of a firm is

$$\begin{aligned}\Pi(\xi, n, p, \varepsilon, \mathcal{I}_{-1}, \nu) &= q(\xi, n, p, \varepsilon^D)\left[p - c(\xi, \varepsilon^S)\right] - FC(\mathcal{I}_{-1}, \nu) \\ &= \pi(\xi, n, p, \varepsilon) - FC(\mathcal{I}_{-1}, \nu)\end{aligned}$$

This profit function is made of a variable profit and fixed costs. Despite a CES demand, this variable profit could be negative because of the dynamic nature of the pricing decision of the firm: some firms could set a price lower than their marginal costs to foster future demand. The second part of the profit function comes from the fixed costs of exporting $FC(\mathcal{I}_{-1}, \nu)$ that depend on the past presence of the firm on the market and the profit shock ν, which allow the empirical model to explain entry and exit decisions of firms that cannot be rationalized by variable profits.

This profit is earned if the firm decides to be active on the market at this period. To study this decision, figure 5 defines the timeline of a typical period, which provides the timing at which decisions are made and the information sets available to the firms when they take decisions. At the beginning of the period, the firm observes its exogenous variables, λ, ϕ, n and X. After realization of the profit shock ν, it decides whether to export in the market. If the firm decides to export, it optimally chooses the mark-up to charge over their marginal costs.[33] Finally, sales and prices are obtained after observing the realization of the non-persistent shocks ε and the firm has to exit with probability δ if it has decided to export this year.[34]

Therefore, denoting μ the multiplicative mark-up of the firm such that $p = \mu c$, the value function of the firm can be defined as the following:

$$\begin{aligned}V(\xi, n, \mathcal{I}_{-1}) &= (1 - \delta\mathcal{I}_{-1})E_\nu \max\left\{V_I(\xi, n) - FC(\mathcal{I}_{-1}, \nu)\ ;\ V_O(\xi)\right\} + \delta\mathcal{I}_{-1}V_O(\xi) \\ \text{with} \qquad V_I(\xi, n) &= \max_\mu\left\{E_\varepsilon\left\{\pi(\xi, n, \mu, \varepsilon) + \beta EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1)\right\}\right\}, \\ V_O(\xi) &= \beta EV'(\xi, n_0, 0), \\ EV'(\xi, n', \mathcal{I}) &= \int_{\xi'} V(\xi', n', \mathcal{I})\, dF(\xi'|\xi).\end{aligned} \tag{7}$$

For each market, the firm chooses between exporting $V_I(\xi, n) - FC(\mathcal{I}_{-1}, \nu)$ and being inactive

[33] Choosing the mark-up rather than the price facilitates the computation of the solution, while allowing for structural shocks ε in demand and costs.

[34] Theses assumptions are mostly driven by the construction of the empirical model. The realizations of the shocks ε after the markup decision generate structural errors that can explain observed sales and prices variations. Moreover, the exogenous exit is necessary to explain the occasional exit decision of very large exporters in the data.

$V_O(\xi)$. By being inactive, the firm makes no profit today but retains the possibility to update its decision in the next period. In contrast, when exporting, it obtains a present profit that depends on the shocks ε and the mark-up chosen by the firm. Moreover, the firm has a continuation value, $EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1)$, characterized by a stock of consumer n' and lower fixed costs to pay in the next period. This continuation value is constructed from the transition of the exogenous variables $F(\xi'|\xi)$, and the expected value of $V(\xi, n', I)$.

3.4 Firms' decisions: entry and pricing.

After defining the problem of the firm, I can now derive the optimal entry and pricing decisions of the firm. Because the accumulation of consumers is based on the sales of the firm, the optimal price charged by the firm deviates from a standard constant mark-up. Instead, firms optimally reduce their mark-up to account for the accumulation of consumers. Because this pricing decision is taken once the firm has decided to enter, I start by describing the optimal mark-up charged by the firm. By backward induction, I then infer the expected profit of the firm and solve for the value and probability of exporting.

Optimal mark-up The firm's choice of mark-up is made after entry, in order to maximize the sum of the present profit and the continuation value of exporting:

$$V_I(\xi, n) = \max_{\mu} \Big\{ E_\varepsilon \Big\{ \pi(\xi, n, \mu, \varepsilon) + \beta EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1) \Big\} \Big\}$$

such that the optimal price chosen by the firm is:

$$p(\xi, n) = \frac{\sigma}{\sigma - 1} \frac{1}{1 + \beta \int \omega(\varepsilon) \eta_1 (1 - n')^{\psi} \frac{\partial EV'(\xi, n', 1)}{\partial n'} dF(\varepsilon)} c(\xi, n) \tag{8}$$

with $\omega(\varepsilon) = \frac{\exp(\varepsilon^D + (1-\sigma)\varepsilon^S)}{\int \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) dF(\varepsilon)}$.[35] The optimal mark-up charged by the firm has two components. First, the firm applies the standard CES mark-up $\frac{\sigma}{\sigma-1}$ based on the price-elasticity of demand. Second, the firm applies a discount factor based on the dynamic incentives it has to lower its price to attract more consumers in the future. This factor depends on two elements: first, how much this increase in sales increases its consumer share tomorrow, $\eta_1(1 - n')^{\psi}$; this element induces lower mark-ups for small or young firms that benefit from higher returns of accumulation. Second, the extent of this discount also depends on the impact of this increase in the future consumer share on the continuation value $\frac{\partial EV'(\xi, n', 1)}{\partial n'}$. This effect is not linear but hump shaped with the profitability of the firm:[36] young firms that are unlikely to survive have no incentives to invest in future consumers, while firms that can use extra consumers to increase their survival probability get the largest benefits from increasing their consumer share. Finally, note that this equation defines the unique optimal price charged by the firm but only through an implicit function, since the future share n' depends on the price charged by the firm.[37]

[35] See appendix C for derivations.

[36] This comes directly from the probability of exit that makes the value function of the firms convex for low profitability firms, and concave for higher profit firms.

[37] In the estimation procedure, the dynamic problem of the firm is solved on a grid. Therefore, I do not use this formula to find the optimal mark-up but instead, pick the mark-up in the grid that maximizes the value function

Consequently, the accumulation of consumers implies heterogeneous mark-ups across firms, depending on their current share of consumers, and their expectations on future profits. Having described the optimal mark-up of the firm, I can now study its entry decision.

Entry condition Knowing the expected option values of being active or inactive, I can now study the entry decision of the firm. Firm pick the most profitable option, after observing the shock ν that affects the fixed costs of being active on a market. From the logistic distribution, the expected value of the firm before observing the shock ν is

$$V(\xi, n, \mathcal{I}_{-1}) = (1 - \delta\mathcal{I}_{-1})\sigma_\nu \log\left[\exp\left(\frac{1}{\sigma_\nu}(V_I(\xi, n) - f)\right) + \exp\left(\frac{1}{\sigma_\nu}V_O(\xi)\right)\right] + \delta\mathcal{I}_{-1}V_O(\xi) \quad (9)$$

in which f and σ_ν can be respectively f^e or f^c and σ_ν^e or σ_ν^c, depending on the value of $\mathcal{I}_{-1}$. This equation closes the dynamic problem of the firm, by providing the fixed point that defines the value function $V(\xi, n, \mathcal{I}_{-1})$. Moreover, the probability for a firm to be active, before the realization of the fixed cost shock ν, is,

$$P(\mathcal{I} = 1|\xi, n, \mathcal{I}_{-1}) = \frac{1 - \delta\mathcal{I}_{-1}}{1 + \exp\left(-\frac{1}{\sigma_\nu}(DV(\xi, n) - f)\right)} \quad (10)$$

with $DV(\xi, n) = V_I(\xi, n) - V_O(\xi)$. This last equation predicts the probability of entry of a firm, conditional on its current characteristics, described by ξ, n and $\mathcal{I}_{-1}$. While n and $\mathcal{I}_{-1}$ are endogenous, ξ are exogenous and unobservables variables. Therefore, to finish the derivation of the model, it is necessary to describe the evolutions of these exogenous variables across time. These evolutions will be important to compute the expectation of the value functions, $EV'(\xi, n, \mathcal{I}_{-1})$, as well as disciplining the variations of sales and prices across times in the empirical application.

3.5 Evolution of exogenous variables

Most of the hysteresis in exporting decisions is likely to come from the persistence over time of firms characteristics Therefore, it is necessary to allow these processes to be time-varying and persistent. Therefore, I assume that exogenous variables follow AR(1) processes, with flexible parameters, such that:

$$\begin{aligned} \lambda_{ft} &= \rho_\lambda \lambda_{ft-1} + \sigma_\lambda \varepsilon_{ft}^\lambda \\ \phi_{ft} &= \mu_\phi + \rho_\phi \phi_{ft-1} + \sigma_\phi \varepsilon_{ft}^\phi \\ X_{dt} &= \mu_X + \rho_X X_{dt-1} + \sigma_X \varepsilon_{dt}^X \end{aligned} \quad (11)$$

where the ε shocks follow a normal distribution with zero mean and unit variance. Note that, by normalization, λ is centered around zero: since both X and λ enters linearly in the demand function, it is not possible to separately identify their respective means.

Finally, I need to impose distributional assumptions on the initial conditions of these unobservables. I assume that the distributions of product appeal and productivity are stable over time such that the initial distributions are constrained by a stationary assumption. However,

of the firm. See appendix D for details.

I assume that the variation in aggregate demand across destinations does not arise from a stationary distribution, such that $X_{d0} \sim N(\mu_{X_0}, \sigma_{X_0})$. Moreover, I assume that the initial share of consumers follow a Beta distribution with parameters 1 and 5.[38]

This concludes the derivation of the model. Each firm observes exogenous variations in its export profitability through variation in its productivity, product appeal and the demand in each destination. Based on these variations, firms decide to enter or exit various destinations, in which they decide at which price to sell their good.

3.6 Restricted model

In order to assess the importance of consumer accumulation on estimated trade costs and aggregate response to trade, I also estimate a restricted version of the model that does not feature this mechanism. This restricted model is equivalent to assuming that exporters have a consumer share n_{fdt} equal to one when they are active on the market. As a consequence, firms do not have incentives to deviate from the CES pricing, and mark-ups are similar across all firms.

This restricted version of the model can be seen as the canonical model used in the literature. In this model, firm-level heterogeneity and entry costs of exporting explain the hysteresis in exporting. This model can be seen as a dynamic version of Melitz (2003), as estimated by Das et al. (2007). Estimating this restricted model is essential to assess the importance of consumer accumulation on the outcomes of the estimation and the aggregate implications of the model.

4 Estimation

In this section, I describe the procedure to estimate the parameters of the model. I start by describing the likelihood of the problem, based on the three structural equations linked with the observable variables (sales, prices and participation to export). I then turn to the algorithm, showing the advantages of a MCMC estimator to account for persistent and unobserved heterogeneity and solve the dynamic problem of the firm. Finally, I provide intuition behind the identification of parameters and unobservables of the model.

4.1 Likelihood

I start by presenting the likelihood that is obtained from the three main equations of the model: the price and demand equations that feature the stock of consumers and the dynamic mark-up charged by the firm, and the entry probability that describes the exporting decision on each destination.

First of all, the demand and price equations (5), (6) and (8) are taken in logarithm to obtain

$$\log s_{fdt} = \log n_{fdt} + \lambda_{ft} + X_{dt} + (1-\sigma)\log p_{fdt} + \varepsilon^D_{fdt}$$
$$\log p_{fdt} = -\phi_{ft} + \alpha\lambda_{ft} + \log \mu(\xi, n_{fdt}) + \gamma_d + \varepsilon^S_{fdt}$$

[38]This only matters for firms that record positive sales the year before the beginning of the sample. An alternative specification would be to allow this initial stock to be correlated with initial unobservables. However, given the small number of firms in this case and the length of the panel (14 periods), this assumption has little consequences on the estimation.

This block constitutes the first part of the likelihood. Assuming that ε follows a bivariate normal distribution with variance Σ, I define this likelihood block as $L_{\varepsilon}(s_{fdt}, p_{fdt}|\xi_{fdt}, n_{fdt}; \Theta)$, with Θ being the full set of parameters, such that

$$
\begin{aligned}
L_{\varepsilon}(s_{fdt}, p_{fdt}|\xi_{fdt}, n_{fdt}; \Theta) = G_{\Sigma}\Big(& \log s_{fdt} - \log n_{fdt} - \lambda_{ft} - X_{dt} - (1-\sigma)\log p_{fdt} \\
& ;\ \log p_{fdt} + \phi_{ft} - \alpha\lambda_{ft} - \log \mu(\xi_{fdt}, n_{fdt}) - \gamma_d \Big)
\end{aligned} \tag{12}
$$

where G_{Σ} is the density function of a bivariate normal distribution with means zero and variance matrix Σ.

The second block of the likelihood is based on the entry decision of the firm. Equation (10) defines the probability to enter for a firm, based on its set of unobservables ξ, its stock of consumer n and its past exporting activity. I denote this function $L_{\nu}(\mathcal{I}_{fdt}|\xi_{fdt}, n_{fdt}, I_{fdt-1}; \Theta)$ that is obtained from the binary choice made by the firm

$$
\begin{aligned}
L_{\nu}(\mathcal{I}_{fdt}|\xi_{fdt}, n_{fdt}, \mathcal{I}_{fdt-1}; \Theta) = & \left[\frac{1 - \delta\mathcal{I}_{fdt-1}}{1 + \exp\left(\frac{1}{\sigma_{\nu}}\left(-DV(\xi_{fdt}, n_{fdt}) + f\right)\right)} \right]^{\mathcal{I}_{fdt}} \\
& \times \left[\delta\mathcal{I}_{fdt-1} + \frac{1 - \delta\mathcal{I}_{fdt-1}}{1 + \exp\left(\frac{1}{\sigma_{\nu}}\left(DV(\xi_{fdt}, n_{fdt}) - f\right)\right)} \right]^{1-\mathcal{I}_{fdt}}
\end{aligned} \tag{13}
$$

where function $DV(\xi_{fdt}, n_{fdt})$ and f are defined as previously. Therefore the total likelihood for a given observation $D_{fdt} \equiv \{s_{fdt}, p_{fdt}, \mathcal{I}_{fdt}\}$ is the product of the two densities $L_{\varepsilon}(.)$ and $L_{\nu}(.)$.

To obtain the unconditional likelihood, that does not depend on the unobservables, it is necessary to integrate out this set of unobservables. Since these unobservables are persistent over time, the likelihood of the entire dataset D is obtained by repeatedly integrating the unobservables from period T to 0:

$$
\begin{aligned}
L(D|\Theta) = \int_{n_{-1}} \int_{\xi_0} ... \int_{\xi_T} \prod_{f,d} & L(D_{fdT}|D_{fdT-1}, \xi_{fdT}) \times ... \times L(D_{fd0}|D_{fd-1}, \xi_{fd0}, n_{fd-1}) \\
& dF(\xi_{fdT}|\xi_{fdT-1}) \times ... \times dF(\xi_{fd0})\, dF(n_{fd-1})
\end{aligned}
$$

where $F(\xi_{fd0})$ and $F(n_{fd-1})$ are defined by the initial unobservables density function, and D_{fd-1} the observables previous to the estimation sample. After describing the likelihood of the problem, I now turn to the estimation procedure to obtain the posterior distribution of parameters Θ.

4.2 Algorithm

To estimate the model, I develop a Markov Chain Monte Carlo (MCMC) estimator to tackle the two important difficulties in evaluating the likelihood: integrating the numerous integrals and solving the dynamic problem of the firm.[39]

[39] The literature on dynamic discrete choices model, starting from Rust (1987), is mostly devoted to this second problem. This problem can be largely simplified using the mapping between conditional choice probabilities and value functions, as highlighted in Hotz and Miller (1993). However, in my application, state variables are not observed, hence complicating the estimation of conditional choice probabilities.

In order to circumvent these difficulties, I employ a particle MCMC estimator, taking advantage of recent Bayesian techniques to sample the posterior distribution of the parameter Θ, conditional on the data. The choice of a Bayesian estimator relies on two recent methods developed in the Bayesian literature. First, I employ a particle filter to perform the integration of the unobservables. In particular, I use the particle Gibbs with ancestor sampling described in Lindsten et al. (2014) following an influential article from Andrieu, Doucet, and Holenstein (2010): the idea is to use a particle filter to update the set of unobservables in a Gibbs fashion, conditional on current unobservables and parameters. This sampling technique allows me to develop a MCMC estimator in which parameters and unobservables are alternatively sampled conditional to each other. Second, to overcome the computational burden of solving the value functions in the likelihood, Imai, Jain, and Ching (2009) and Norets (2009) show how to take advantage of the iterative feature of the MCMC estimator to solve the value functions, by iterating the Bellman equation at each iteration of the Markov chain. The intuition is that the value function can be approximated at the early stages of the Markov chain: by using this value function as initial value in the next iteration, the value function will converge toward the fixed point defined by the contraction mapping as the Markov chain converges and explores the posterior distribution of Θ.

Overall, the MCMC estimator explores the posterior distribution of the parameters Θ. This distribution is proportional to the product of the likelihood and the prior distribution such that

$$P(\Theta \,|\, D) \propto \int_{\xi} L(D \,|\, \xi, \Theta)\, dF(\xi \,|\, \Theta)\, P(\Theta) \tag{14}$$

To avoid the influence of priors in the parameters estimation, I assume flat priors except for values of parameters that do not satisfy theoretical or stationarity constraints.[40] The goal of the Markov Chain is to repeatedly sample from the posterior distribution according to (14). Given the large number of parameters (30), this is achieved by sequentially updating blocks of parameters and unobservables: in particular, I divide parameters in four blocks. One consists of the parameters from L_ε, which characterize the supply and demand equations, one is related to the law of motion of n(.), one of the parameters of the different laws of motion of ξ, and a final block consists of the dynamic parameters from L_ν. Therefore, a typical iteration s of the Markov Chain consists of updating three different objects: the value function $V^{(s)}(\Theta^{(s)})$, the set of unobservables, $\xi^{(s)}_{fdt} = (\lambda^{(s)}_{ft}, \phi^{(s)}_{ft}, X^{(s)}_{dt})$ and the parameter $\Theta^{(s)}$. I first sample unobservables $\xi^{(s)}$ using the particle Gibbs sampler conditional on $\Theta^{(s-1)}$ and $V^{(s-1)}(\Theta^{(s-1)})$. Then, I update the value function using equation (7) to obtain $V^{(s)}(\Theta^{(s-1)})$. Finally, I sample Θ^s conditional on $\xi^{(s)}$, using either a Metropolis Hastings or Gibbs sampler for the four different blocks of parameters described above.[41]

Two important points are worth point out regarding the algorithm. First, the iteration of the value functions that allows the evaluation of the likelihood are obtained on a grid that is updated throughout the algorithm. From the value function, the specific values of $DV(.)$ and $\mu(.)$ that

[40] I exclude from the support of Θ (or equivalently assigned a prior probability of zero for these values), negative values for the variance parameters, as well as values beyond -1 and 1 for the autocorrelation parameters. I also impose the fixed costs parameters (f, fe) and the parameter ψ to be positive.

[41] More details are provided in appendix D.

enter the likelihood function are obtained by interpolation to be evaluated at any point in the state space.[42] Second, due to the complexity of the estimation procedure, I do not estimate the value of β, the discount rate, and σ the demand elasticity. I set their values at 0.95 and 2 respectively.[43]

After describing the details of the estimation procedure, I provide, in the next section, intuition about the sources of identification of the parameters and the unobservables.

4.3 Identification intuition

Despite the complexity of the algorithm, estimating this model using micro data and a full information estimator provides simple intuitions of parameters' identification.

To describe the sources of identification, it is useful to distinguish the identification of unobservables and parameters. Let's assume first that the parameters are known. In this case, the identification of the unobservables mostly come from a variance decomposition of the demand shifters and prices: the demand shifter is decomposed between a firm-year component (the product appeal λ_{ft}), a destination-year component (the aggregate demand X_{ft}), and a firm-destination-year component (the consumer base n_{fdt}). Once the product appeal is known, the productivity ϕ_{ft} is identified from price variations across firms. Therefore, the identification of the unobservables mostly comes from a decomposition of observables variables, which is straightforward if the parameters are known. Moreover, the hierarchical structure and the entry decisions bring additional information to identify the posterior distribution of these unobservables. For instance, if a firm is not exporting one year, the information from previous and future years will help identify the potential value of the unobservables. Similarly, if a firm only exports to one destination at a given year, the fact that it does not export somewhere else provides information about its product appeal or productivity.

Turning to parameters identification, they can be divided in three groups. The identification of the 15 parameters related to the laws of motion of the unobservables can be easily identified once knowing the values of these unobservables. Regarding the 6 parameters entering the demand and pricing equations, their identification is similar to traditional demand and supply equations: correlation between sales and prices, and prices with destination dummies and product appeal, while the parameters of the variance matrix are obtained from the variance of the unexplained variation in prices and sales. Finally, the 9 parameters related to the entry problem are obtained by comparing potential profits and firms' observed decisions: the number of exporters identifies the per-period fixed costs, the persistence in exporting the entry costs, and the remaining variance in exporting decisions identifies the required variance of these fixed costs' shocks.

Consequently, the identification of the unobservables conditional to the parameters, and of the

[42] I use a grid of 20 points in each dimension of the value function, which leads to a state space of size $20^4 \times 3$ (3 being the number of destination groups.). In appendix D, I show that the procedure performs well when implemented on data simulated with a grid larger than the one used in the estimation.

[43] In theory, the mark-up decisions could generate variations in prices that are orthogonal to the demand shocks ε, and therefore identify the price-elasticity. However, given the difficulty of the estimation procedure, and the importance of this parameter, I rely on a calibrated value. Moreover, note that a value of 2 is in the lower range of estimates found in the literature. The reason is that this elasticity will be amplified by the consumer margin such that the demand elasticity of the firm will be larger in the long-run. Because this feature does not exist in the restricted version of the model, I use a value of 3 in the restricted version, such that the average mark-up is equal across models.

parameters conditional to the unobservables are quite straightforward. The goal of the MCMC estimator is to repeatedly sample each component conditional to the other, in order to obtain their joint distribution. After a necessary period of convergence, the Markov Chain describes the posterior distributions of the parameters from which confidence intervals can be obtained.[44]

5 Results

I implement the estimation on a set of wine exporters from France; the choice of this industry is based on two criteria. First, wine producers only export wine. Therefore, it is reasonable to assume that the entry decisions into foreign destinations are made at the firm level, and it is possible to aggregate sales and prices at the firm level for each destination. Second, the wine industry is a large industry in France and, therefore, I can obtain a large enough sample of exporters with a large set of destinations. In appendix A.2, I describe the specific selection procedure to obtain the estimation sample, and provide summary statistics.

I start by describing the fit of the model relative to the exporters' dynamics presented earlier. Then I present the estimated values of the parameters, and in particular the decrease in entry costs induced by the introduction of the consumer margin.

5.1 Fit of the model

I report in this section the fit of the model regarding the survival rates, sales and prices of the firm-destination pair at different ages. Figure 6 reports the predictions of the model relative to the data. I also report the results of the restricted version of the model, which does not contain a consumer margin.

As reported in figure 6, the full model with consumer accumulation can reproduce the growth in sales across ages (top left figure). This is not surprising since the full model allows for destination-specific growth in sales through consumer accumulation. As the sales of exporters increase with age, their profit also increases. Consequently, the full model can perform better at explaining lower survival rates for young exporters: the average prediction gives a survival rate of 70 percent the first year, and close to 90 percent after 8 years in the foreign markets. However, this growth in sales is not enough to fully explain the low survival rates of young exporters, and, therefore, does not entirely solve the puzzle linked with young exporters dynamics. Importantly, the model cannot match these low survival rates for young firms because of the nature of the estimator: a full-information estimator does not only target moments, but all observations. In this case, the model estimates a large variance in fixed costs to explain the large variance in exit decisions across firms, which also limits differences in survival rates across ages. In this respect, the use of a full information estimator implies a more robust test of the model.

In comparison, the restricted model cannot explain this rise in sales and even less in survival rates: in the restricted model, the predicted survival rate is flatter across ages, between 75 and 85 percent, which is similar to the average survival rate in the sample. However, the predictions

[44]Specifically, I perform 60 000 iterations of the MCMC and discard the first 10 000. Given the complexity of the estimation, I checked its validity using a sample of simulated data. Even though I cannot perform Monte Carlo simulation on many samples, because of the duration of the estimation, the ability of the procedure to recover the true parameters is reassuring. See appendix D for details.

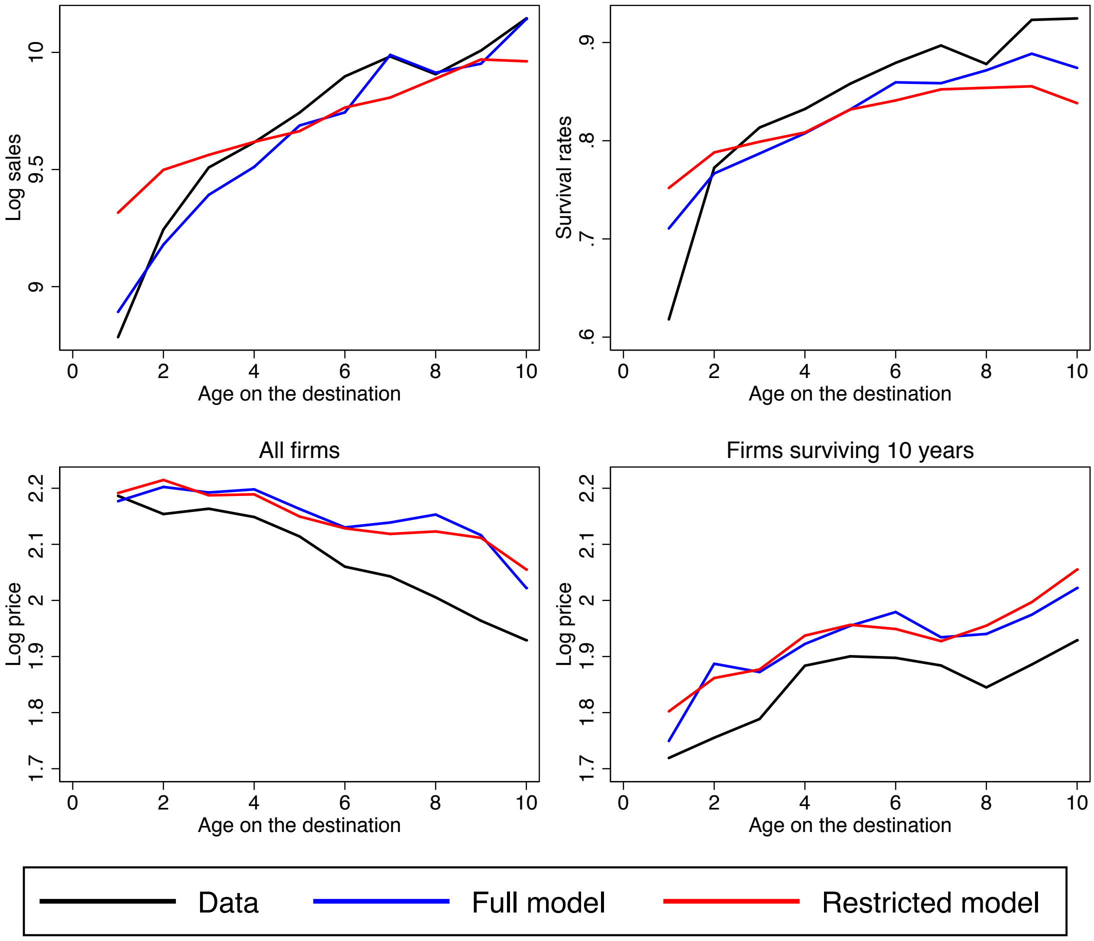

FIGURE 6: Predictions of survival rates, sales and prices across ages.

on prices appear quite similar across models (bottom figures). Both of them can reproduce the decrease in prices with age. When looking at firms surviving 10 years, we can see that the full model can do slightly better in explaining the rise in price with the age of the firm.[45]

After describing the fit of the model, I now turn to the description of the estimated values of the parameters.

5.2 Estimated parameters

The results of the estimation of the model are reported in table 2. I report for each parameter the median of its posterior distribution, and its 90 percent confidence interval. I also report the trace plots of the Markov chain in figure 17 in appendix E, which confirms the good mixing and convergence properties of the chain.

First, looking at the law of motion of the consumer margin, we note that the initial share of consumers at entry (n_0) is relatively small, equal to 6 percent, which leaves a large potential for firms to grow through the accumulation of consumers. This growth is driven both by the past

[45]Prices do not display a significant trend with age in the structural sample. Therefore, it is quite difficult to distinguish and interpret the differences between the two models.

TABLE 2: Estimated parameters

Parameter		**Estimate**	**90% Confidence Interval**	
			Lower bound	*Upper bound*
Continuation fixed costs	Europe	9555	8729	10498
(in euros)	Americas	11255	10212	12556
	Asia/Oceania	12581	11095	14388
Entry fixed costs	Europe	35144	31871	38136
(in euros)	Americas	26708	23989	29467
	Asia/Oceania	27374	24685	30153
Variance of continuation costs	σ_ν^c	13962	11570	16180
Variance of entry costs	σ_ν^e	4541	4033	5059
Exit rate	δ	0.03	0.02	0.04
Law of motion of n	n_0	0.061	0.054	0.068
	$\underline{n}$	0.024	0.020	0.027
	$\eta_1(10^{-6})$	5.60	4.80	6.53
	η_2	0.039	0.014	0.075
	ψ	0.65	0.55	0.74
Law of motion of appeal	ρ_λ	0.97	0.96	0.97
	σ_λ	0.21	0.19	0.23
Law of motion of productivity	ρ_ψ	0.98	0.97	0.98
	σ_ψ	0.08	0.07	0.09
	μ_ψ	-0.03	-0.04	-0.02
Law of motion of agg. demand	ρ_X	0.92	0.88	0.95
	σ_X	0.07	0.06	0.08
	μ_X	1.09	0.62	1.57
	μ_{X_0}	12.94	12.77	13.10
	σ_{X_0}	0.31	0.21	0.46
Elasticity cost of appeal	α	0.41	0.35	0.50
Cost dummies	γ_2	0.23	0.21	0.25
	γ_3	0.18	0.15	0.20
Variance matrix	Σ_{11}	1.30	1.26	1.33
	Σ_{12}	0.16	0.15	0.17
	Σ_{22}	0.15	0.15	0.15

sales of the firm (η_1), as well as the past shares of consumers (η_2), since these coefficients are both significantly larger than zero. Moreover, we can see that the degree of concavity of the law of motion is significant, with a median of the posterior distribution of the coefficient ψ equal to 0.65.

Second, the other unobservables of the model - appeal, productivity and aggregate demand - depict strong degrees of persistence. The coefficients of autocorrelation of the AR(1) processes are estimated to be in average 0.97, 0.98 and 0.92, respectively for the product appeal, the productivity of the firm, and the aggregate demand of the destination. Moreover, quality is costly for the firm as we could expect: the coefficient on quality in the marginal cost function (α)) is 0.41 which implies that firms with higher quality also display higher marginal costs of production and prices.

Finally, because I estimate a structural model of entry, the model is able to deliver euro estimates of the entry and per-period fixed costs paid by an exporter. We see that the obtained fixed costs are relatively low, with the estimated entry cost to an European destination being around 35 000 euros.[46] In addition, a firm will have to pay 9 500 euros every year to keep exporting to this destination. As an element of comparison, the average export value of a firm in my sample to an European destination is 42 000 euros, while the median value is 13 000. One of the reasons for these relatively low numbers is the small variance parameter of these entry costs' shocks, whose the median of the posterior distribution is 4 540. This low number reflects the ability of the model to correctly predict the entry of firms, such that a large variance of these entry costs' shocks is not necessary to rationalize entry decisions. Finally, the model also predicts an exogenous exit rate of 3 percent.

In order to confirm the small magnitudes of these entry fixed costs relative to the literature, I compare in table 3 these parameters with the estimates of the restricted version of the model, which does not feature a consumer margin. The comparison between the two models highlights that entry costs are much larger in the version without consumer margin. For instance, the average entry costs to export to Europe jump from 35 000 to 60 000 euros. Similarly, entry cost estimates to Americas or Asia/Oceania doubles when shutting down consumer accumulation. Moreover, the variance of these entry shocks also increase between the two models. On the contrary, estimates of the continuation fixed costs are roughly similar in the two models.

TABLE 3: Estimated parameters (comparison between models)

Parameter		**Estimates**	
		Full model	*Restricted model*
Continuation fixed costs	Europe	9 555	10 146
	Americas	11 255	11 981
	Asia/Oceania	12 581	13 848
Entry fixed costs	Europe	35 144	60 365
	Americas	26 708	48 158
	Asia/Oceania	27 374	53 984
Variance of continuation costs	σ_ν^c	13 962	17 033
Variance of entry costs	σ_ν^e	4 541	11 844
Exit rate	δ	0.03	0.00
Elasticity cost of appeal	α	0.41	0.20

This decrease in entry costs is explained by two main factors. First, since the model accounts for the fact it will take time for firms to grow, become successful and make large profits in foreign markets, high entry costs are not necessary to deter entry into these markets: accounting for lower sales and lower survival rates in the first years of exporting reduces the expected value of being an exporter, hence reducing the entry costs necessary to match the observed entry patterns. Second, the consumer margin captures some of the state dependence in exporting status, reducing the role played by entry costs in explaining the hysteresis in export decisions. This result will be very important when looking at the models' predictions in response to shocks. Estimating large entry costs to export implies a substantial option value of exporting: large entry costs make entering so difficult that firms will hesitate to exit this market in case of adverse shocks. I study

[46] Prices are normalized across years using a national consumer price index, such that the values are expressed as euros from the year 2000.

these consequences in the next section when comparing the predictions of these models under simulated and observed trade shocks. Interestingly, Dickstein and Morales (2015) find a similar reduction in entry costs when using an estimation method that does not impose restrictions on the information set of exporters.

Another important difference between these two models arises from the estimates of the cost of appeal. In the full model with consumer margin, product quality is relatively costly, with a cost elasticity of 0.41. However, the model without consumer margin identifies product appeal with a lower impact on prices, with an average estimate of 0.2. This difference shows that the consumer margin, by capturing some of the variance in sales, modifies the definition of product appeal: in the presence of consumer accumulation, product appeal becomes more related to product characteristics rather than distribution network for instance. As a consequence, this interpretation of product appeal is more connected to marginal costs than the one identified in the restricted model.[47]

5.3 Other outcomes of the model

I now discuss the evolution with export experience of two important objects introduced in this model: the consumer shares and the mark-up charged by firms.

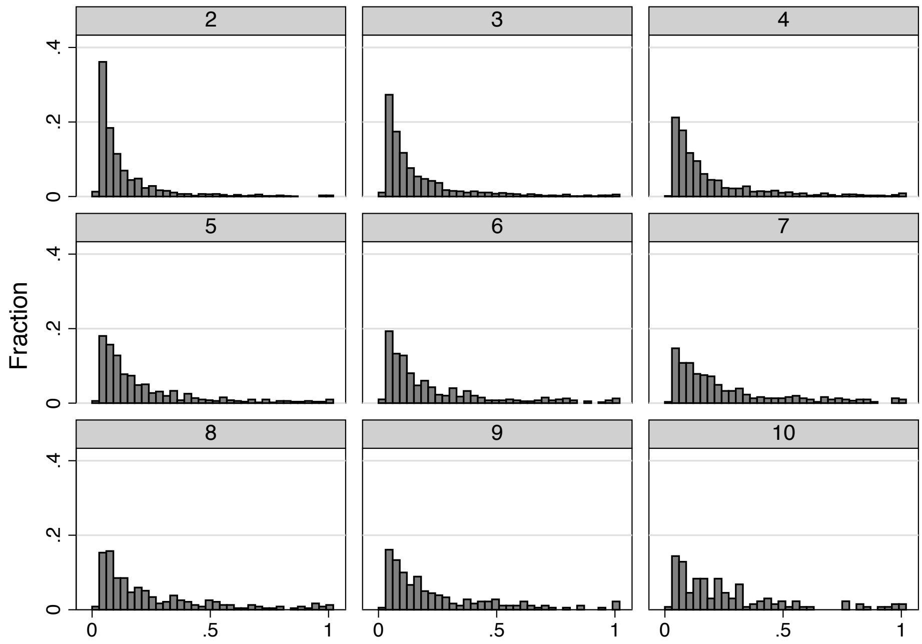

FIGURE 7: Distribution of consumer shares by age

Figure 7 provides the distribution of consumer shares for each age of the firm. Remember that when firms enter, they all have an initial share of approximately 6 percent, which explains why the figure provides distributions from ages 2 to 10. Figure 7 illustrates that the distribution tends to shift toward the right as age increases. One can see that most of the firms have a small

[47]The full set of parameter estimates for the restricted model can be found in table 9 in appendix E.

consumer share at age 2: only a small fraction of them are larger than 25 percent. However, as age increases, more and more firms reach a larger size. Therefore, at age 10, a significant number of them has a consumer share that is larger than 50 percent. However, there is also a large amount of heterogeneity within ages: some firms are large at ages 2 or 3, but a large fraction of them are still small in terms of consumer shares when reaching years 9 or 10. As a result, the overall distributions appear to flatten as age increases, rather than translate toward the left. This implies that the process of consumer accumulation is not identical across firms, and very much relies on the individual sales of the firm rather than an exogenous increase of consumers with age. Some firms will never reach a large fraction of consumers, because it is not profitable for them to do so.

I then turn to the distributions of mark-ups charged by the firms, that can be used by firms to foster consumer accumulation. Figure 8 reports the distributions of mark-ups, separately for each age from 1 to 9. One can see that, similar to the consumer shares, there is a large heterogeneity in mark-ups across ages, but also within ages: the model does not imply a mechanical correlation between mark-ups and age. However, we can see that firms tend to price more aggressively at a young age, in comparison to more established firms. The reason is twofold: first, these firms are small and therefore benefit from large returns of higher sales on consumer accumulation. Second, because these firms are small and young, additional consumers are crucial to survive in the following years.

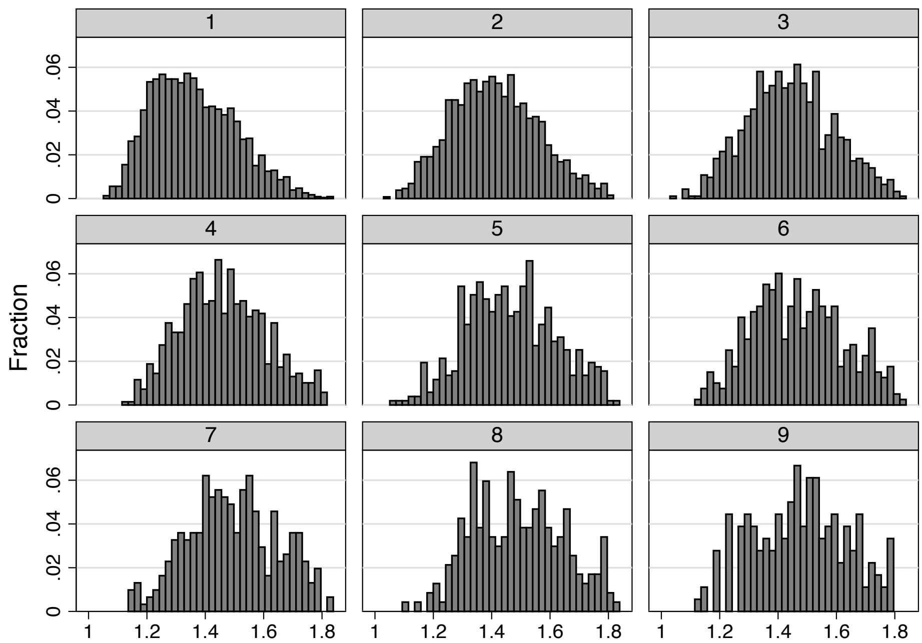

FIGURE 8: Distribution of mark-ups by age

However, it also appears that the pricing behavior does not vary mechanically with ages: some old firms appear to price almost as aggressively as young firms. This result is due to the fact that survival rates do not change that much across ages, hence forcing firms to maintain low prices after a few years to increase their likelihood to survive. Overall, when disentangling the contributions of different factors in prices dynamics, the dynamic pricing behavior leads to a 10% increase in prices between age 1 and 10.

In the next section, I explore the implications of the estimated models on patterns of international trade.

6 Aggregate implications

In this section, I use simulations and out-of-sample predictions to demonstrate the importance of the model regarding aggregate trade responses to shocks. The introduction of the consumer margin generates a sluggish response of trade flows, as it will take time for firms to reach new consumers. Moreover, low entry costs imply a stronger response of firms' entry and exit to shocks. As a consequence, the model can replicate, in response to a positive trade shock, a discrepancy between the short and long run trade elasticities. Moreover, I show that the model generates out-of-sample predictions that better match the extensive margin response French exporters to exchange rate movements in the Brazilian markets.

6.1 Sluggish trade response

The accumulation of consumers by firms generates frictions in growing on foreign markets. As a consequence, the trade response to shocks will be slow at the microeconomic and aggregate level. This documented pattern can explain the discrepancy between values of the trade elasticity at different horizons. International macro economists use elasticities around 1 or 2 in order to match trade responses to price variations at a high frequency. However, international trade economists use elasticities ranging from 6 to 8 to explain variations in trade flows across countries, or trade responses after a trade liberalization episode.[48]

In order to quantitatively evaluate the ability of the model to generate this discrepancy in trade elasticities between horizons, I simulate a decrease of 10 points on tariffs applied to French exports to the US. I simulate the trajectories of the firms in the estimation sample following this tariff reduction, and compare them to a counterfactual scenario without the tariff decrease. I apply this experiment to the full model, as well as the restricted model that does not feature consumer accumulation. Figure 9 reports the log-deviation relative to the counterfactual scenario without the tariff change of the total trade to the US.

As we can see from figure 9, the predictions of the two models are significantly different. In the model without consumer margin, trade increases instantaneously as the shock occurs: with lower tariffs, exporters prices decrease and trade increase, and new exporters enter the market. Because the price elasticity is larger in this model ($\sigma = 3$), we obtain a trade elasticity of roughly 2.4 immediately after the tariff change.[49] After the first year, no further adjustment occurs. In comparison, the model with consumer margin depicts a slower adjustment to trade as it takes 5 years to observe the full effect of the reduction in tariff. In the short-run, the trade response is smaller but the long-run effect is roughly 2.5 times the effect recorded after one

[48] See Ruhl (2008) or Alessandria et al. (2013) for studies of this discrepancy.

[49] Figure 19 in appendix E provides a decomposition of this response.

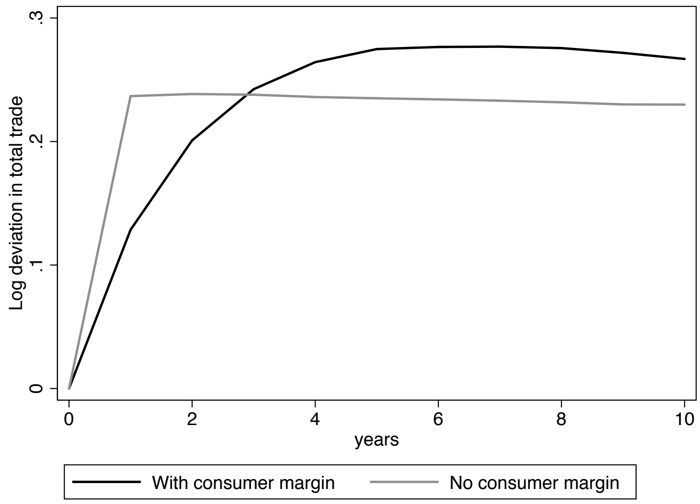

FIGURE 9: Effect of a permanent 10 points tariffs decrease.

year. Therefore, the model with consumer margin can generate this discrepancy between trade elasticities measured in the short and long run, as documented in the literature.

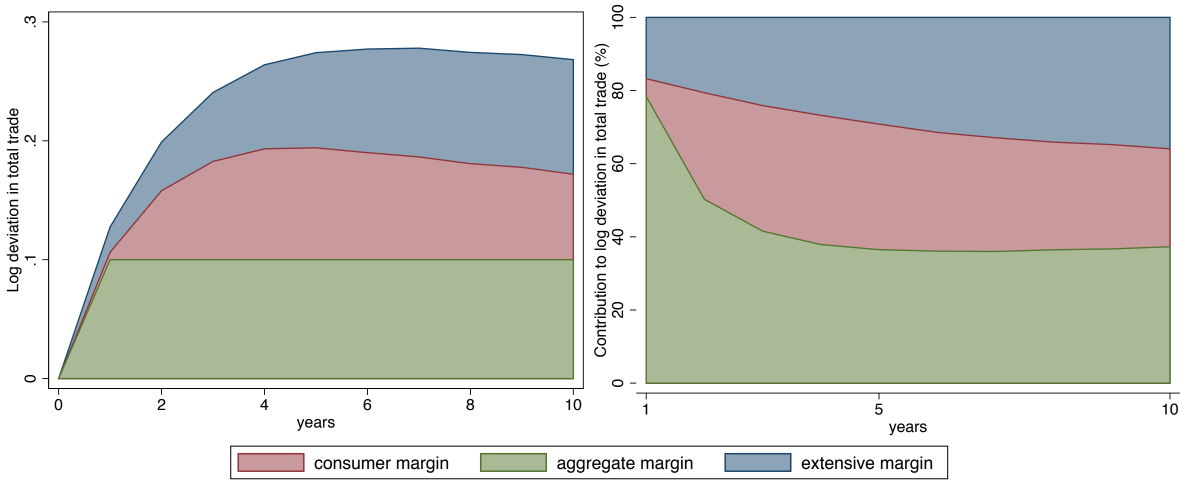

FIGURE 10: Decomposition of a permanent 10 points tariffs decrease.

In order to understand more precisely the mechanisms at play during this episode, I decompose the growth in trade between several margins. Figure 10 reports this decomposition in the left panel, and the relative contribution of each margin in the right panel.[50] This decomposition in the left figure highlights the contribution of each margin in this slow adjustment: in the first year, most of the increase in trade is due to the decreasing tariff that leads to lower prices and larger sales. However, in the following years, both the intensive and extensive margins amplifies the trade elasticity as it takes time for existing and new exporters to reach their optimal number of consumers. Moreover, due to decreasing returns of consumer accumulation, the model can

[50]The decomposition follows a methodology identical to Hummels and Klenow (2005). I provide the decomposition across all margins in figure 18 in appendix E but only reports the main margins at play in figure 10.

also explain why the extensive margin displays a increasing contribution throughout the trade expansion. This result is consistent with recent findings documented in Kehoe and Ruhl (2013) and Alessandria et al. (2013): in their empirical study, the latter manuscript finds that the contribution of the extensive margin goes from zero to around 50% during a trade expansion following a devaluation episode. By contrast, the right figure shows a relative contribution of the extensive margin of respectively 15 and 35 percent in the short and long run.

6.2 Out-of-sample predictions: export response to exchange rate variations in Brazil.

In order to further demonstrate the relevance of the model with consumer margin, I compare its out-of-sample predictions relative to the standard model. I take advantage of additional destinations, that have not been previously used in the estimation, to test the ability of the model to correctly predict the exporting behavior of the French exporters contained in my sample.

I apply this methodology to the Brazilian wine market during my sample period.[51] The choice of the Brazilian market is based on two reasons: first, it is a large market such that a large enough number of French wine producers export to Brazil. Second, the Brazilian wine market has recorded during the sample period two important shocks that affected the Brazilian demand for French wine: the devaluation of the Brazilian currency, the real, in 1999, that has been followed by a strong depreciation of the currency in the following years, and the Argentinian devaluation in 2002, which led to a strong growth in wine export to Brazil. These shocks respectively generated a strong increase in the price of French wines in local currency and an important drop of the price index on the Brazilian wine market.

Therefore, I take advantage of these variations in exchange rates, which can be arguably seen as exogenous to French exporters behavior, as sources of variation in the aggregate demand received by French firms. The model relies on five state variables that characterize the entry and sales of exporters: the appeal λ_{ft} and productivity ϕ_{ft} of the firms, their consumer shares n_{fdt}, the aggregate demand from a destination X_{dt} and their previous export activity $\mathcal{I}_{fdt-1}$. Because the quality and productivity of the firms are common across destinations, I can use the estimated individual qualities and productivities from the estimation procedure. Moreover, the variables n_{fdt} and $\mathcal{I}_{fdt-1}$ are obtained from the predictions of the model, such that only initial conditions are required for these variables. Therefore, I can construct the aggregate demand X_{dt} from Brazil, using variations in real exchange rates and the Brazilian GDP,[52] and feed this variable in the model to deliver predictions of entry, sales and prices in the Brazilian market for each firm I used in the estimation.

The results of these predictions are displayed in figure 11. This figure compares the actual data (separately from the estimation sample and the full sample of French firms) and the median

[51] My sample period goes from 1997 to 2010. However, I stop my predictions in 2007, since the trade collapse generated a strong decrease in trade that is difficult to account for.

[52] From the demand equation used in the model, $X_{dt} = \log Y_{dt} - (1-\sigma)\log P_{dt} + (1-\sigma)\log(\tau_{dt} e_{dt})$. Therefore, I use the Brazilian GDP, the BRA/FRA exchanges rates and exchange rates movements of main wine exporters to Brazil to construct variations in the price index. To obtain the level of $X_{BRA,t}$, I set $X_{BRA,t}$ such that the sales of the median prediction equals the realized sales on the market in 1998, the year before the shock. Appendix F provides more details.

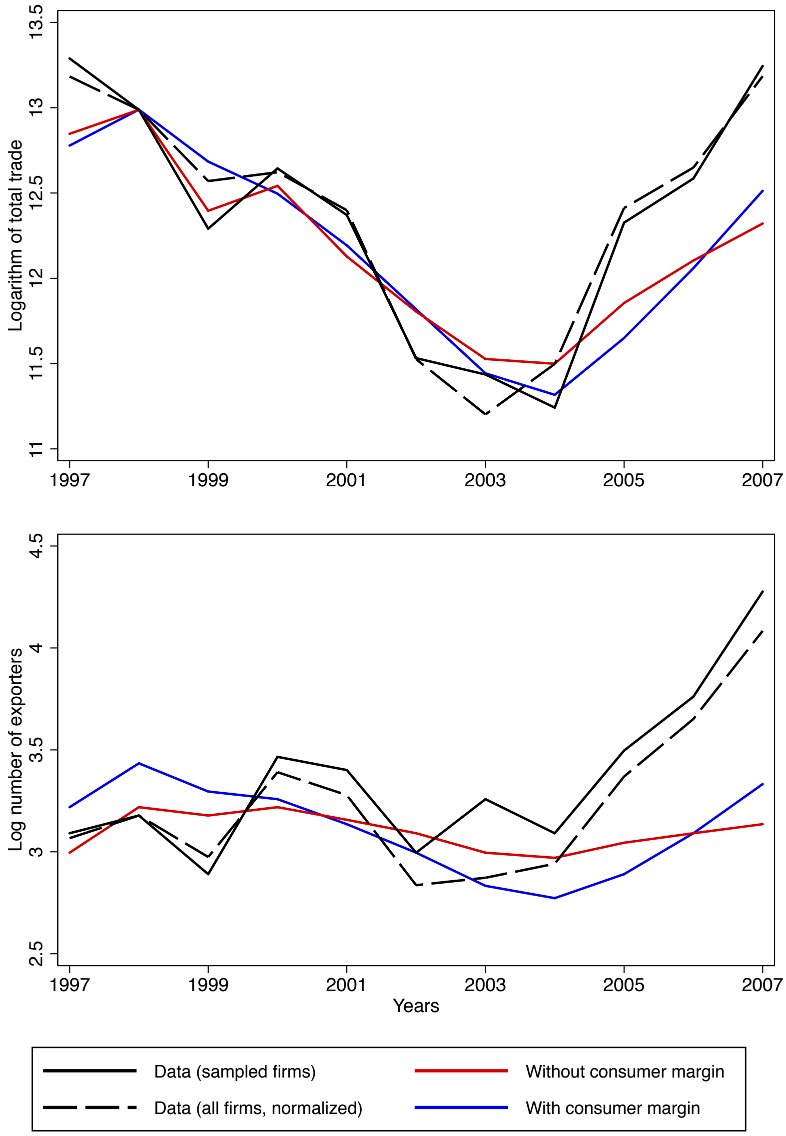

Figure 11: Total exports of wine to Brazil from selected firms

predictions from the two models of the total trade and number of exporters in the sample.[53] The top figure reports the decrease in French wine exports to Brazil that occurred between 1998 to 2003. This decrease is explained by the Brazilian devaluation in 1999, and the growth of Argentinian exports led by their devaluation in 2002. However, total exports increase after 2003 as a result of the improvement in economic conditions in Brazil at this period. Regarding the predictions of the models, we can see that both models perform relatively well at predicting the decrease in trade and the rebound after 2004. However, the model without consumer margin does not react as much to the changes in exchange rates. Specifically, the reduction in trade is not as large as what happened in reality, and the rebound is not as steep as what the data suggest: the variation in relative prices does reduce export values, but not in the same magnitude as in the data. In comparison, the model with consumer margin does marginally better at predicting the extent of the trade reduction due to the exchange rate movements: the decrease in trade is

[53] In appendix F, figure 21 also includes the confidence intervals of those predictions.

slightly more pronounced and the rebound is more consistent with the actual data.

The bottom figure highlights the reason for this difference in trade responses across models. The full model with consumer accumulation predicts a larger adjustment along the extensive margins: more firms decide to leave the Brazilian market after the Brazilian devaluation, and a larger number of firms enter as economic conditions improve after 2003. This adjustment of the extensive margin does not appear as much in the restricted model: because of large entry costs, firms will prefer to lose money temporarily when a negative shock occurs, in order to keep the option value of exporting in the next years. On the contrary, with consumer accumulation and lower entry costs, some firms decide to leave the market because they know it will be affordable to potentially reenter in the future. This stronger response in terms of entry/exit explains why the model with consumer margins can replicate the growth in trade after 2004, while the response from the restricted model is less consistent with the data.

7 Conclusion

In this paper, I develop and estimate a dynamic empirical model of trade that features state dependence in demand through the accumulation of consumers in foreign markets. Estimating the model using a set of French wine exporters, I show that accounting for this dependence is critical to understand the entry and exit decisions of firms in foreign markets, but also for the estimation of the costs of exporting: on average, estimated entry costs are half of those estimated in the standard model without consumer accumulation. Moreover, I demonstrate using simulations and out-of-sample predictions that this consumer margin, and the associated fall in entry costs, matters for aggregate predictions: the model can generate a slow response of aggregate trade to shocks and can correctly replicate the contribution of the extensive margin throughout a trade liberalization episode.

These results shed new light on the nature of the barriers to trade at the firm level. While existing models emphasize the role of large sunk entry costs as the main barrier to trade to explain the persistence in export markets, this paper shows that dependence in demand is responsible for a significant share of this persistence. In fact, the ability to reach a large and stable demand for a product appears to be one of the primary sources of success for firms in foreign markets. Therefore, this study improves our understanding of the determinants of trade dynamics at the microeconomic and aggregate levels, which has important implications for countries aiming to improve the export performance of their industries.

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.

Akhmetova, Z. and C. Mitaritonna (2012): "A Model of Firm Experimentation under Demand Uncertainty with an Application to Multi-Destination Exporters," *University of New South Wales Working Paper.*

Albornoz, F., H. F. C. Pardo, G. Corcos, and E. Ornelas (2012): "Sequential exporting," *Journal of International Economics*, 88, 17–31.

Alessandria, G. and H. Choi (2007): "Do Sunk Costs of Exporting Matter for Net Export Dynamics?" *The Quarterly Journal of Economics*, 122, 289–336.

——— (2014): "Establishment heterogeneity, exporter dynamics, and the effects of trade liberalization," *Journal of International Economics*, 94, 207–223.

Alessandria, G., H. Choi, and K. Ruhl (2014): "Trade Adjustment Dynamics and the Welfare Gains from Trade," Working Paper 20663, National Bureau of Economic Research.

Alessandria, G., S. Pratap, and V. Z. Yue (2013): "Export dynamics in large devaluations," *Manuscript.*

Andrieu, C., A. Doucet, and R. Holenstein (2010): "Particle markov chain monte carlo methods," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72, 269–342.

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumers Margin in International Trade," *Journal of political economy*, 118, 1151–1199.

——— (2015): "A Unified Theory of Firm Selection and Growth," *The Quarterly Journal of Economics.*

Aw, B. Y., M. J. Roberts, and D. Yi Xu (2011): "R&D Investment, Exporting, and Productivity Dynamics," *The American Economic Review*, 101, 1312–1344.

Berman, N., V. Rebeyrol, and V. Vicard (Forth.): "Demand learning and firm dynamics: evidence from exporters," *Review of Economics and Statistics.*

Bernard, A. B., E. A. Boler, R. Massari, J.-D. Reyes, and D. Taglioni (2017): "Exporter dynamics and partial-year effects," *American Economic Review*, 107, 3211–28.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in International Trade," *The Journal of Economic Perspectives*, 105–130.

Berthou, A. and V. Vicard (2015): "Firms' Export Dynamics: Experience Versus Size," *The World Economy*, 38, 1130–1158.

Blevins, J. R. (2015): "Sequential Monte Carlo Methods for Estimating Dynamic Microeconomic Models," *Journal of Applied Econometrics*.

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

Dickstein, M. J. and E. Morales (2015): "What do exporters know?" Tech. rep., National Bureau of Economic Research.

Drozd, L. A. and J. B. Nosal (2012): "Understanding international prices: Customers as capital," *The American Economic Review*, 102, 364–395.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010): "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 41, 417–445.

Eaton, J., M. Eslava, D. Jinkins, C. Krizan, M. Kugler, and J. Tybout (2014): "A Search and Learning Model of Export Dynamics," *Manuscript*.

Fitzgerald, D., S. Haller, and Y. Yedid-Levi (2016): "How exporters grow," Tech. rep., National Bureau of Economic Research.

Flury, T. and N. Shephard (2011): "Bayesian inference based only on simulated likelihood: particle filter analysis of dynamic economic models," *Econometric Theory*, 27, 933–956.

Foster, L., J. Haltiwanger, and C. Syverson (2016): "The slow growth of new plants: Learning about demand?" *Economica*, 83, 91–129.

Gourio, F. and L. Rudanko (2014): "Customer Capital," *Review of Economic Studies*, 81, 1102–1136.

Heckman, J. J. (1981): "Heterogeneity and State Dependence," *NBER Chapters*, 91–140.

Hottman, C. J., S. J. Redding, and D. E. Weinstein (2016): "Quantifying the sources of firm heterogeneity," *The Quarterly Journal of Economics*, 131, 1291–1364.

Hotz, J. and R. Miller (1993): "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 60, 497–529.

Hummels, D. L. and P. Klenow (2005): "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95, 704–723.

Imai, S., N. Jain, and A. Ching (2009): "Bayesian estimation of dynamic discrete choice models," *Econometrica*, 77, 1865–1899.

Kehoe, T. J. and K. J. Ruhl (2013): "How important is the new goods margin in international trade?" *Journal of Political Economy*, 121, 358–392.

Khandelwal, A. (2010): "The long and short (of) quality ladders," *The Review of Economic Studies*, 77, 1450–1476.

Li, S. (2018): "A structural model of productivity, uncertain demand, and export dynamics," *Journal of International Economics.*

Lincoln, W. F. and A. H. McCallum (2018): "The rise of exporting by US firms," *European Economic Review*, 102, 280–297.

Lindsten, F., M. I. Jordan, and T. B. Schön (2014): "Particle Gibbs with ancestor sampling," *The Journal of Machine Learning Research*, 15, 2145–2184.

Macchiavello, R. (2010): "Development uncorked: Reputation acquisition in the new market for Chilean wines in the UK," *Manuscript.*

McCallum, A. H. (2015): "The Structure of Export Entry Costs," *Manuscript.*

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

Morales, E., G. Sheu, and A. Zahler (2015): "Extended gravity," *NBER Working Paper*, 21351.

Nguyen, D. X. (2012): "Demand uncertainty: Exporting delays and exporting failures," *Journal of International Economics*, 86, 336–344.

Norets, A. (2009): "Inference in dynamic discrete choice models with serially orrelated unobserved state variables," *Econometrica*, 77, 1665–1682.

Osborne, M. (2011): "Consumer learning, switching costs, and heterogeneity: A structural examination," *Quantitative Marketing and Economics*, 9, 25–70.

Pierce, J. R. and P. K. Schott (2012): "Concording US Harmonized System Codes over Time," *Journal of Official Statistics*, 28, 53–68.

Rauch, J. E. and J. Watson (2003): "Starting small in an unfamiliar environment," *International Journal of Industrial Organization*, 21, 1021–1042.

Roberts, M. J., D. Y. Xu, X. Fan, and S. Zhang (2012): "A Structural Model of Demand, Cost, and Export Market Selection for Chinese Footwear Producers," Working Paper 17725, National Bureau of Economic Research.

Rodrigue, J. and Y. Tan (2015): "Price and Quality Dynamics in Export Markets," *Manuscript.*

Ruhl, K. J. (2008): "The international elasticity puzzle," *Manuscript.*

Ruhl, K. J. and J. L. Willis (2017): "New exporter dynamics," *International Economic Review*, 58, 703–726.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

Shocker, A. D., M. Ben-Akiva, B. Boccara, and P. Nedungadi (1991): “Consideration set influences on consumer decision-making and choice: Issues, models, and suggestions,” *Marketing letters*, 2, 181–197.

Timoshenko, O., D. Dias, and P. Bastos (2017): “Learning, Prices, and Firm Dynamics,” Tech. rep., Society for Economic Dynamics.

Timoshenko, O. A. (2015): “Learning versus sunk costs explanations of export persistence,” *European Economic Review*, 79, 113–128.

Van Beveren, I., A. B. Bernard, and H. Vandenbussche (2012): “Concording EU trade and production data over time,” Working Paper 18604, National Bureau of Economic Research.

Appendices
An empirical dynamic model of trade with consumer accumulation

A Constructions of the samples

The dataset used in the paper is initially disaggregated at the monthly level. From this raw dataset, a number of steps are implemented to improve the reliability and consistency of the data. First, I describe the operations implemented for the first empirical exercise, that uses a wide set of products. Then, I describe the procedures implemented to obtain the final sample used in the structural estimation.

A.1 Data appendix for the reduced-form exercise

I implement two important steps to prepare the data for the regressions displayed in the reduced-form exercise. First, I clean outliers and product categories that do not provide a meaningful and consistent unit of count across years. Second, I correct for the partial-year bias.

Cleaning and harmonization I make three different operations to clean the dataset from potential outliers or measurement errors.

- First, I use the algorithm from Pierce and Schott (2012) and Van Beveren, Bernard, and Vandenbussche (2012) to account for changes in product categories at the eight digit level. This algorithm allows me to obtain categories that are consistent across the sample years (1996-2010).
- Second, I drop product categories that meet one of the following criteria:
  - the counting unit is changing across years.
  - the counting unit is not identical within the category (because of the previous step, the current product category can contain eight digit categories with different units).
- Finally, because unit values, constructed as export values divided by quantities, are a source of measurement errors, I winsorize them at the eight-digit product category$\times$country$\times$year level. Specifically, I set at the values of the 5th and 95th percentiles the prices that are beyond these two thresholds.

Correction for partial-year bias As described in Berthou and Vicard (2015) and Bernard, Boler, Massari, Reyes, and Taglioni (2017), a firm will sell less in average during its first calendar year as exporter. This is because calendar years do not necessarily match the beginning of the exporting activity. In order to correct for this potential bias, I reconstruct the dataset to align calendar exporting years of each exporter. The idea is to define a new year for each spell of export, setting the first month of this year as representative of a regular year, and constructing exporting spells based on this new starting month.

Specifically, the following procedure is applied to each firm-destination-product triplet: for the earliest observation in 1996, if no observation is seen in 1995, a new spell is defined: the month of this first flow is probabilistically drawn based on the number of flows observed during the following 12 months. Then, the year is set to 1996 or 1997 depending on whether the initial month is earlier or later than July. The following observations are adjusted accordingly to preserve the duration between monthly export flows, as long as there is no discontinuity in the exporting activity according to the newly defined calendar years. In case of discontinuity, the next observation becomes a new reference point, and the same procedure is applied for this observation and the following ones.

Once this adjustment is implemented, I aggregate the data at the yearly-level. Specifically, I sum values exported within each newly created calendar year at the firm-product-category level. Moreover, I obtain yearly prices using an export-weighted average of monthly prices. In case of missing prices, I assume a weight of zero for this observation.[54] If this observation is the only observation within a firm-destination-product- year combination, I drop all the observations within the firm-destination-product triplet.

This procedure leaves me with sales and prices measured at the firm-product-destination-year level, with no missing observation in prices, and adjusted for the existence of partial-year of exporting.

A.2 Data appendix for the structural estimation

The procedure to clean the data for the structural estimation is different than the reduced-form exercise. I describe in this subsection the choice of the wine industry and the set of destinations I use for implementing my estimation. Then, I describe the cleaning procedure implemented on the wine producers and provide summary statistics on the final sample of firms used in the estimation.

A.2.1 Wine industry

The decision to implement this estimation on wine exporters relies on two constraints. First of all, I study the entry decision made at the firm level. This level of analysis is explained by the fact that brands and reputation are often defined by the firm that produces the good. Therefore, this requires to study firms that display a small level of heterogeneity in terms of goods. A car producer for instance, that also exports car pieces, or engines for other vehicles, is difficult to analyze as a single-product firm. However, a wine producer mostly export wines, and specifically bottles of wine, whose prices are easy to define, and aggregate at the firm level. For these reasons when defining my sample, I exclusively use wine producers that do not export any other goods outside of wine. A large share of the trade in wine is made by wholesalers who export other types of items, and for which the study at the level of the firm is irrelevant. In addition to this homogeneity constraint, my estimation procedure requires enough firms which export to several destinations. As a major exporting industry from France, the wine industry meets both of these conditions: a large number of exporters, exporting a precisely defined good.

[54] Since quantities are sometimes missing, I compute an average price rather than computing the price from the ratio of yearly values and quantities.

In addition to imposing restrictions on the set of firms included in the final sample, I only use a restricted set of destinations.

A.2.2 Selection of destinations

I select 15 different destinations on which I analyze the behaviors of French exporters. These destinations have been selected among the 20 most popular destinations for wine exports from France, excluding countries with large import/export platforms such as Denmark and Singapore, while reflecting some heterogeneity in terms of location. Moreover, I divide these destinations in three groups, for which I estimate different entry and fixed costs of exporting.. The list of these destinations can be found in table 4.

Table 4: List of destination countries included in the structural sample

	Group 1 *Europe*		**Group 2** *Americas*	**Group 3** *Asia/Oceania*
Great-Britain	Germany	Belgium	(Brazil)	Australia
Netherlands	Italy	Spain	Canada	China
Ireland	Sweden	Switzerland	United States	Japan

Note that I do not include Brazil in the estimation sample. The observations related to this destination are used in the out-of-sample exercise and are excluded so that it does not affect the estimation procedure.

A.2.3 Aggregation

Because the estimation is conducted at the firm-destination-year level, it is necessary to aggregate the sales and quantities exported across products exported by the firm. The choice of the wine industry is crucial here since bottles of wines are quantities that can be easily aggregated. An industry producing differentiated goods would have made this aggregation less straightforward. The aggregation of prices and sales are the following:

$$p_{fdt} = \sum_{h=1}^{H_{fdt}} w_{fhdt} \frac{s_{fhdt}}{q_{fhdt}} \qquad \text{with} \quad w_{fhdt} \equiv \frac{s_{fhdt}}{\sum_h s_{fhdt}}$$

$$s_{fdt} = \sum_{h=1}^{H_{fdt}} s_{fhdt}$$

where H_{fdt} is the number of 8-digit observations for each firm-destination-year triplet. Moreover, note that there is a certain number of missing quantities in the data. Therefore, I assign a weight w_{fhdt} equal to zero to the observations that have quantities or values exported equal to one or zero. When this observation is the only one at the firm-destination-year level (no other product is sent to this market by this firm this year), I dropped all the observations related to this firm from the sample.

A.2.4 Partial-year bias

Similar to the sample used in the reduced form exercise, I will correct for the partial-year bias, by redefining the entry months of all entering exporters. As a consequence, I shift all the subsequent flows to maintain the same sequence in the exports of the firm. Therefore, exports during the first year will look similar to the subsequent years of exporting.

A.2.5 Cleaning

I clean the data to avoid the potential existence of outliers in prices. In particular, I exclude observations that display extreme prices along two dimensions. First, I flag observations which log difference is larger than two, or lower than -2 relative to the previous year. Second, I flag prices which log value is larger than two or lower than two relative to the average price of the firm-year pair. After having flagged those observations, I dropped all observations of a firm that contains at least one flagged observation.

Finally, a last criterion for a firm to be included in the final sample is based on the number of observations. Many firms export one year to one market during the sample period, and this does not provide enough information to analyze their exporting behavior. Therefore, I only keep firms that recorded at least 10 exporting events, and have exporting activity in at least two destinations. Note that with 14 destinations and 14 years of data, the maximum number of observations by a given firm is 196. This selection process could present a problem as it is likely to affect the estimates of entry and fixed costs of exporting, by only looking at successful firms. However, this procedure will tend to select firms that survive several years, rather than short-lived exporters: as a consequence, it tends to go against the theory of consumer accumulation that can accommodate small and short-lived exporters relative to the standard model. Finally, I only keep firms that have some exporting activity to Brazil during the sample period.

A.2.6 Final sample

Once these cleaning steps were implemented, I obtain a sample of 233 firms. The following tables present summary statistics regarding this sample.

TABLE 5: Description of the sample used in the structural estimation

Statistics:	*pc5*	*median*	*pc95*	*mean*	N
# observations per firm	10	28	107	40.2	233
av. # destinations per firm-year	1.27	3.13	8.17	3.85	2 339
av. # years per firm-destination	2.2	4.75	8.92	4.94	1 783

Table 5 provides information regarding the number of observations provided by the sampled firms, as well as the number of destinations they export to in an average year. One can see that the firms selected are relatively large, with a minimum number of export episodes equal to 10 by the sampling procedure. However, the median firm only records 28 export episodes, while the maximum number of episodes in the dataset is 196 (14 × 14). Moreover, they are relatively diversified in terms of destinations since the median firm exports to 3.13 destinations in an average year.

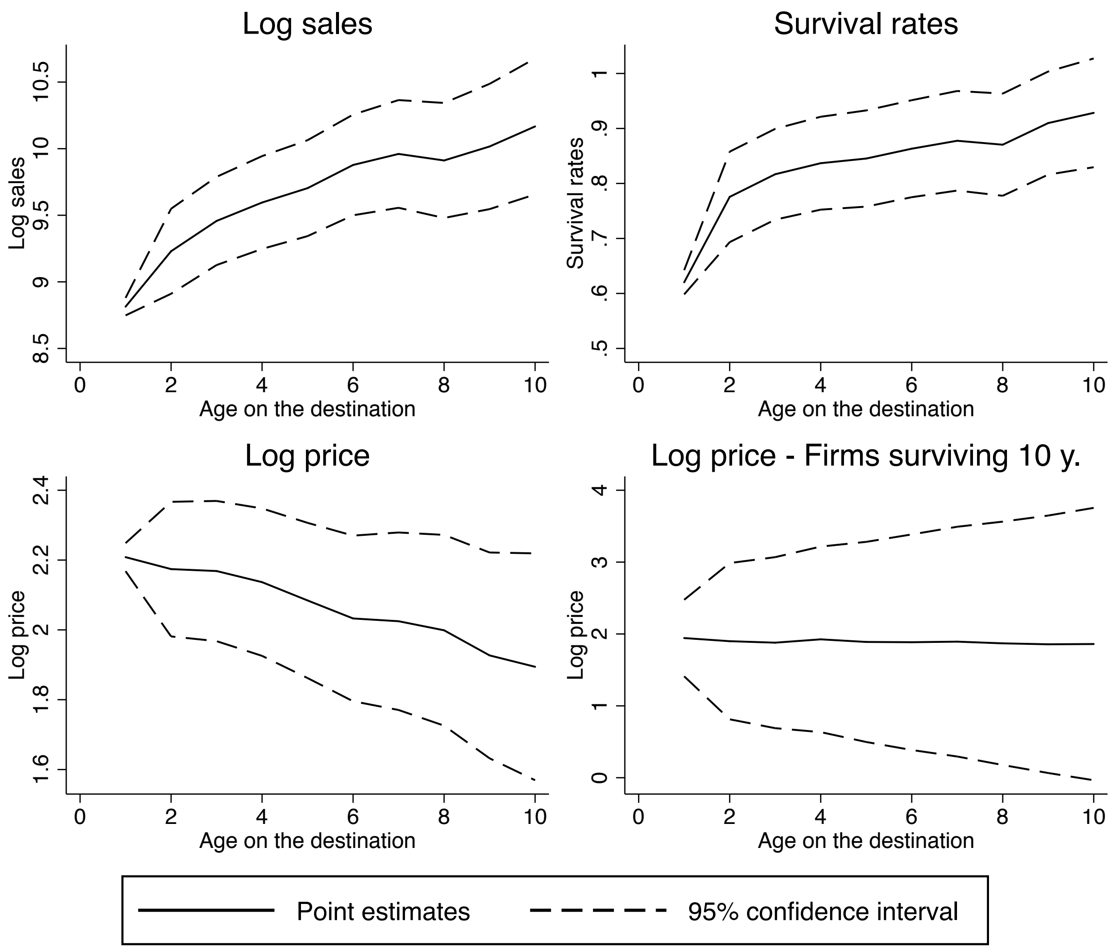

FIGURE 12: Sales, prices and survival rates across ages (Wine producers)

Notes: The figure reports the average log sales, log prices and survival rates of wine producers in a destination at different ages. The estimates are obtained from the regression of these dependent variables on a set of age dummies and destination×year fixed effects. The age in a destination is defined as the number of years a firm has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors estimates clustered at the firm-destination level.

In order to inspect how this sampling procedure affects the trajectories of the exporters, I replicate the regressions on age dummies I perform in section 2. Figure 12 reports the results of these regressions for sales, prices and survival rates.[55] The patterns of sales and prices are very similar to the ones observed using the comprehensive sample: sales appear to increase in the early years, with the an average growth rate of 30 percent the first year. However, we can see that the survival rates in the structural sample are larger than the ones displayed in the exhaustive

[55]Table 6 provides the tables related to these regressions.

data. While the survival rate was close to 35 percent in the full sample, it is around 60 percent in this restricted sample. This arises because of the requirement made during the selection of exporters: because the estimation procedure requires firms with several observations, this tends to eliminate firms with very large attrition rates that do not records many episodes of exporting activity. Note that this difference in survival rates between exhaustive and restricted samples will play against the story I develop in this paper. Large attrition rates will be consistent with a story that emphasizes strong dependence in demand rather than an important role for sunk costs of entry. Finally, prices tend to decrease with ages when we do not control for selection. However, they tend to be flat when we restrict the sample to firms surviving 10 years.

Table 6: Age regressions using the structural sample

	No fixed effects			Year x destination fixed effects		
	(1) Log sales	(2) Log prices	(3) Surv. rates	(4) Log sales	(5) Log prices	(6) Surv. rates
age 2	0.46***	-0.026	0.16***	0.42***	-0.034*	0.16***
	(0.04)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)
age 3	0.71***	-0.016	0.20***	0.64***	-0.040	0.20***
	(0.04)	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)
age 4	0.83***	-0.022	0.22***	0.78***	-0.072*	0.22***
	(0.05)	(0.03)	(0.02)	(0.05)	(0.03)	(0.02)
age 5	0.95***	-0.053	0.24***	0.89***	-0.12***	0.22***
	(0.06)	(0.03)	(0.02)	(0.06)	(0.03)	(0.02)
age 6	1.11***	-0.11**	0.26***	1.06***	-0.18***	0.24***
	(0.07)	(0.04)	(0.02)	(0.07)	(0.04)	(0.02)
age 7	1.18***	-0.12**	0.27***	1.15***	-0.18***	0.26***
	(0.07)	(0.04)	(0.02)	(0.08)	(0.05)	(0.02)
age 8	1.12***	-0.16***	0.26***	1.10***	-0.21***	0.25***
	(0.09)	(0.05)	(0.02)	(0.10)	(0.06)	(0.03)
age 9	1.24***	-0.21***	0.31***	1.20***	-0.28***	0.29***
	(0.1)	(0.05)	(0.02)	(0.1)	(0.06)	(0.03)
age 10	1.39***	-0.24***	0.32***	1.35***	-0.31***	0.31***
	(0.1)	(0.06)	(0.03)	(0.1)	(0.08)	(0.03)
N	7631	7631	6828	7631	7631	6828
R^2	0.093	0.0060	0.065	0.21	0.14	0.13

Notes: Firm x destination clustered standard errors between parentheses. * p<0.05, ** p<0.01, *** p<0.001

B Additional age regressions

In this section, I provide additional results and describe alternative specifications to look at the correlation between sales or prices and age in an export market.

B.1 Additional results

In table 13, I present results looking at variations in sales with export experience, when including firm-product-year fixed effects. In this specification, sales dynamics are identified across destinations and the results confirm that exported values strongly increase with age in foreign markets.

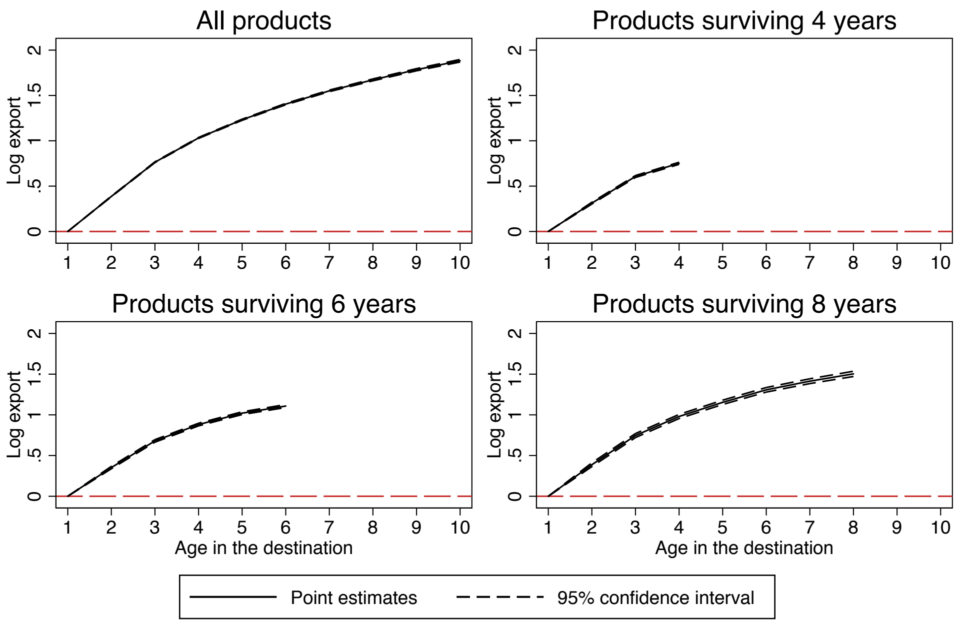

FIGURE 13: Sales across export ages, across destinations

Notes: The figure reports the cumulative growth of sales compared to age one, of a firm-product category pair in a destination at different ages. The regression uses logarithm of sales as dependent variable, and includes product category×destination×year and firm×product category×year fixed effects. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product level.

B.2 Additional specification

Both specifications shown in the paper use cross sectional variations (either across firms within markets, or within firm across destinations) to identify sales and prices dynamics. However, an intuitive way to capture these dynamics would be to identify trends within exporting spells. Therefore, a natural specification would include firm-destination-product-spell fixed effects to capture the heterogeneity across firms, products and markets. However, including this set of fixed effects makes it impossible to identify a trend in prices or sales across ages: it can capture deviations from the trend, but not the trend itself. To understand why, consider a sample of firms on a given market *pdt*. Because of the market-level fixed effect, their average price is normalized to zero. Now consider the same set of firms one year later, when all firms aged by one year. The market-level fixed effect means that their average price will be normalized to zero as well. Therefore, a trend in prices cannot be identified. Intuitively, the fact that all firms are getting one year older every year implies the absence of control groups. Therefore, this specification does not allow to test for a trend in sales or prices, but can identify for which ages the growth is more pronounced.

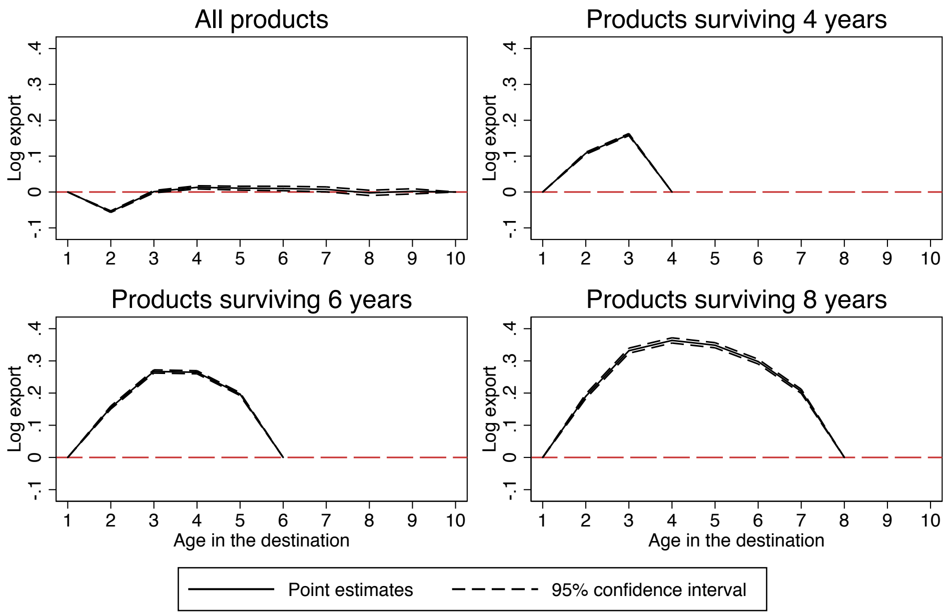

FIGURE 14: Sales across export ages, within variation

Notes: The figure reports the cumulative growth of sales compared to age one, of a firm-product category pair in a destination at different ages. The regression uses logarithm of sales as dependent variable, and includes product category×destination×year and firm×product category×destination×spell fixed effects. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product-destination-spell level.

Figures 14 and 15 report the results of this specification for sales and prices. As we can see, even sales are not increasing with age in this specification: in fact, the coefficients are negative for young firms, which means that sales increase less between years 1 and 2. Similar patterns appear for prices. However, regressions including a constant set of firms are more informative. We can see that both prices and sales tend to increase faster in the first years of exporting, which is consistent with the findings from the two main specifications used in the main text.

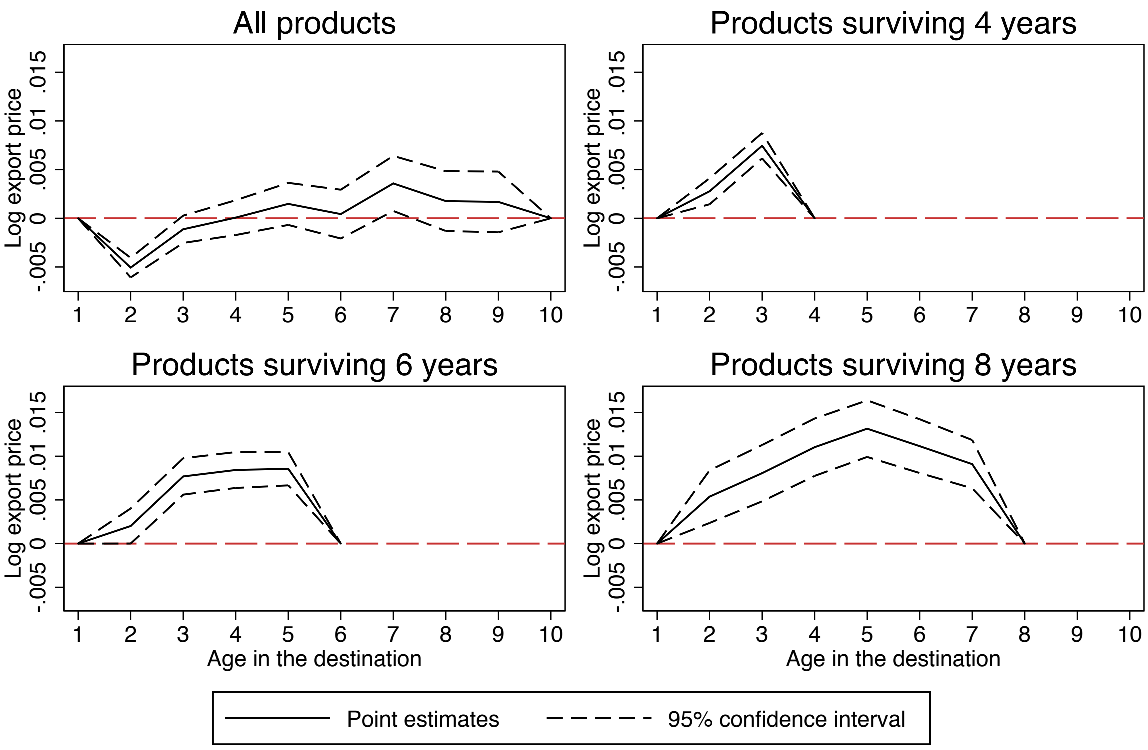

Figure 15: Prices across export ages, within variation

Notes: The figure reports the cumulative growth of prices compared to age one, of a firm-product category pair in a destination at different ages. The regression uses logarithm of price as dependent variable, and includes product category×destination×year and firm×product category×destination×spell fixed effects. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product-destination-spell level.

B.3 Tables of results

Table 7: Age regressions (full sample)

	(1) Survival	(2) Log sales	(3) Log price	(4) Log sales	(5) Log price	(6) Log sales	(7) Log price
age 2	0.24***	0.59***	0.022***	-0.055***	-0.0051***	0.39***	0.012***
	(0.0003)	(0.001)	(0.0006)	(0.001)	(0.0005)	(0.002)	(0.0008)
age 3	0.33***	1.02***	0.038***	0.00090	-0.0011	0.76***	0.025***
	(0.0004)	(0.002)	(0.0009)	(0.002)	(0.0007)	(0.002)	(0.0010)
age 4	0.38***	1.32***	0.048***	0.013***	0.000069	1.03***	0.034***
	(0.0005)	(0.002)	(0.001)	(0.002)	(0.0009)	(0.003)	(0.001)
age 5	0.40***	1.54***	0.057***	0.011***	0.0015	1.23***	0.039***
	(0.0006)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)
age 6	0.42***	1.74***	0.059***	0.0098**	0.00043	1.40***	0.042***
	(0.0007)	(0.003)	(0.001)	(0.003)	(0.001)	(0.004)	(0.001)
age 7	0.44***	1.92***	0.068***	0.0073*	0.0036*	1.55***	0.044***
	(0.0008)	(0.004)	(0.002)	(0.004)	(0.001)	(0.004)	(0.002)
age 8	0.45***	2.08***	0.072***	-0.0025	0.0018	1.67***	0.041***
	(0.0009)	(0.005)	(0.002)	(0.004)	(0.002)	(0.005)	(0.002)
age 9	0.46***	2.22***	0.072***	0.0018	0.0017	1.78***	0.043***
	(0.001)	(0.006)	(0.003)	(0.004)	(0.002)	(0.006)	(0.002)
age 10	0.47***	2.33***	0.074***	.	.	1.88***	0.041***
	(0.001)	(0.007)	(0.003)			(0.007)	(0.003)
N	17 503 978	19 311 381	19 311 381	10 780 577	10 780 577	12 871 145	12 871 145
R^2	0.29	0.37	0.78	0.86	0.96	0.74	0.92
Spell FE	N	N	N	Y	Y	N	N
Firm-Prod-Year FE	N	N	N	N	N	Y	Y

Notes: Firm x product x destination clustered standard errors between parentheses. Year x product x destinations fixed effects are included in all regressions. * p<0.05, ** p<0.01, *** p<0.001

Table 8: Age regressions (sample of surviving firms)

	(1) Log sales	(2) Log price	(3) Log sales	(4) Log price	(5) Log sales	(6) Log price
Products surviving 4 years						
age 2	0.21***	0.0024**	0.11***	0.0028***	0.31***	0.0026
	(0.002)	(0.0009)	(0.002)	(0.0007)	(0.005)	(0.002)
age 3	0.36***	0.0067***	0.16***	0.0075***	0.61***	0.012***
	(0.003)	(0.001)	(0.001)	(0.0007)	(0.005)	(0.002)
age 4	0.29***	-0.00057	.	.	0.75***	0.014***
	(0.004)	(0.002)			(0.006)	(0.002)
N	3 886 412	3 886 412	3 851 329	3 851 329	2 497 364	2 497 364
R^2	0.37	0.81	0.88	0.96	0.79	0.95
Products surviving 6 years						
age 2	0.27***	0.0020	0.16***	0.0020*	0.35***	0.0011
	(0.003)	(0.001)	(0.002)	(0.001)	(0.007)	(0.003)
age 3	0.48***	0.0071***	0.27***	0.0077***	0.68***	0.0070*
	(0.004)	(0.002)	(0.003)	(0.001)	(0.008)	(0.003)
age 4	0.59***	0.0072***	0.26***	0.0084***	0.88***	0.013***
	(0.005)	(0.002)	(0.002)	(0.001)	(0.008)	(0.003)
age 5	0.63***	0.0061*	0.20***	0.0086***	1.02***	0.012***
	(0.006)	(0.003)	(0.002)	(0.0010)	(0.009)	(0.003)
age 6	0.55***	-0.0043	.	.	1.11***	0.013***
	(0.007)	(0.003)			(0.009)	(0.003)
N	2 686 575	2 686 575	2 678 763	2 678 763	1 763 426	1 763 426
R^2	0.37	0.81	0.86	0.96	0.80	0.95
Products surviving 8 years						
age 2	0.28***	0.0048*	0.19***	0.0054***	0.39***	0.0083
	(0.005)	(0.002)	(0.004)	(0.002)	(0.01)	(0.004)
age 3	0.51***	0.0075**	0.33***	0.0081***	0.75***	0.0099*
	(0.007)	(0.003)	(0.004)	(0.002)	(0.01)	(0.005)
age 4	0.63***	0.0093**	0.36***	0.011***	0.98***	0.018***
	(0.009)	(0.004)	(0.004)	(0.002)	(0.01)	(0.005)
age 5	0.70***	0.0100*	0.35***	0.013***	1.15***	0.016**
	(0.01)	(0.004)	(0.004)	(0.002)	(0.01)	(0.005)
age 6	0.73***	0.0058	0.30***	0.011***	1.31***	0.017**
	(0.01)	(0.005)	(0.004)	(0.002)	(0.01)	(0.005)
age 7	0.72***	0.00020	0.21***	0.0091***	1.41***	0.016**
	(0.02)	(0.006)	(0.003)	(0.001)	(0.02)	(0.006)
age 8	0.60***	-0.014	.	.	1.50***	0.016**
	(0.02)	(0.007)			(0.02)	(0.006)
N	1 690 783	1 690 783	1 690 783	1 690 783	1 073 287	1 073 287
R^2	0.38	0.81	0.86	0.96	0.82	0.95
Spell FE	N	N	Y	Y	N	N
Firm-Prod-Year FE	N	N	N	N	Y	Y

Notes: Firm x product x destination clustered standard errors between parentheses. Year x product x destinations and firm x products fixed effects are included in all regressions. * p<0.05, ** p<0.01, *** p<0.001

C Derivations

The firm chooses the optimal mark-up μ to maximize the value of exporting:

$$\begin{aligned}\mu &= \operatorname{argmax} V_I(\xi, n, \mu)\\ &= \operatorname{argmax} E_\varepsilon \Big\{ \pi(\xi, n, \mu, \varepsilon) + \beta EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1) \Big\}\\ &= \operatorname{argmax} \int_\varepsilon \pi(\xi, n, \mu, \varepsilon) + \beta EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1) dF(\varepsilon)\end{aligned}$$

Therefore, the first order condition of the problem is

$$\int_\varepsilon \frac{\partial \pi(\xi, n, \mu, \varepsilon)}{\partial \mu} + \beta \frac{\partial EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1)}{\partial \mu} dF(\varepsilon) = 0$$

First, profit function is

$$\begin{aligned}\pi(\xi, n, \mu, \varepsilon) &= n \exp(\lambda + X + \varepsilon^D) \mu^{-\sigma}(\mu - 1) c(\xi, n, \varepsilon^S)^{1-\sigma}\\ \Rightarrow \frac{\partial \pi(\xi, n, \mu, \varepsilon)}{\partial \mu} &= \left[(1-\sigma)\mu^{-\sigma} + \sigma \mu^{-\sigma-1}\right] n \exp(\lambda + X + \varepsilon^D) c(\xi, n, \varepsilon^S)^{1-\sigma}\end{aligned}$$

Second, the continuation value can be rewritten $EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1) = EV'(\xi, n'(s, n), 1)$ where s are the sales of the firm. Therefore,

$$\begin{aligned}\frac{\partial EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1)}{\partial \mu} &= \frac{\partial EV'(\xi, n'(s, n), 1)}{\partial \mu}\\ &= \frac{\partial s}{\partial \mu} \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\\ &= (1-\sigma)\mu^{-\sigma} n \exp(\lambda + X + \varepsilon^D) c(\xi, n, \varepsilon^S)^{1-\sigma} \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\end{aligned}$$

Therefore, the first order condition can be rewritten

$$\begin{aligned}&\int_\varepsilon n \exp(\lambda + X + \varepsilon^D) c(\xi, n, \varepsilon^S)^{1-\sigma} \Big[(1-\sigma)\mu^{-\sigma} + \sigma\mu^{-\sigma-1}\\ &\qquad\qquad + (1-\sigma)\mu^{-\sigma} \beta \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\Big] dF(\varepsilon) = 0\\ \Leftrightarrow\quad &\int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) \Big[\sigma + (1-\sigma)\mu\Big(1 + \beta \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\Big)\Big] dF(\varepsilon) = 0\\ \Leftrightarrow\quad &\mu(\sigma - 1) \int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) \Big(1 + \beta \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\Big) dF(\varepsilon)\\ &\qquad\qquad = \sigma \int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) dF(\varepsilon)\\ \Leftrightarrow\quad &\mu = \frac{\sigma}{\sigma - 1} \frac{\int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) dF(\varepsilon)}{\int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S)\Big(1 + \beta \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\Big) dF(\varepsilon)}\\ \Leftrightarrow\quad &\mu = \frac{\sigma}{\sigma - 1} \frac{1}{\int_\varepsilon \omega(\varepsilon)\Big(1 + \beta \frac{\partial n'}{\partial s} \frac{\partial EV'(\xi, n', 1)}{\partial n'}\Big) dF(\varepsilon)}\end{aligned}$$

with $\omega(\varepsilon) = \frac{\exp(\varepsilon^D + (1-\sigma)\varepsilon^S)}{\int_\varepsilon \exp(\varepsilon^D + (1-\sigma)\varepsilon^S) dF(\varepsilon)}$.

D Details of the algorithm

I describe in this section of the appendix the MCMC algorithm I implement. I start by describing how the Markov chain is initialized, before describing a given iteration of the chain, involving the update of the unobservables and parameters.

D.1 Initial values

I start by describing how the unobservables are obtained, before describing the initial parameters. From the value of the price elasticity of demand, I can obtain $\log s_{fdt}+\sigma p_{fdt} = \log n_{fdt}+X_{dt}+\lambda_{ft}$. I can then decompose this term using firm-year and destination-year fixed effect to obtain $\lambda_{ft}^{(0)}$ and $X_{dt}^{(0)}$. In order to obtain $\phi_{ft}^{(0)}$, I run the regression $\log p_{fdt} - \frac{\sigma}{\sigma-1}$ on $\lambda_{ft}^{(0)}$. This allows me to obtain $\alpha^{(0)}$, and the residual is regressed on firm-year fixed effects to obtain $\phi_{ft}^{(0)}$. Having in hand initial values for the unobservables, I can use linear regressions to obtain the AR(1) coefficients for the unobservables, and use nonlinear least square to estimate $\underline{n}^{(0)}$, $n_0^{(0)}$, $\eta_1^{(0)}$ and $\eta_2^{(0)}$ after arbitrarily setting $\psi^{(0)} = 0.5$. Finally, I set values for the fixed costs parameters, and the variance parameter of the fixed cost shocks. I arbitrary set $f^{(0)} = s_v^{(0)} = 10\,000$ and $f^{(0)} = 30\,000$ for the three different groups of countries.

After setting these initial values, I implement 500 iterations that does not account for the dynamic problem of the firm. Therefore, I sample unobservables and parameters assuming a constant mark-up and only taking advantage of the realized sales and prices. This step allows me to obtain initial conditions for the parameters and unobservables that are closer to their true values, although biased because they do not account for the dynamic problem.

Given this initial set of parameters and unobservables, I can start the iterative procedure described below.

D.2 Creation of the grid

In order to solve for the value function as a function of Θ, I need to create a grid describing the state space of the problem. Note that the state space is made of (λ, ϕ, n, X). Consequently, I need a grid that is relatively more precise for values of the unobservables that are more prevalent. Consequently, I create the four-dimensional grid as following

- $\lambda_g \sim N(0, 3\,\mathrm{std}(\lambda_{ft}))$
- $\phi_g \sim N(\mathrm{mean}(\phi_{ft}), 3\,\mathrm{std}(\phi_{ft}))$
- $X_g \sim N(\mathrm{mean}(X_{ft}), 3\,\mathrm{std}(X_{ft}))$
- $n_g \sim U[\underline{n}\ ;\ 1]$

Note that this grid is updated when the standard deviations or averages of the current unobservables are 20 percent larger or smaller than the ones used for the current grid, such that the grid will follow the potential change in the distribution of the unobservables. I set the size of the grid to be 20 on each dimension, such that the value function will be iterated at 20^4 different grid points.

Moreover, in order to solve the optimal mark-up of the firm, I also need to specify a set of grid points for the optimal mark-up term. I create a set of grid points mk_g of size $g_m = 20$, such that $\mu = \frac{\sigma}{\sigma-1}\frac{1}{mk_g}$, with $mk_g \equiv \{1 \cup \{\exp(\log(0.01) + \log(1.5)\frac{i}{g_m-2})\}_{i=0..g_m-2}\}$.

D.3 Iteration

Three different objects are updated at each iteration of the Markov Chain:

- the value function $V^{(s)}(\Theta^{(s)})$,
- the set of unobservables $\xi^{(s)}_{fdt} = (\lambda^{(s)}_{ft}, \phi^{(s)}_{ft}, X^{(s)}_{dt})$,
- the parameter vector $\Theta^{(s)}$.

I perform 60,000 iterations of the Markov chain, discarding the first 10,000 iterations. In the next paragraphs, I describe each of these following steps. I start by describing the step that aims to compute the value functions since they define objects that are used in the other steps. I then turn to the sampling of unobservables, and the sampling of parameters.

Update of the value function The value functions are obtained from the Bellman equation, iterated from the previous iteration of the value functions. From section 3, we have

$$V_I(\xi, n) = \max_{\mu}\Big\{E_\varepsilon\Big\{\pi(\xi, n, \mu, \varepsilon) + \beta EV'(\xi, n'(\xi, n, \varepsilon, \mu), 1)\Big\}\Big\}$$

Therefore, the value function is updated the following way:

$$V^{(s+1)}(\xi_g, n_g, \mathcal{I}, \Theta^{(s)}) = (1 - \delta^{(s)}\mathcal{I})s_v^{(s)} \log\Bigg[\exp\left(\frac{1}{s_v^{(s)}} EV_O(\xi_g)\right) \tag{15}$$

$$+ \exp\left(\frac{1}{s_v^{(s)}} \max_{mk \in mk_g}\Big\{E_\varepsilon \pi(\xi_g, n_g, mk_g, \Theta^{(s)}) - f^{(s)} + EV_I(\xi_g, mk_g)\Big\}\right)\Bigg] + \delta^{(s)}\mathcal{I}EV_O(\xi_g)$$

with

$$EV_I(\xi_g, mk_g) = \frac{\sum_{\xi \in \xi_g}\sum_{n \in n_g} V^{(s)}(\xi, n, I, \Theta^{(s)})\, P_n(n \,|\, \xi_g, mk_g)\, P_\xi(\xi \,|\, \xi_g)}{\sum_{\xi \in \xi_g}\sum_{n \in n_g} P_n(n \,|\, \xi_g, mk_g)\, P_\xi(\xi \,|\, \xi_g)},$$

$$EV_O(\xi_g) = \frac{\sum_{\xi \in \xi_g} V^{(s)}(\xi, n_0, 0, \Theta^{(s)})\, P_\xi(\xi \,|\, \xi_g)}{\sum_{\xi \in \xi_g} P_\xi(\xi \,|\, \xi_g)}$$

$P_\xi(.|.)$ being the transition probability of the unobservables at the current parameters, and $P_n(n \,|\, \xi_g, mk_g)$ the probability of obtaining a share n in the next period given the current unobservables ξ_g and the mark-up decision mk_g.[56] In practice, I iterate several times the Bellman equation, in order to reduce the error coming from the use of the previous value functions. In this case, I iterate not using the (s)-th value function anymore, but the current value function.

[56] This probability is obtained from the shock ε that makes the sales of the firms, and therefore the future share of consumers, non-deterministic.

In addition to updating the value function, I define, during this iteration, two objects that will be used in the sampling of parameters and unobservables. First, I save the optimal mark-up chosen by the firm. This object, evaluated on the grid, is defined as

$$mk_g^* \equiv \text{argmax}\Big\{E_\varepsilon\Big\{\pi(\xi,n,mk,\varepsilon)+\beta EV'(\xi,n'(\xi,n,\varepsilon,mk),1)\Big\}\Big\}.$$

Second, I create the difference in expected value functions, $DEV()$, that is defined as

$$DEV(\xi_g,n_g)=EV_I(\xi_g,mk_g^*)-EV_O(\xi_g).$$

This object will be convenient when computing the difference in value functions for each firm.

Sampling of unobservables I sample unobservables using the particle Gibbs with ancestor sampling (PGAS) sampler described in Lindsten et al. (2014), which relies heavily on the particle MCMC introduced in Andrieu, Doucet, and Holenstein (2010). The idea of particle MCMC is to develop techniques that use particle filtering in MCMC algorithm. Specifically, the PGAS iteratively sample parameters conditional to a particle and update that particle conditional to parameters. Importantly, the particle filter allows the sampling of the particle, proportional to the likelihood function, with good mixing properties. An important point of this sampler is that the current unobservables need to survive all the resampling steps of the filter. Moreover, I use ancestor sampling when choosing the specific set of particles, in order to further improve the mixing of the sampler.

To further describe the sampling of unobservables, I take the example the sampling of the unobservables λ_{ft} and ϕ_{ft}, conditional to the current unobservables at iteration s, $X_{dt}^{(s)}$. These unobservables are proportional to their prior distribution, $F_\lambda()$ and $F_\phi()$, and the conditional likelihood $L(D_{fdt}|D_{fdt-1},\lambda_{ft},\phi_{ft},X_{dt}^{(s)})$. The steps are the following:

- Starting from period 0, and for each firm, I generate r=1..20 particles $(\lambda_{f0}^r,\phi_{f0}^r)$ from their prior distribution $F_\lambda()$ and $F_\phi()$.
- I compute the likelihood of each of these particles for each firm:

$$L_{f0}^r \equiv \prod_d L(D_{fd0}|D_{fd-1},\lambda_{f0}^r,\phi_{f0}^r,X_{d0}^{(s)})$$

  using extrapolations from the functions $DEV(\xi_g,n_g)$ and $mk^*(\xi_g,n_g)$ to obtain the difference in value functions and the mark-up necessary to compute the likelihood.

- for each period t=1..T:
  - I resample 20 $(\lambda_{ft-1}^r,\phi_{ft-1}^r)$ proportionally to L_{ft-1}^r and replace $(\lambda_{ft-1}^{20},\phi_{ft-1}^{20})$ by $(\lambda_{ft-1}^{(s)},\phi_{ft-1}^{(s)})$.
  - generate 20 new particles, for each firm, from the prior distribution based on the resampled $(\lambda_{ft-1}^r,\phi_{ft-1}^r)$.
  - I compute the particle-specific likelihood $L_{ft}^r \equiv \prod_d L(D_{fdt}|D_{fdt-1},\lambda_{ft}^r,\phi_{ft}^r,X_{dt}^{(s)})$

- I retain one specific particle $(\lambda_{ft}^{(s+1)}, \phi_{ft}^{(s+1)})$ by using the ancestor sampling from period T to period 0.
  - Sample one particle for each firm $(\lambda_{fT}^{(s+1)}, \phi_{fT}^{(s+1)})$ proportionally to L_{fT}^{r}
  - for each period t=T-1..0:
    * sample $(\lambda_{ft}^{(s+1)}, \phi_{ft}^{(s+1)})$ proportionally to $\frac{L_{ft}^{r} F(\phi_{ft+1}^{(s+1)}|\phi_{ft}^{r}) F(\lambda_{ft+1}^{(s+1)}|\lambda_{ft}^{r})}{\sum_r L_{ft}^{r} F(\phi_{ft+1}^{(s+1)}|\phi_{ft}^{r}) F(\lambda_{ft+1}^{(s+1)}|\lambda_{ft}^{r})}$

This procedure gives me a new set of unobservables $(\lambda_{ft}^{(s+1)}, \phi_{ft}^{(s+1)})$ that have sampled proportionally to the likelihood function. Then, I perform a similar procedure to sample $X_{dt}^{(s+1)}$ conditional to the parameters and unobservables.

Sampling of parameters The sampling of parameters is made more complicated by the fact that functions $DEV()$ and $mk()$ need to be reevaluated for a new Θ. Therefore, sampling parameters requires to perform a Metropolis-Hastings step for which it is necessary to iterate the value functions for this new parameter Θ, similarly to the step updating the value functions. Formally, the sampling of a given block of parameter Θ takes the following steps:

- A new parameter Θ^* is drawn using a proposal function.
- The value function $V(\xi_g, n_g, I, \Theta^*)$ is obtained from equation (15) and the functions $DEV(\xi_g, n_g)$ and $mk_g(\xi_g, n_g)$ are computed.
- I obtain by interpolation DV_{fdt} and μ_{fdt}, allowing me to compute the likelihood function for the parameter Θ^*.
- $\Theta^{(s+1)}$ is set to be Θ^* with probability $\max\left\{1, \frac{\prod_t \prod_d \prod_f L_{fdt}(D, \xi_{fdt}^{(s+1)}; \Theta^*)}{\prod_t \prod_d \prod_f L_{fdt}(D, \xi_{fdt}^{(s+1)}; \Theta^{(s)})}\right\}$.

All parameters of the model, with the exception of the initial mean and variance of X, enter the dynamic problem of the firm. Therefore, all parameters enter the value function and should be evaluate using this Metropolis-Hastings step. However, because this step is relative time-consuming in the algorithm, I decide to only run the full Metropolis-Hastings step for some parameter blocks. For others, that are less likely to play a significant role in the dynamic problem, I use a simple Gibbs sampler that relies on specific parts of the likelihood function. Specifically, I update the different blocks of parameters as following:

- Parameters from the supply equation (α, γ_1,γ_2) are obtained from a Gibbs sampler based on the Bayesian regression of prices on quality λ_{ft} and country group dummies.
- The parameters of the variance matrix of demand and supply shocks (Σ_{11}, Σ_{12}, Σ_{22}) are sampled from an inverse Wishart distribution, based on the demand and supply residuals.
- The parameters of the AR(1) processes (ρ_λ, σ_λ, μ_ϕ, ρ_ϕ, σ_ϕ, μ_X, ρ_X, σ_X, μ_{X_0}, σ_{X_0}) are directly sampled from the Bayesian regression of the unobservables on their lags.
- The fixed costs parameters, variances of fixed costs and exit rate (f_1^c, f_2^c, f_3^c, f_1^e, f_2^e, f_3^e, σ_ν^c, σ_ν^e, δ) are sampled from the full Metropolis Hastings step as described above, using a random walk proposal function that targets an acceptance rate of 0.2.

- The parameters of the law of motion of n (η_1, η_2, n_0, $\underline{\mathrm{n}}$, ψ) are sampled from the full Metropolis Hastings step as described above, using a random walk proposal function that targets an acceptance rate of 0.2.

Overall, this procedure is doable thanks to parallelization using GPU computing. On average, an iteration of the Markov chain takes less than 3 seconds, which implies a total computing time of two days for 60 000 iterations.

D.4 Test on simulated data

To test my empirical procedure, I simulate a set of data following the data generating process assumed in the model. Then, I implement my estimation procedure to test the validity of the estimation. However, because of the complexity of the estimation, I cannot perform a full Monte Carlo study of the estimation method. Therefore, I cannot test whether my estimator consistently recovers the true value of the parameters, but instead whether the true value of the parameters belongs to the confidence interval obtained from the estimation. I simulate data for 200 firms, 14 years and 14 destinations and run 60 000 iterations of my algorithm, discarding the first 10 000, as I do in the estimation procedure. Moreover, in order to evaluate the errors made through the interpolation procedure, I simulate data using a grid of 25 points, while only 20 points are used in the estimation.[57] I report in figures 16 the Markov chains for all parameters, as well as the true value of the parameters displayed by the red lines. As displayed on these figures, the estimation provides confidence intervals that are consistent with the true value of the parameters. Even though I do not perform a true Monte Carlo experiment with this estimator, this is reassuring regarding the validity of the procedure.

[57] Unfortunately, I was unable to run the problem with a grid larger than 25^4.

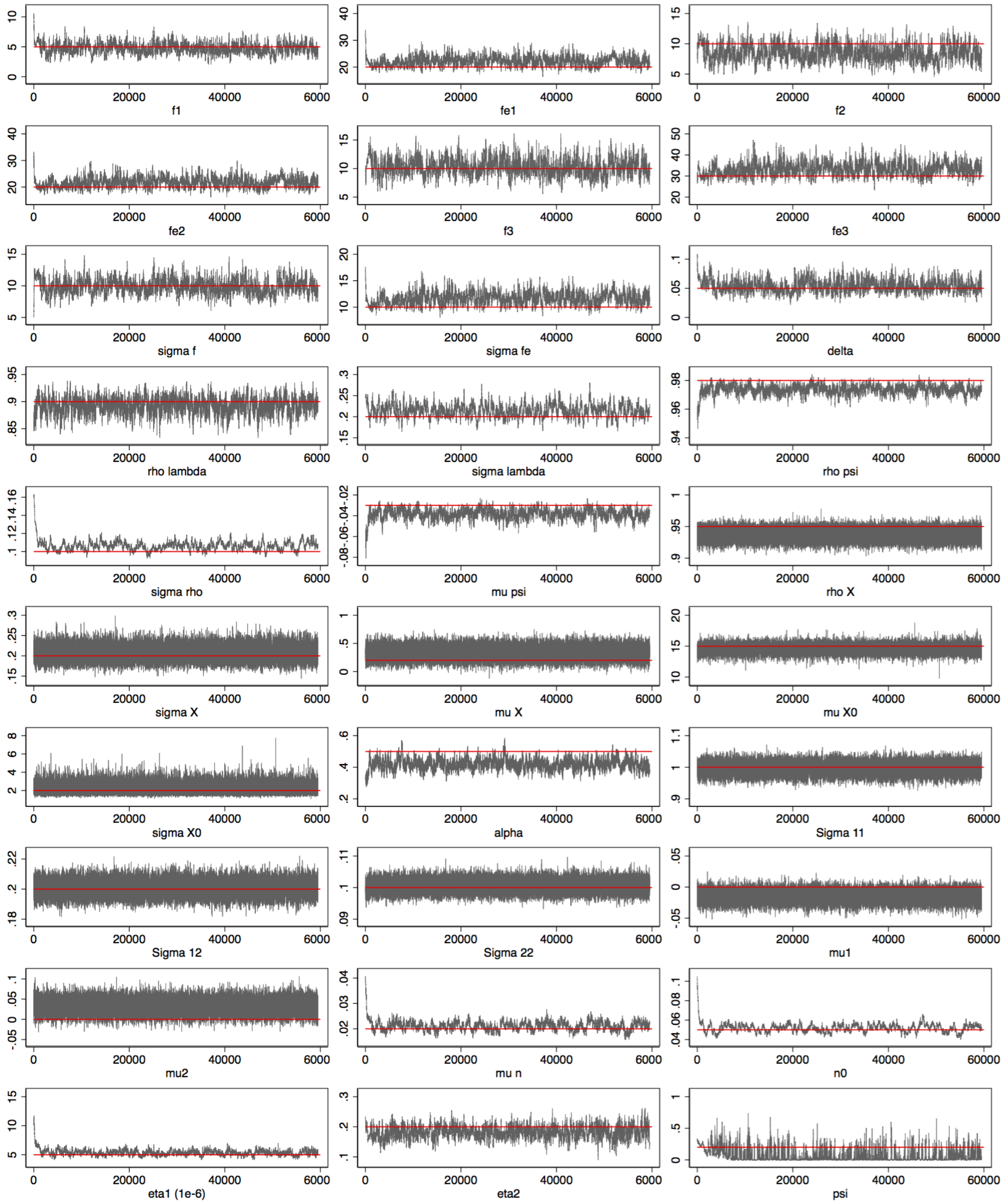

FIGURE 16: Markov Chains from the estimation on simulated data.

E Additional results

E.1 Markov Chains

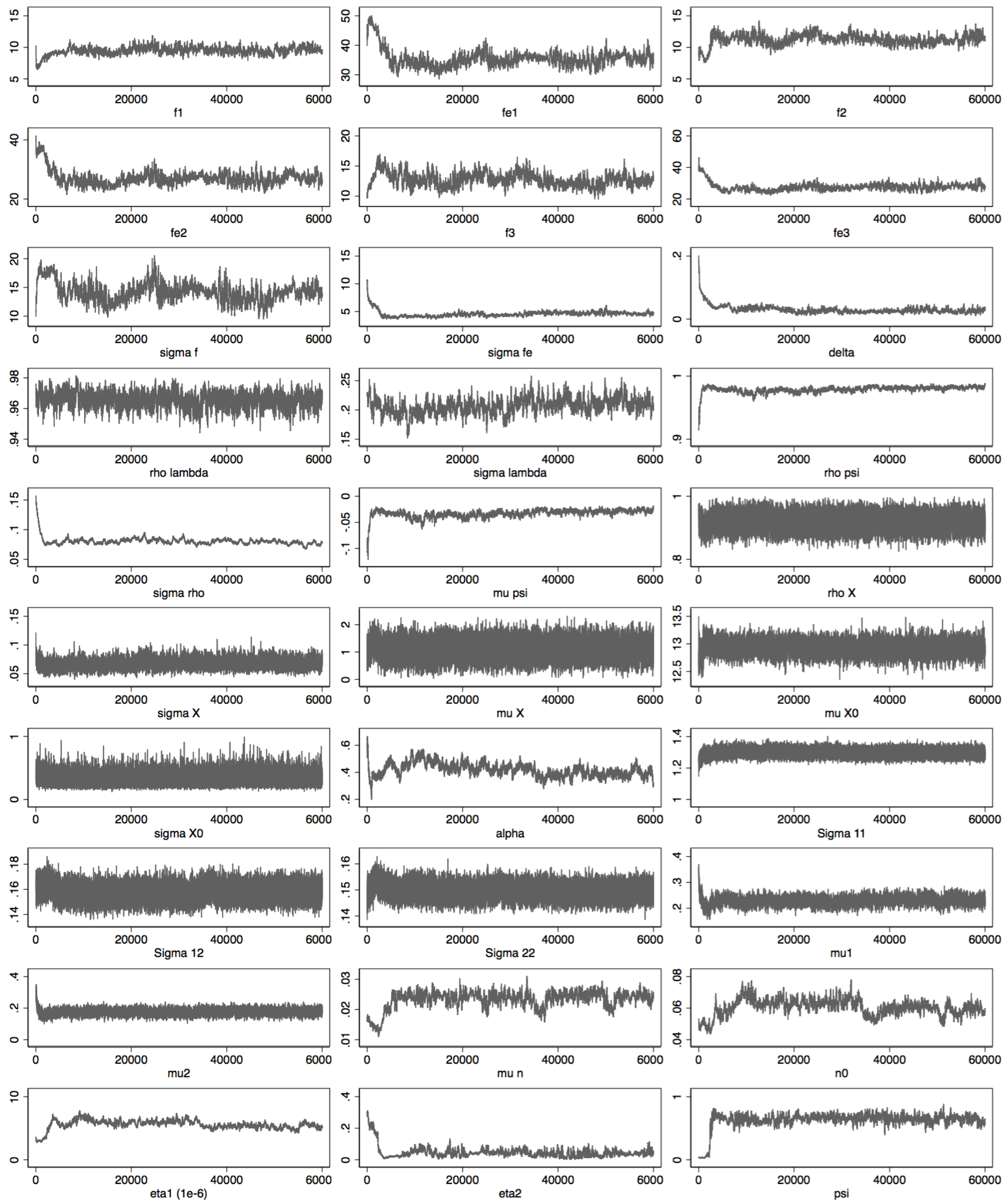

FIGURE 17: Markov Chains from the estimation.

E.2 Full results of the restricted model

Table 9: Estimated parameters - restricted model

Parameter		**Estimate**	**90% Confidence Interval**	
			Lower bound	*Upper bound*
Continuation fixed costs	Europe	10 146	9 308	10 789
(in euros)	Americas	11 981	10 779	13 013
	Asia/Oceania	13 848	12 304	15 288
Entry fixed costs	Europe	60 365	55 436	65 400
(in euros)	Americas	48 158	43 978	52 610
	Asia/Oceania	53 984	49 687	58 608
Variance of continuation costs	σ_ν^c	17 033	14684	19 178
Variance of entry costs	σ_ν^e	11 844	10 866	12 903
Exit rate	δ	0.00	0.00	0.01
Law of motion of appeal	ρ_λ	0.98	0.97	0.98
	σ_λ	0.34	0.31	0.37
Law of motion of productivity	ρ_ψ	0.97	0.97	0.98
	σ_ψ	0.10	0.10	0.11
	μ_ψ	-0.04	-0.05	-0.03
Law of motion of agg. demand	ρ_X	0.96	0.93	0.99
	σ_X	0.15	0.13	0.18
	μ_X	0.54	0.23	0.86
	μ_{X_0}	11.67	11.27	12.05
	σ_{X_0}	0.79	0.56	1.10
Elasticity cost of appeal	α	0.20	0.12	0.24
Cost dummies	γ_2	0.30	0.28	0.31
	γ_3	0.28	0.25	0.30
Variance matrix	Σ_{11}	2.20	2.14	2.26
	Σ_{12}	0.29	0.28	0.30
	Σ_{22}	0.14	0.14	0.15

E.3 Additional figures

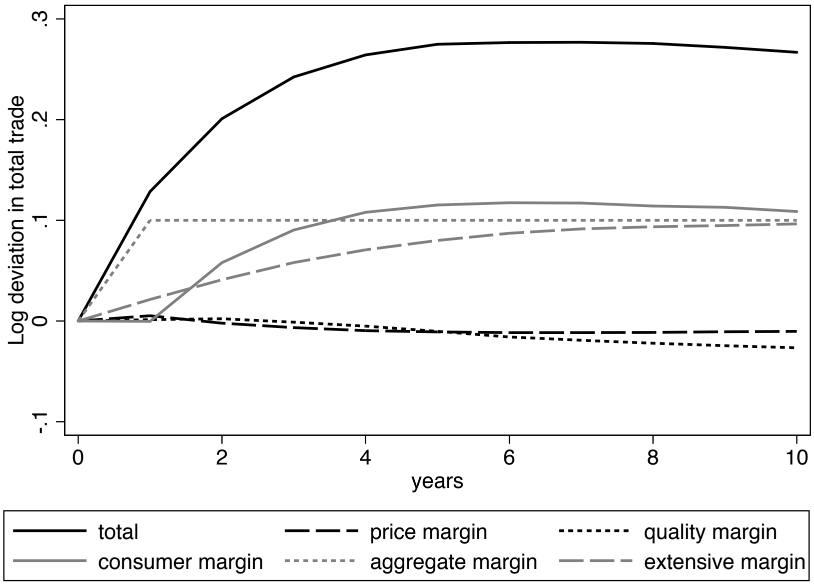

Figure 18: Effect of permanent 10 points tariffs decrease (All margins).

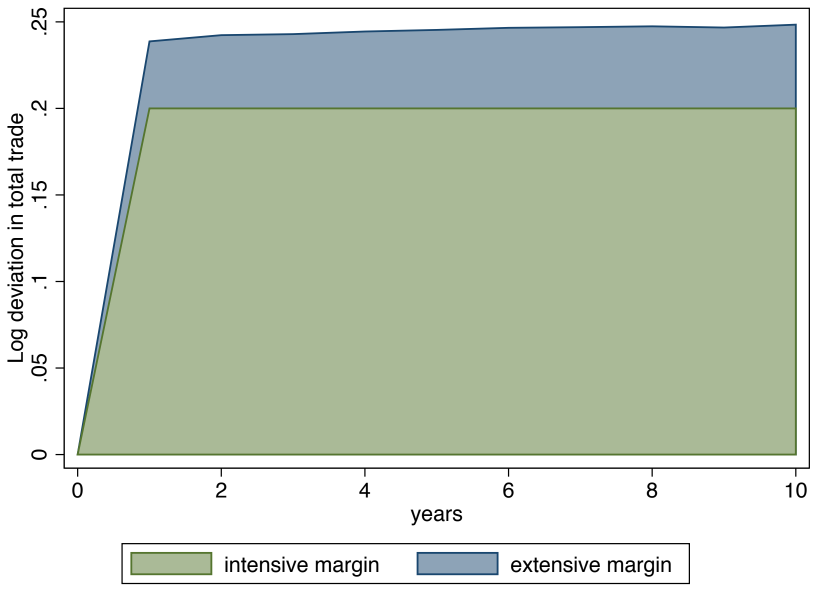

Figure 19: Effect of permanent 10 points tariffs decrease (Restricted model).

F Details on out-of-sample predictions

In order to perform out-of-sample predictions, I construct variations in the aggregate demand from Brazil. This variable is defined from the model as $X_{dt} = \log Y_{dt} - (1-\sigma)\log P_{dt} + (1-\sigma)\log(\tau_{dt} e_{dt})$. Importantly, I only need to construct variations of aggregate demand since the level will be chosen to exactly match the total exports of French firms to Brazil. Therefore, in addition to using the Brazilian GDP, the exchange rate between France and Brazil, I also need to construct a proxy for variations of the price index in the Brazilian market. In order to do so, I use variations in exchange rates from the five main countries exporting to Brazil. Table 10 describes these countries and their respective market shares.

TABLE 10: Top market shares

Country	**Average market share**
France	22.1 %
Italy	20.4 %
Chile	19.6 %
Portugal	15.6 %
Argentina	13.5 %

Notes: Calculations made from BACI. Average market share is the average market share among the Brazilian imports, over the period 1997-2007, for the 4-digit category 2204 '*Wine of fresh grapes*'.

Note that the next largest wine exporter to Brazil has a market share of less than 2 percent and is therefore not included in the computation of the price index. Therefore, I construct variations in $X_{B,t}$ as following:

$$\begin{aligned}X_{B,t} - X_{B,98} =& \log Y_{B,t} - \log Y_{B,98} - (1-\sigma)\left[\log P_{B,t} - \log P_{B,98}\right] \\ &+ (1-\sigma)\left[\log(\tau_{F,t} e_{F,t}) - \log(\tau_{F,98} e_{F,98})\right] \\ =& \log Y_{B,t} - \log Y_{B,98} - \log\left(\sum_i \omega_{i,98}\left(\frac{e_{i,t}}{e_{i,98}}\right)^{1-\sigma}\right) + (1-\sigma)\log\left(\frac{e_{F,t}}{e_{F,98}}\right)\end{aligned}$$

where the difference in Y is computed from changes in the Brazilian GDP, $\omega_{i,98}$ are the import shares of each of the five countries displayed in the table 10 and $e_{i,t}$ their exchange rates with Brazil. The obtained variations in aggregated demand for French wine is described in figure 20, which highlights the impact of the Brazilian and Argentinian devaluations.

I then perform 500 simulation of trajectories using the median product appeals and productivities obtained from the estimation, and the constructed variation in aggregated demand. I set the level of X_{dt} such that the median prediction exactly predicts the right amount of export from France to Brazil in 1998. Trajectories differ because I need to simulate demand and supply shocks (ε) and fixed costs shocks (ν) in order to obtain predictions for each firm. Predictions reported in the text are based on the median trajectories, and I report in figure 21 the 90% confidence interval of these predictions.

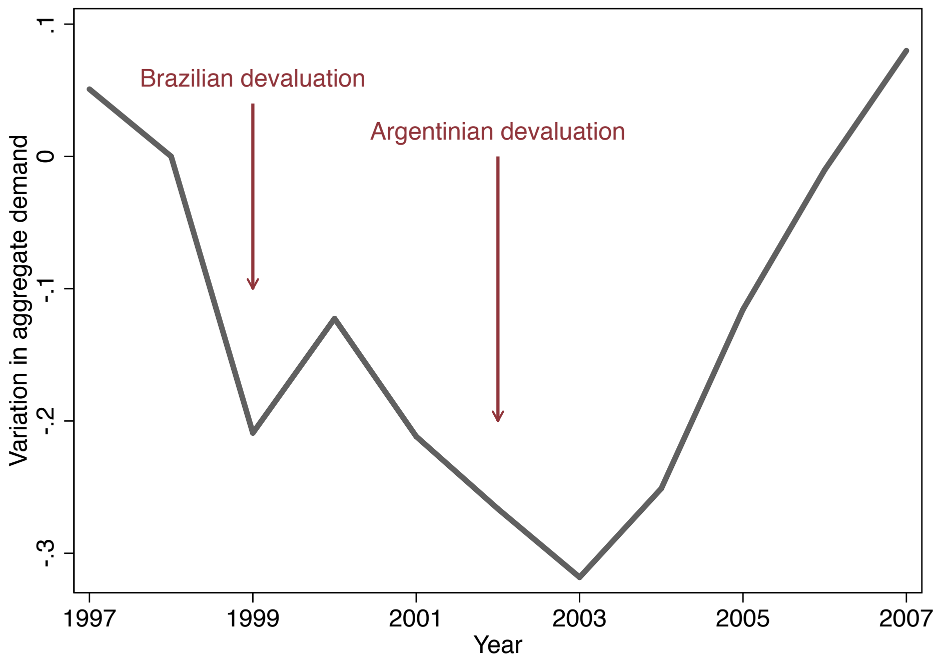

FIGURE 20: Computed variations in aggregate demand for French wine from Brazil.

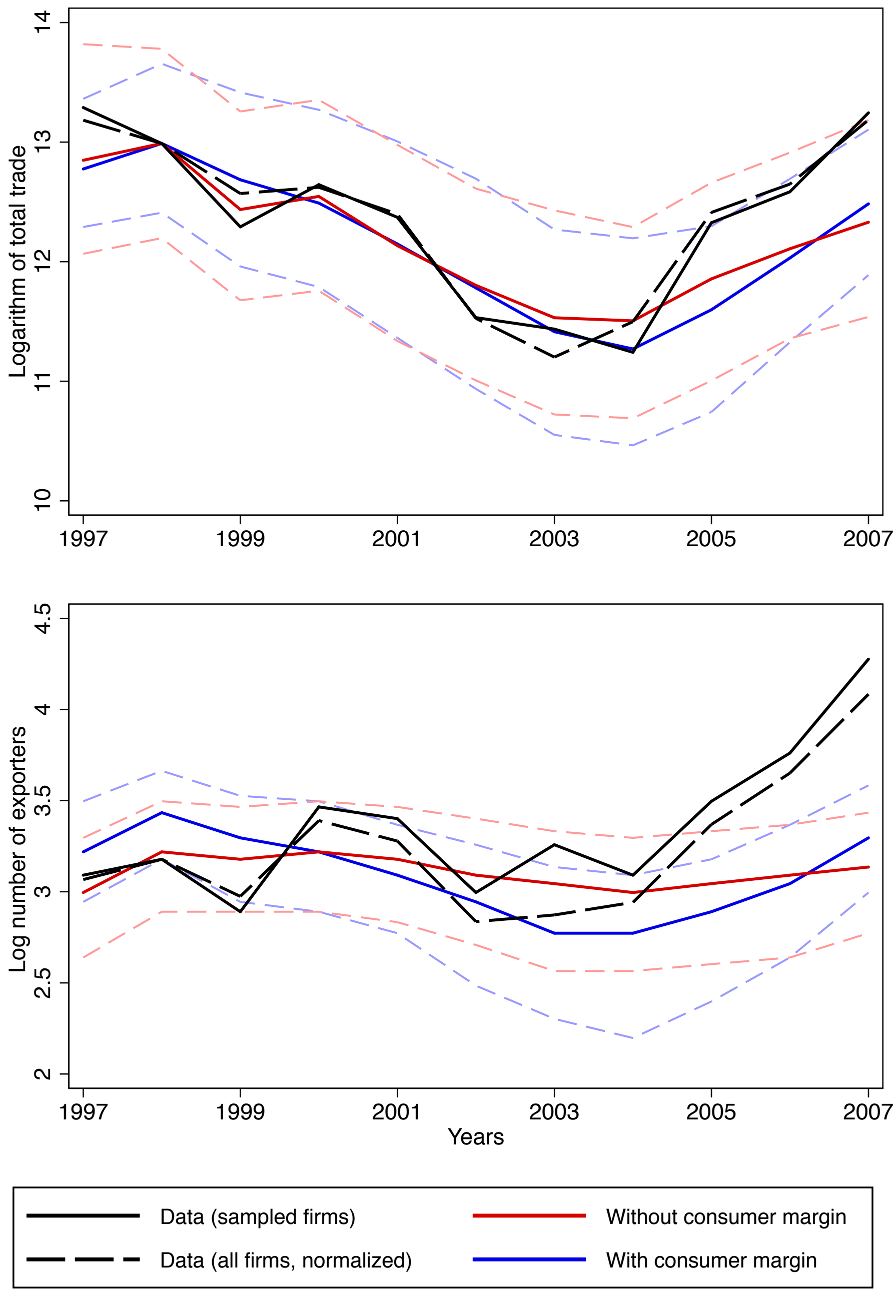

Figure 21: Total exports of wine to Brazil from selected firms