An empirical dynamic model of trade with consumer accumulation*

Job Market Paper

Paul Piveteau

Department of Economics, Columbia University[†]

November 8, 2015

updates at: http://www.columbia.edu/~pp2382/research.html

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 on foreign markets. This paper introduces consumer inertia in a dynamic model of trade through the accumulation of consumers in foreign markets, which is consistent with the increase in sales and prices over time that I document for new exporters. Estimating the model using export data from individual firms, the model correctly predicts lower survival rates for new exporters and also estimates much lower entry costs of exporting - about one third of those estimated in the standard model. These results have important implications at the aggregate level. In contrast to the standard model, this model correctly replicates the slow response of trade to shocks and the increasing contribution of the extensive margin in this response. Moreover, out-of-sample predictions demonstrate that the model better predicts actual trade responses to an observed shock than the standard model.

*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, Ildiko Magyari, Thierry Mayer, Antonio Miscio, Ferdinando Monte, Jean-Marc Robin, Bernard Salanie, Gabriel Smagghue, Daniel Xu and seminar audiences at Columbia University and Sciences Po Paris for comments and suggestions. Part of this research was conducted while I was visiting the economics departments of the ENS Cachan and Sciences Po Paris, I thank them for their hospitality. I also would like to thank the Alliance program and Columbia University CIBER for financial support and the CNIS and French customs for data access. All remaining errors are mine.

†1022 International Affairs Building, 420 West 118th Street, New York, NY 10027. Email: pp2382@columbia.edu. Website: http://www.columbia.edu/~pp2382/.

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% of the decline of international trade with geographical distance is due to a reduction 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 exporters' dynamics that have been recently documented in the literature: most new exporters start small and only a small fraction survives and expands on 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 exporters' dynamics, I develop a Markov Chain Monte Carlo (MCMC) 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 about one third 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 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 exporters' 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 will have a different stock of consumers, depending on its history in the

[1]This finding is consistent with recent papers that show the importance of demand characteristics as source firm heterogeneity (Hottman, Redding, and Weinstein, Forthcoming; Roberts, Xu, Fan, and Zhang, 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.

foreign market; this will shape 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 will start with low levels of sales and profits when entering a new destination. As they survive and accumulate consumers, their sales and profits will 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 exporters 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 of the model, I employ a Markov Chain Monte Carlo (MCMC) estimator that will account for this unobserved heterogeneity, and facilitates the solution of the dynamic problem of the firm. Therefore, this estimator will allow me to obtain value estimates of the entry and per-period fixed costs of exporting, which will be identified by rationalizing the actual entry and exit patterns of exporters on the different export markets.

The results of the estimation demonstrate the importance of the accumulation of consumers to replicate exporters' dynamics. The introduction of state dependence in demand improves the ability of the model 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 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 33 000 euros, around 78% of the average revenue in these destinations.[5] Because the accumulation of consumers accounts for an important part of the dependence in export decisions, large entry costs become unnecessary to rationalize the hysteresis in export markets. To confirm this finding, I estimate a version of the model without consumer accumulation and obtain an estimate of the average entry cost to European destinations of 98 000 euros, roughly three times the estimates of the full model.

These results have important implications at the aggregate level. In particular, the model will generate aggregate adjustments in response 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

[3] Recent empirical evidence for this type of mechanism on domestic market was found by Foster et al. (2012) who studied the behavior of new firms producing homogeneous goods.

[4] Khandelwal (2010) at the product level or Hottman, Redding, and Weinstein (Forthcoming) at the micro level, also define appeal or quality as the demand shifter after controlling for prices in a demand equation. However, I assume that appeal does not vary across destinations.

[5] Or equivalently 2.7 times the median yearly revenue on these destinations.

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 will be larger in the long-run than the short-run. In my simulations, the ratio between the long and the short-run elasticities is around three, a value that is consistent with the ratio of elasticities used in the international trade and international macroeconomics literature. Second, the model can predict the increasing contribution of the extensive margin during a trade expansion. Recent papers, Kehoe and Ruhl (2013) and Alessandria et al. (2013) in particular, document how the extensive margin tends to have a small contribution in the short-run but plays a significant role in the long run in explaining trade growth. The model with consumer accumulation generates a relative contribution of the extensive margin two to three times larger in the long-run than in the short-run. Because the technology for accumulating consumers displays decreasing returns, new exporters will record larger growth than established exporters in the years following the shock, hence increasing their contribution to trade relative to older exporters 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 on 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 is able to replicate, unlike the standard model without consumer accumulation, the decrease 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) will predict larger inflows and outflows of exporting firms, and therefore larger variations in total trade.

This paper is closely related to the literature investigating exporters and firms 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. Many recent studies have documented and studied the specific dynamics of new exporters. Nguyen (2012), Albornoz et al. (2012), Berman et al. (2015) and Timoshenko (2015) emphasize the role of demand uncertainty and experimentations to explain exporters 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. Eaton et al. (2014) also develop a structural approach to study exporters' dynamics: they use an importer-exporter matched dataset to estimate a structural model of export entry in which exporters grow through

[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 (2015) similarly shows the prevalence of entry costs when estimating fixed costs of exporting for US firms.

the search of foreign distributors and the learning of their own ability.[8] However, while they do not allow for other margins of firms growth on foreign markets, my model will feature 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 exporters' 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 exporters' dynamics. Drozd and Nosal (2012) and Gourio and Rudanko (2014) show how convex adjustment costs of market shares can respectively explain several puzzles in international macroeconomics and adjustments of important variables 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 elasticities of trade.[9] This series of papers develops macroeconomics 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, Choi, and Ruhl, 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 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 allow 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), that allows me to account for the existence of persistent unobservables, as well as solve the full solution of the DDCM.[10]

The outline of the paper is the following: in the next section, I will present stylized facts about the trajectories of exporters, that will emphasize the importance of demand in exporters' 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 will inspect the aggregate implications of the estimated results through simulations and out-of-sample predictions, and section 7 will conclude.

[8] See also Akhmetova and Mitaritonna (2012) and Li (2014) that show the importance of demand uncertainty, and Aw et al. (2011) looking at the impact of R&D activities on exporter decisions.

[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 estimation method in Industrial Organization can be found in Osborne (2011).

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 the age on the foreign market, even after controlling for survival. Third, prices also increase with the age of exporters.

These facts are consistent with the empirical model I will present in the next section: first of all, the high level of attrition across age will require the model to account for endogenous selection. Moreover, the rise in sales, while prices increase on average, 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 will start small, and will 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 I used in this paper is provided by the French customs services. These data record yearly values and quantities exported by French firms from 1995 to 2010.[11] Yearly trade flows are disaggregated at the firm, country and eight-digit product category of the combined nomenclature (CN). This dataset will be used to present stylized facts about new exporters in this section, and a restricted sample 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, and improve the reliability of the unit values. The partial-year bias comes from the mismatch between calendar years and exporting years: because trade data are based on calendar years, the first year of activity of a new exporter will report lower sales on average, since this exporter potentially entered anytime during that year.[12] These partial years will imply an overestimation of the growth rate between the first and second year of export. To correct for this bias, I readjust the dataset using information available at the monthly level. For each new entry by a firm on a new destination, I readjust the month of entry, and adjust accordingly the dates of the subsequent exporting flows for that firm. Aggregating this adjusted dataset at the yearly level, I obtained a transformed dataset that does not display this bias. Second, in order to improve the reliability of the unit values, I drop all the product categories that use weight as unit of measure. Even though the weight of a product is sometimes the relevant unit for that product, it appears that it is used as unit when the type of product in a category is not homogeneous, and therefore casts some doubt on the use of these quantities to create unit values.[13] In addition to these two important adjustments, Appendix A describes additional procedures implemented on the dataset to improve its reliability.

[11]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 her 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.)

[12]See Berthou and Vicard (2015) and Bernard, Massari, Reyes, and Taglioni (2014) for papers investigating the extent and consequences of this bias.

[13]The main patterns displayed in the next subsection, in particular relatively to prices, appears to hold when using the products that use weights as units.

Table 1 provides some information on the distributions of the number of observations along different dimensions. Trade flows are known to be sparse across firms and destinations, and trade flows from France are no exception. This is true for a firm across destinations or product categories in a given year, since the median exporting firm records two 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 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	8.49	1	1	2	5	28	671 403
by firm-CN8-year	2.21	1	1	1	2	8	2 581 098
by firm-dest-year	2.60	1	1	1	2	8	2 189 506
Exporting spells duration (years)							
firm-dest-CN8 level	1.67	1	1	1	2	5	3 413 456
firm-dest level	2.01	1	1	1	2	7	1 091 995

Notes: CN8 denotes an eight-digit category from the Combined nomenclature, after normalization following Pierce and Schott (2012). An exporting spell is defined as a set of consecutive yearly exporting flows.

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 on 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 will be augmented with fixed effects that will 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 specifications are the following

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

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

where age_{fpdt} is defined as the number of consecutive years a firm f has been selling the good p to destination d. Y_{fpdt} will be respectively the logarithm of export sales, the logarithm of prices (unit values),[15] and a dummy equal to one if the firm is still exporting to the market the following year. μ_{pdt} will be a market×year-specific fixed effect such that the variations that identify 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, imply that firms surviving 10 years are substantially different than firms who started to export. Consequently, the variations that the regressions will capture when comparing old and new firms will mostly come from a selection effect comparing different set of firms, rather than changes across ages for a given set of firm. In order to partially account for this dynamic selection, I also present the results when only looking at firm-product-destination triplet that survive 10 years on their specific markets. Even though this only partially accounts for selection, since surviving firms are also firms with specific trajectories, it will show that the observed relationships are not only due to dynamic selection, but also comes from variations within a constant set of firms.

Another possibility to partially account for this dynamic selection would be to use firm-product fixed effects, or first difference transformations. These transformations would control for the heterogeneity across firms, and only capture variation within a firm-product-destination triplet across ages. However, the identification of a trend with age is not possible using variations within a given triplet. The intuition comes from the fact that the increase of age is a treatment that applies to all firms, and therefore cannot be separately identified from a cohort effect. I discuss related specifications at the end of the section.

2.3 Results

I respectively present three important facts about exporters, namely the growths of the survival rates, exported values, and prices with export experience on 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] However, I show that these facts still hold after controlling for the partial-year bias highlighted by Berthou and Vicard (2015) and Bernard, Massari, Reyes, and Taglioni (2014). Moreover, the increase of prices has not been documented, to my knowledge, using a comprehensive trade dataset, even though Foster, Haltiwanger, and Syverson (2012) 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 on a market, i.e. to export on this market the following year, is very low for the average exporter. Figure 1 displays the average survival rate for a

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

[16] See for instance Ruhl and Willis (2008) for a presentation of these facts and the associated puzzles.

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 only roughly 35 percents. However, this survival probability rapidly increases once exporters have survived several years: this rate is larger than 50 percents at age 2, and close to 75 percents at age 6. This result reflects the same idea highlighted in the previous section that most export spells are very short lived.

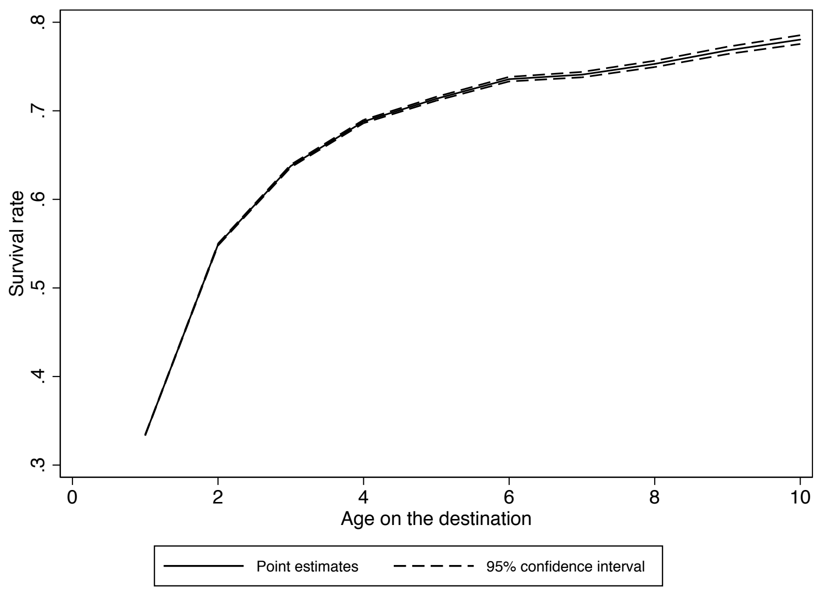

FIGURE 1: Survival rates across export ages

Notes: The figure reports the average survival rate of a firm-product category pair on a destination at different ages. The estimates are obtained from the regression (1) that uses as dependent variable a dummy equal to one when the firm-product pair is exporting to the destination in the following year, and includes product category×destination×year fixed effects. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors estimates clustered at the firm-product-destination level.

These low and increasing survival rates will 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, 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's age on 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 by plotting the results from regression (1), after normalizing the average log sales at age one to be zero. When comparing exported values, exporters which are in their third year of exporting will export more than twice more relatively to a new exporter. This difference reaches an order of 7 when comparing a 10 years old exporter with a new exporter. However, it is important to note that these differences are mostly due to a strong selection across exporters: old exporters, who by definition managed to survive on foreign markets, were initially larger than the average new exporter. The right panel in Figure 2 emphasizes this point by looking at the relationship when restricting the set of exporters to those surviving 10 years. 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 percents between ages one and two. Moreover, this growth appears to continue the first six years: at this age, exporters tend to be in average twice larger relatively to their first year of exporting.

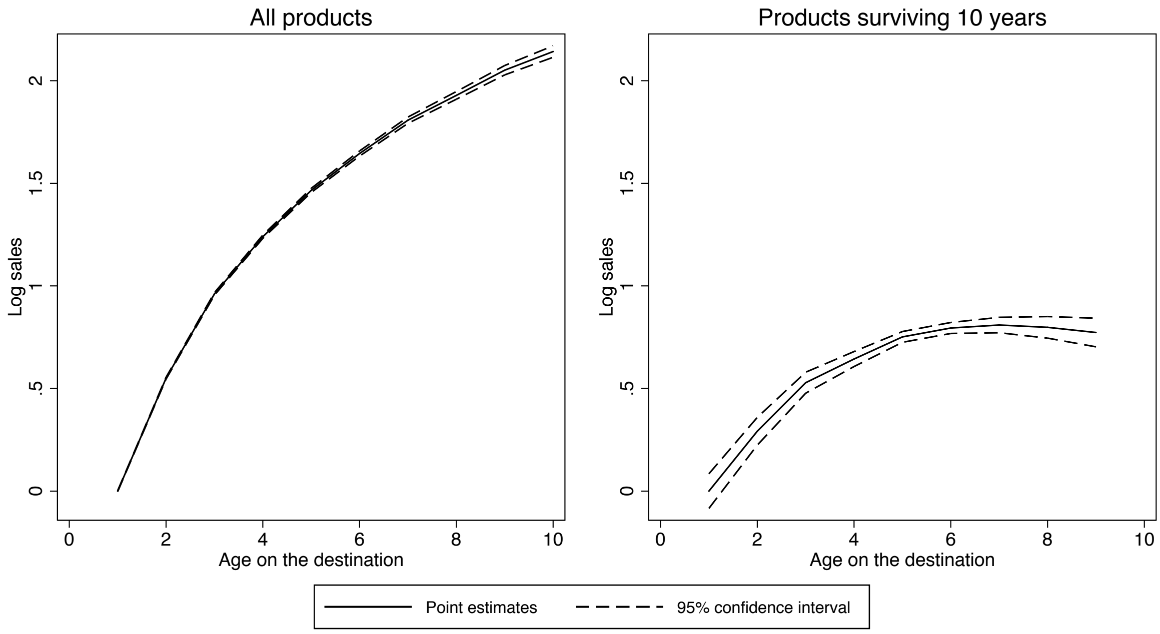

FIGURE 2: Sales across export ages

Notes: The figure reports the cumulative growth of sales relatively to age one, of a firm-product category pair on a destination at different ages. The estimates are obtained from the regression (1) that uses logarithm of sales as dependent variable, and includes product category×destination×year fixed effects. The left panel reports the results of this regression on the entire sample, while the right panel reports the result from an estimation using only the sample of firms that reach age 10. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors estimates clustered at the firm-product-destination level.

In conclusion, we observe substantial growth rates of sales during the first years of exports. These growth rates are large but appears 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, Massari, Reyes, and Taglioni (2014). 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 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 role of the structural model introduced in the next section.

Fact 3: Export prices tend to increase with firm's age on a destination, even more so conditional on surviving.
One potential reason for the growth in sales could be productivity improvements that lead to a reduction in the prices of the good exported, and therefore an increase in its sales. On the contrary, it appears that prices also increase with the experience of the firm on the export market.

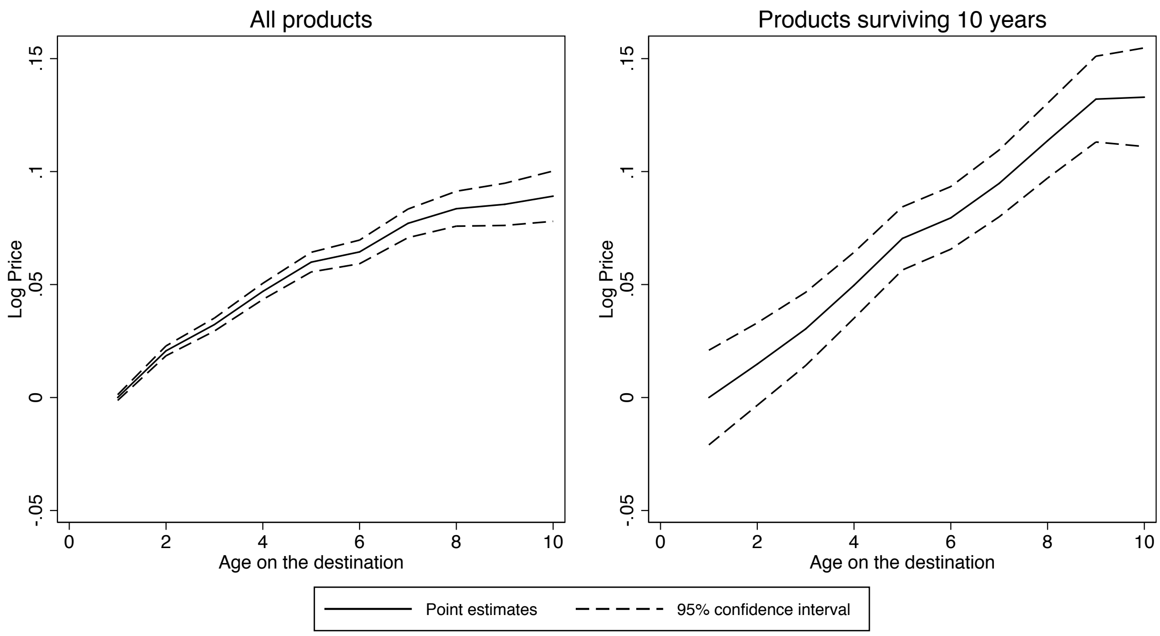

FIGURE 3: Prices across export ages

Notes: The figure reports the cumulative growth of prices relatively to age one, of a firm-product category pair on a destination at different ages. The estimates are obtained from the regression (1) that uses logarithm of unit values as dependent variable, and includes product category×destination×year fixed effects. The left panel reports the results of this regression on the entire sample, while the right panel reports the result from an estimation using only the sample of firms that reach age 10. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors estimates clustered at the firm-product-destination level.

Figure 3 reports the estimated parameters of regression (1) in which the average price at age one is normalized to zero. The left figure shows that the price of a 10 years old exporter is in average 9 percents higher than the price of a new exporter. Similarly to sales, this effect could come from a selection effect of the exporting activity: a selection process driven by the quality of the product for instance, would imply that older firms which managed to survive, have higher prices than young exporters. However, when controlling for selection by looking at surviving firms (Right panel of Figure 3), it appears that the growth of prices is even larger compared to the regression using the full sample: the price after 10 years appears to be in average 12 percents larger than the price charged by the same firm at age one.

Observing a larger growth of prices when looking at a constant sample of firms has two important implications. First of all, it means that costs are the main driver of the selection process: high price firms tend to disappear more in the first years such that the positive correlation between prices and age is weakened when using the full sample. Second, it implies that this positive correlation cannot be only driven by dynamic selection. Therefore, an additional mechanism is necessary to explain why firms tend to increase their price during their exporting life. The structural model presented in the following section will introduce such a mechanism, through the dynamic pricing of the firms.

There exists other methods that can partially account for the endogenous selection across ages. However, within variations cannot be used in this context as it is not possible to separately identify the role of experience, cohort and trend effects. In appendix B, I describe results from two specifications that use related sources of identification. The first one includes a set of firm-product-destination fixed effects, such that the identification only comes triplets that exit the market and reenter a few years later.[17] This specification documents a decreasing trend for prices and a hump shape of sales, which confirms that high price products tend to survive less in average. The second specification introduces a set of firm-product fixed effects such that the variation is obtained from the same firm-product pair which is selling to different destinations, with different ages. A potential issue with this specification comes from the endogenous sorting across destinations: older destinations are also the ones to which the firm has decided to export first. The results appear consistent with this mechanism: sales appear to grow faster with this specification, while growth in prices are smaller but still positive. Detailed results are provided in appendix B.

This section introduced simple facts about exporters' dynamics that will guide the empirical model developed below. We can draw three conclusions from these figures. First, survival rates are very low on export markets and grow with the age of the firm. This result has two consequences: it implies that the entry decision needs to be accounted for when studying the dynamic problem of the firm. Moreover, this fact is contradictory with a world where the main barrier to export is made of sunk entry costs: in such a world, exporters would tend to keep exporting once they have overcome this important barrier. Second, sales of exporters grow rapidly in the first years of exporting. These large growth rates are also present when accounting for dynamic selection across firms. Third, this increase in sales is driven by a growth in the demand of the firm: prices variations cannot explain this large increase, implying the importance of demand characteristics as main drivers of this increase in sales. On the contrary, it appears that prices tend to rise across ages, even more so when controlling for dynamic selection. This pattern could be explained by a dynamic behavior of the firms that foster their growth in the early years by reducing their prices.

Despite these conclusions, it is difficult to make strong causal statements by comparing firms across different ages. This brings to light a second benefit of developing and estimating a structural model to study the entry and growth of exporters: in addition to understanding the dynamic decisions of firms, it will allow the model to control for the endogenous sorting and attrition of firms, and recover the different processes that drive the observables variables of the model. The next section introduces this model.

[17] They are the only triplets that go 'backward' in age, and therefore are the only sources of variation.

3 Structural model of export entry

This section describes an empirical model of entry into foreign markets in which the accumulation of consumers constitutes a new source of dependence in the dynamic problem of the firm. This model aims to identify the different sources of profit on foreign markets to understand the export decisions made by firms. Therefore, it is crucial to allow for heterogeneity across firms and destinations, but also to allow this heterogeneity to be persistent over time. Indeed, persistent heterogeneity will be the main competing hypothesis to sunk entry costs to explain the persistence in export decisions. As a consequence, this model will feature 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 will depend on four characteristics: productivity, product appeal, aggregate demand and consumer share.[18]

The introduction of consumer accumulation will imply two deviations from the standard dynamic model, which will be consistent with the stylized facts presented earlier: first, firms will 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 accumulation of consumers comes from sales, firms will have dynamic incentives to lower their prices in the first years of exporting 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 the entry and pricing decisions.[19] In particular, the optimal price charged by the firm will depart from a constant mark-up over marginal costs to take into account the dynamic impact of prices on consumer accumulation.

3.1 Demand

There exists a wide range of mechanisms that can give rise to inertia in consumption and state dependence in demand. A large literature in industrial organization has found empirical evidence of this type of behavior and have studied their consequences on the market equilibria and the pricing behaviors of firms. 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. One can cite the existence of habits in consumption, the fact that searching new products is costly, or the failure of perfect information for the consumers about goods as examples of economic explanations that leads to state dependence in the demand formed by an agent (see for instance Dubé, Hitsch, and Rossi (2010) for a paper distinguishing and measuring the contribution of these different mechanisms).

[18] Therefore, I will assume that entry decisions are independent across destinations, once controlling for firms' characteristics, which will keep the state space of the dynamic problem relatively small. McCallum (2015) provides support for this assumption by finding that entry costs of exporting are mostly country specific. See also Morales et al. (2014) for a paper that use moments inequalities to maintain such a large state space.

[19] Note that I do not study the choices made by the firms for each product she could potentially export. Firms are seen as single-good producers in this model, and will be considered as such in the empirical application using wine producers.

In order to keep the model tractable, I will introduce state dependence in demand through the existence of a firm-specific customer base on each destination. This customer base, denoted n_{fdt}, describes the share of consumers, on a destination d at time t, that includes the product f in its consideration set. This representation follows the marketing literature that defines a consideration set as the set of products that consumers consider when making purchase decisions.[20] It is also consistent with the idea of customer margin introduced in the macroeconomic and international trade literature.[21]

Therefore, I will assume that a new exporter has an initial share of consumer n_0 when it enters a new foreign destination. In the subsequent years, the consumer awareness of the products will be propagated through two mechanisms. First, the sales of a product will increase its awareness in the next period. Specifically, an euro increase in the sales of a product will increase by η_1 the potential share of consumers in the next period. This acquisition of consumers can arise in a situation in which consumers have imperfect information about product characteristics, and therefore use sales as a signal for the expected utility gain obtained from consuming a good.[22] Second, another source of consumer accumulation will come from word-of-mouth: I will assume that each aware consumer will share its awareness with η_2 consumers. Both of these mechanisms will 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 will be 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' is

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

This specification is largely inspired from the marketing literature as described in Arkolakis (2010): the accumulation of consumers has decreasing returns such that it is more difficult for an established firm to accumulate more consumers relatively to a firm with a small initial share. Indeed, for established firms, a significant share of these newly reached consumers will already be part of their consumer share, hence not contributing to its growth. Therefore, the parameter ψ will describe the importance of these decreasing returns, and the two parameters η_1 and η_2 will characterize the importance of the two different sources of growth in the accumulation process.

These two different margins of growth will capture different mechanisms of consumer accumulation, but more importantly will generate different optimal responses by the firm. In a world with word-of-mouth, where consumers learn from their neighbors, the growth of this consumer share could be seen as exogenous, only based on the past share of consumers. In this world, firms cannot affect this accumulation with their pricing decisions.[23] However, in a world where

[20] See for instance Shocker et al. (1991) for an article studying the importance of consideration sets in consumers' decisions.

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

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

[23] This model does not take into account advertising as a source of growth, even though this could be a natural candidate to foster consumer accumulation. The inability to observe this type of expenditures in trade datasets makes it difficult for an empirical model to account for this channel.

consumers face uncertainty regarding product characteristics and sales are seen as a signal, firms will have incentives to reduce its price in order to foster the accumulation of consumers. This distinction between these two sources of growth brings back to the distinction between structural and spurious structural dependences (Heckman, 1981), that generate different optimal responses by the firm.

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 will depend on the sales and the share of consumers of the previous period in this market.

This share of consumer will act as a demand shifter for the firm since it will scale the amount of demand the firm will receive from each destination. To obtain the total demand of the firm, it is necessary to solve the consumption problem of the consumers. Because not all consumers know about all products, consumers will display CES preferences over a limited set of goods. Denoting Ω_i the set of goods in the consideration set of a given consumer i, the utility function is

$$U_i = \left[\int_{\omega\in\Omega_i} \exp\left(\frac{1}{\sigma}\lambda(\omega)\right) q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}} \qquad \sigma > 1,$$

where $q(\omega)$ is the quantity consumed and $\lambda(\omega)$ the appeal of the product. Consequently, for a price $p(\omega)$, the sales to consumer i of a good ω are

$$s_i(\omega) = \begin{cases} \exp(\lambda(\omega))\tilde{p}(\omega)^{1-\sigma}P^{\sigma-1}y & \text{if } \omega \in \Omega_i \\ 0 & \text{if } \omega \notin \Omega_i \end{cases}$$

with P the price index, and y the amount spent by consumer i for this consideration set.[24] Therefore, I obtain the demand for the firm f at time t for destination d as

$$q_{fdt} = q(\lambda_{ft}, X_{dt}, n_{fdt}, p_{fdt}, \varepsilon^D_{fdt}) = n_{fdt}\exp(\lambda_{ft} + X_{dt} + \varepsilon^D_{fdt})p^{-\sigma}_{fdt} \tag{4}$$

where X_{dt} will capture all the aggregate variables of the demand shifter,[25] 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 capture that firms will provide different product appeal in different destinations, as long as these differences are common across firms. This assumption will be fundamental to explain the

[24] Note that by having different sets of goods, each consumer would have 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 have the same price index defined as $P = \left[\int_{\omega\in\Omega} n(\omega)\exp(\lambda(\omega))\tilde{p}(\omega)^{1-\sigma}d\omega\right]^{\frac{1}{1-\sigma}}$.

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

identification 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 that are 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 of all, I assume constant marginal costs of production. These marginal costs are a decreasing function of the productivity of the firm ϕ_{ft}, and will depend on the appeal of the good produced through a parameter α that characterizes the cost elasticity of appeal. Moreover, I assume the existence of non-persistent productivity shocks ε^S_{fdt}, and I allow costs to vary with the destination market by including a set of coefficients γ_d. Formally, the marginal cost function of the firm is

$$c_{fdt} = c(\phi_{ft}, \lambda_{ft}, \varepsilon^S_{fdt}) = \exp(-\phi_{ft} + \alpha\lambda_{ft} + \varepsilon^S_{fdt} + \gamma_d) \tag{5}$$

In addition to these production costs, I will 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

$$FC_d + \nu_{fdt} = \begin{cases} f_d + \nu_{fdt} & \text{if} \quad \mathcal{I}_{fdt-1} = 1 \\ f_d + fe_d + \nu_{fdt} & \text{if} \quad \mathcal{I}_{fdt-1} = 0 \end{cases}$$

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} is a random shock on fixed costs. I will assume that this shock ν_{fdt} will follow a logistic distribution with variance parameter σ_ν. The addition of this shock will allow the model to rationalize all observed decisions made by the firms. Moreover, it is important to note that the amplitude of these fixed costs will vary across destinations. However, I will restrict this variation in the estimation, by allocating each foreign destination to specific groups sharing the same value of fixed costs.[26]

This achieves the definition of the demand and supply characteristics of the firm. I now turn to the definition of the profit and value functions associated to the exporting activity of firm.

3.3 Profit and value function

From the demand received by the firm, and the costs associated with production, I derive the potential profit of the firm for each destination market. After defining the timing of a typical period, I can define the entry problem of the firm, and the associated value functions. This

[26] For instance I will assume that entry and per-period fixed costs will be similar for all European countries. Morales, Sheu, and Zahler (2014) develop a specific empirical procedure that allow them to flexibly estimate entry and fixed costs across destinations.

dynamic problem will depend on five variables that will define the state space of the problem: the exogenous variables, that gathers product appeal λ, productivity ϕ and aggregate demand X, the share of consumer n, and the presence on the market in the previous year $\mathcal{I}_{-1}$.

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

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

where $\mathcal{I}_{-1}$ is a dummy equal to one if the firm was selling on the market in the previous year. This profit function is made of a variable profit and fixed costs. Despite having CES preferences, 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 $FC(\mathcal{I}_{-1})$ that will depend on the past presence of the firm on the market. Finally, the profit shock ν will allow the empirical model to explain the entry and exit decisions of firms that cannot be rationalized by the values of the variable profit and fixed costs.

However, this profit will only be obtained by the firm if it decides to be active on the market at this period. In order to study the problem of the firm, it is necessary to define the timeline of a typical period. This timeline should provide the timing at which decisions are taken and the information sets available to the firms when they take these decisions. Figure 4 displays the timeline of a period that defines the dynamic problem of the firm.

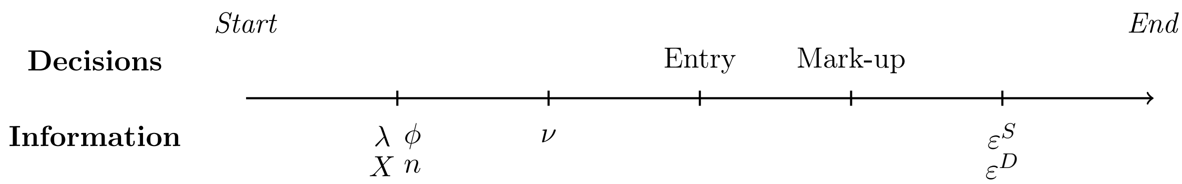

Figure 4: Timeline of one period

As described in figure 4, the firm observes at the beginning of the period its exogenous variables, λ, ϕ, n and X. After realization of the profit shock ν, they decide to be active or not on the market. If they decide to enter, they will not decide the price to be charged but the mark-up to apply to their costs.[27] Finally, sales and prices will be obtained after observing the realization of the non-persistent shocks ε.[28]

[27] While being very similar decisions, choosing the mark-up rather the price facilitates the computation of the solution, while allowing for structural shocks ε in demand and costs.

[28] The assumptions made regarding the timing of the shocks and decisions are mostly driven by the construction of the empirical model. The realization of the shock ν before the entry decisions allow the model to rationalize

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:

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

The first line describes the entry problem, in which the firm chooses between being active $V_I(\xi, n) - FC(\mathcal{I}_{-1})$ and inactive $V_O(\xi)$. By being inactive, the firm makes no profit today but keep the possibility to update its decision in the next period. By being active, it obtains a present profit that will depend on the shocks ε and the mark-up decided by the firm. Moreover, the firm will have 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 will be constructed from the transition of the exogenous variables $F(\xi'|\xi)$, and the expected value of $V(\xi, n', I)$.

In order to solve this problem, it is necessary to proceed through backward induction by describing the pricing decision made by the firm once it enters. This optimal pricing decision leads to the expected profit of the firm, and therefore solve for the entry decisions. I describe these optimal decisions and the value functions of the problem in the next subsection.

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 will deviate from a standard constant mark-up. Instead, firms will 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 will infer the expected profit of the firm conditional on this optimal pricing decision, and therefore infer the value and probability of exporting.

Optimal price The choice of the mark-up of the firm is not standard since it involves solving a dynamic problem: by affecting the sales of the firm today, the price charged by the firm affect the share of consumers tomorrow. Therefore, the firm will have incentives to reduce its price today to foster the accumulation of future consumers.

The choice of mark-up of the firm is made after entry, in order to maximize the sum of the present profit and the continuation value of exporting. Formally, the problem and first-order

entry decisions that couldn't be explained otherwise. Similarly, the realizations of the shocks ε after the markup decisions generate structural errors in the sales and prices equations that can explain sales and prices variations.

conditions are the following:

$$V_I(\xi,n) = \max_{\mu}\Big\{E_\varepsilon\Big\{\pi(\xi,n,\mu,\varepsilon) + \beta EV'(\xi,n'(\xi,n,\varepsilon,\mu),1)\Big\}\Big\}$$
$$\Longrightarrow \qquad E_\varepsilon\Big\{\frac{\partial\pi(\xi,n,\mu,\varepsilon)}{\partial\mu} + \beta\frac{\partial n'}{\partial\mu}\frac{\partial EV'(\xi,n',1)}{\partial n'}\Big\} = 0$$

Therefore, the optimal price of the firm is:

$$p(\xi,n) = \mu(\xi,n)c(\xi,n) \tag{6}$$
$$\text{with} \qquad \mu(\xi,n) = \frac{\sigma}{\sigma-1}\frac{1}{1+E_{w(\varepsilon)}\beta\eta_1(1-n')^{\psi}\frac{\partial EV'(\xi,n',1)}{\partial n'}}$$

The optimal mark-up charged by the firm has two components. First, the firm will apply the standard CES mark-up $\frac{\sigma}{\sigma-1}$ based on the price-elasticity of the demand. Second, the firm will apply a discount factor based on the dynamic incentives it has to lower its price to attract more consumers in the future. This factor will depend on two elements: first, how much this increase in sales will increase its consumer share tomorrow, $\eta_1(1-n')^{\psi}$; this element will induce lower mark-ups for small or young firms that benefit from higher returns of accumulation. Second, the extent of this discount will also depend 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 will not be linear but hump shaped with the profitability of the firm:[29] young firms that are unlikely to survive will not have incentives to invest in future consumers. Firms that can use extra consumers to increase their probability of survival will get the largest benefits from increasing their consumer share. However, because of the concavity of the value function conditional on surviving, this effect will be smaller for high profit firms that are likely to survive in the next period. 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' will depend on the price charged.[30]

Consequently, the accumulation of consumer will imply heterogeneous mark-ups by the firms, depending on their current share of consumers, and their expectations on future profits. Having described the optimal mark-up of the firm, it is possible to infer the expected profit of the firm in case of entry. Therefore, I can evaluate the two options of the firm, and study its entry decision.

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

[30]Note that

$$E_{w(\varepsilon)}\beta\eta_1(1-n')^{\psi}\frac{\partial EV'(\xi,n',1)}{\partial n'} \equiv \int_\varepsilon c(\xi,\varepsilon)q(\xi,n,\mu,\varepsilon)\beta\eta_1(1-n')^{\psi}\frac{\partial EV'(\xi,n',1)}{\partial n'}dF(\varepsilon)$$

To overcome the absence of closed form solution for the optimal price, I will use a grid to solve the optimal price of the firm in the estimation procedure. Moreover, solving the dynamic problem of the firm will also be facilitated by assuming that $E_\varepsilon EV'(\xi,n'(\xi,n,\varepsilon,\mu),1) = EV'(\xi,n'(\xi,n,E_\varepsilon\varepsilon,\mu),1)$. This assumption will allow me to redefine the problem such that

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

for which $E_\varepsilon\pi(\xi,n,\mu,\varepsilon)$ admits a closed-form solution that will facilitate the evaluation of the model.

Entry condition Knowing the expected option values of being active or inactive, I can now study the entry decision of the firm. The firm will pick the most profitable option, after observing the shock ν that affects the fixed costs of being active on a market. The logistic assumption for this shock will generate a closed-form solution for the probability of entry, but also for the expected value function before observing this shock. Formally, the expected value of the firm before observing the shock ν is

$$
\begin{aligned}
V(\xi, n, \mathcal{I}_{-1}) &= E_\nu \max \left\{ V_I(\xi, n) - FC(\mathcal{I}_{-1}) - \nu \;\; ; \;\; V_O(\xi) \right\} \\
&= \sigma_\nu \log \left[\exp\left(\frac{1}{\sigma_\nu} \big(V_I(\xi, n) - FC(\mathcal{I}_{-1})\big) \right) + \exp\left(\frac{1}{\sigma_\nu} V_O(\xi) \right) \right].
\end{aligned}
$$

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,

$$
P(\mathcal{I} = 1 | \xi, n, \mathcal{I}_{-1}) = \frac{1}{1 + \exp\left(-\frac{1}{\sigma_\nu} \left(DV(\xi, n) - FC(\mathcal{I}_{-1}) \right) \right)} \tag{7}
$$

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

In order to close the definition of the dynamic problem of the firm, I need to specify the evolution of the exogenous variables of the model. These exogenous variables will be important since they can account for a large amount of the persistence observed in the data. Most of the hysteresis in exporting decisions is likely to come from the persistence over time of fundamental characteristics of the firm such as productivity or product appeal. Therefore, it is necessary to allow these processes to be quite general in their persistence. Moreover, to account for the important attrition rate across ages, it is also necessary to let these processes vary across time, through random shocks. Consequently, one wants to assume general processes that are time variant, and allow for important persistence in their evolution. For these reasons, I will assume that these three variables will follow AR(1) processes, with flexible parameters. Formally, I assume

$$
\begin{aligned}
\lambda_{ft} &= \rho_\lambda \lambda_{ft-1} + \sigma_\lambda \varepsilon^\lambda_{ft} \\
\phi_{ft} &= \mu_\phi + \rho_\phi \phi_{ft-1} + \sigma_\phi \varepsilon^\phi_{ft} \\
X_{dt} &= \mu_{Xd} + \rho_X X_{dt-1} + \sigma_X \varepsilon^X_{dt}
\end{aligned}
\tag{8}
$$

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. Moreover, because X_{dt} describes the aggregate demand from a destination d, I allow the mean μ_{X_d} of this process to change across destination. This will allow the model to capture different trends in aggregate demand across different destinations.

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. Consequently, we have

$$\begin{aligned} \lambda_{f0} &\sim N\left(0, \frac{\sigma_\lambda}{\sqrt{(1-\rho_\lambda^2)}}\right) \\ \phi_{f0} &\sim N\left(\frac{\mu_\phi}{1-\rho_\phi}, \frac{\sigma_\phi}{\sqrt{(1-\rho_\phi^2)}}\right) \end{aligned} \tag{9}$$

However, I will assume that the variation in aggregate demand across destination do not arise from a stationary distribution. Therefore, I will assume a flexible distribution of initial conditions for X_{d0} such as

$$X_{d0} \sim N(\mu_{X_0}, \sigma_{X_0}). \tag{10}$$

Moreover, I will assume that the initial share of consumers, which will apply to firms that records positive sales the year before the beginning of the model, follow a Beta distribution with parameters 1 and 5.[31]

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, the firm decides to enter or exit various destinations where it decides at which prices to sell its good. The more the firm sells on a market, the more consumers will be ready to buy from it in the next period, fostering its demand and profit in the next period. After describing the model, I now describe the restrictions I impose to obtain a model without consumer accumulation, that will behave similarly to standard models used in the literature.

3.6 Restricted model

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

[31] Given the number of firms in this case, and the length of the panel I will use (14 periods), this assumption has no consequences on the estimation.

exporting. This model can be seen as a dynamic version of Melitz (2003), as estimated by Das, Roberts, and Tybout (2007). Estimating this restricted model will be essential to assess the importance of the accumulation of consumers on the outcomes of the estimation and the aggregate implications of the model.

4 Estimation

In this section, I describe the procedure used to estimate the parameters of the model. The likelihood is directly obtained from the three structural equations of the model. However, the evaluation of this likelihood is made cumbersome by the number of persistent and unobservables variables and the dynamic problem of the firm.

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 to show the advantages of a MCMC estimator to facilitate the estimation of the model. Finally, I provide the intuition behind the identification of the 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 demand equation in which the stock of consumers of the firm appears, the pricing equation that features 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 (4), (5) and (6) are taken in logarithm to obtain

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

This block will constitute the first part of the likelihood. Assuming that ε follow a bivariate normal distribution with variance Σ, I define this likelihood block as $L_\varepsilon(s_{fdt}, p_{fdt}|\xi_{fdt}, n_{fdt}; \Theta)$,[32] with Θ being the full set of parameters, such that

$$\begin{aligned}L_\varepsilon(s_{fdt}, p_{fdt}|\xi_{fdt}, n_{fdt}; \Theta) = G_\Sigma\Big(&\log s_{fdt} - \log n_{fdt} - \lambda_{ft} - X_{dt} - (1-\sigma)\log p_{fdt}\ ; \\ &\log p_{fdt} + \phi_{ft} - \alpha\lambda_{ft} - \log\mu(\xi, n) - \gamma_d\Big)\end{aligned} \quad (11)$$

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 will be based on the entry decision of the firm. Equation (7) 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 $L_\nu(I_{fdt}|\xi_{fdt}, n_{fdt}, I_{fdt-1}; \Theta)$

[32] As previously defined, ξ_{fdt} gathers all the exogenous variable of the model, product appeal, productivity and aggregate demand, such that $\xi_{fdt} \equiv \{\lambda_{ft}, \phi_{ft}, X_{dt}\}$

that is obtained from the binary choice made by the firm

$$L_\nu(\mathcal{I}_{fdt}|\xi_{fdt}, n_{fdt}, \mathcal{I}_{fdt-1}; \Theta) = \left[1 + \exp\left(\frac{-DV(\xi_{fdt}, n_{fdt}) + FC(\mathcal{I}_{fdt-1})}{\sigma_\nu}\right)\right]^{-\mathcal{I}_{fdt}} \times \left[1 + \exp\left(\frac{DV(\xi_{fdt}, n_{fdt}) - FC(\mathcal{I}_{fdt-1})}{\sigma_\nu}\right)\right]^{\mathcal{I}_{fdt}-1} \tag{12}$$

where function $DV(\xi_{fdt}, n_{fdt})$ and $FC(\mathcal{I}_{fdt-1})$ are defined as previously. Therefore the total likelihood for a given observation $D_{fdt} \equiv \{s_{fdt}, p_{fdt}, \mathcal{I}_{fdt}\}$ is

$$L(D_{fdt}|D_{fdt-1}, \xi_{fdt}, n_{fdt-1}; \Theta) = L_\nu(\mathcal{I}_{fdt}|\xi_{fdt}, n_{fdt}, \mathcal{I}_{fdt-1}; \Theta) \times L_\varepsilon(s_{fdt}, p_{fdt}|\xi_{fdt}, n_{fdt}; \Theta)$$

To obtain the unconditional likelihood, that does not depend on the unobservables of the model, it is necessary to integrate out this set of unobservables. However, because 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. Formally, the full likelihood is

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

where $F(\xi_{fd0})$ is defined by the density of the initial unobservables defined in equations (9) and (10), and $F(n_{fd-1})$ the beta distribution assumed for firms that were exporting the year before the beginning of the estimation sample, 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 by describing the algorithm aiming to find the posterior distribution of parameters Θ.

4.2 Algorithm

To estimate the model, I develop a Markov Chain Monte Carlo (MCMC) estimator to account for two important difficulties in evaluating the likelihood of this problem: the different sources of persistent and unobservable heterogeneity and the dynamic problem of the firm. First of all, the persistent unobservables characteristics make it necessary to perform a large number of integration in order to evaluate the likelihood. This is particularly cumbersome given the persistent nature of these sources of heterogeneity. The second difficulty comes from the need to solve for the value functions in order to obtain the objects $DV()$ and $\mu()$ and evaluate the likelihood. The literature on dynamic discrete choices model, starting from Rust (1987) is mostly devoted to this specific problem: it requires to find the solution of the Bellman equation through value functions iterations until reaching a fixed point.[33] Therefore, even in the absence of unobservables, the likelihood function is a highly non-linear function of the parameter set Θ, increasing the difficulty, and the computing time, of evaluating the likelihood.

In order to circumvent these difficulties, I employ a MCMC estimator, taking advantage of

[33]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 where state variables are mostly unobserved, obtaining conditional choice probabilities in a first step is not trivial, and likely to be an imprecise exercise.

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 findings from the Bayesian econometrics literature. First, Arellano and Bonhomme (2009) shows how Bayesian hierarchical models nest fixed and random effects models: using a prior distribution of the unobservable of the model, the posterior distribution of the unobservable term will be very precise when many observations are available (for instance when one firm sells to many destinations), such that this posterior distribution will be close to the fixed effects value. When the number of observations is limited (for instance when a firm only sells to one country), the prior distribution of the unobservable variable, as specified by the model, will constrained the value of this variable similarly to the random-effect case. Moreover, using MCMC in this context will allow to perform the integrations by updating unobservables as latent variables of the model. Therefore, the use of a Bayesian estimator allows to perform a fast integration of these unobservables, while correcting for the first-order bias that exists in fixed and random-effects models.[34]

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, by only updating the value functions in the Bellman equation once at each iteration. The intuition is that there is no need to fully solve for the fixed point of the value function at each point of the parameter set. Instead, it is possible to only iterate the Bellman equations a limited number of times at each iteration of the Markov chain, reusing these value functions as initial values for the next iteration. As the Markov chain converges and explores the posterior distribution of Θ, the value function will also converge toward the fixed point that solves the Bellman equation.

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

$$P(\Theta \,|\, D) \propto \int_\xi L(D \,|\, \xi, \Theta)\, dF(\xi \,|\, \Theta)\, P(\Theta) \tag{13}$$

where $L(D \,|\, \Theta) = \int_\xi L(D \,|\, \xi, \Theta)\, dF(\xi \,|\, \Theta)$ is the likelihood of the problem and $P(\Theta)$ is the prior distribution of the parameter set. Because I do not want the priors to influence the posterior distribution of the parameters, I will assume that all the priors are flat, except for values of parameters that do not satisfy theoretical or stationarity constraints.[35] Therefore, the goal of the Markov Chain is to repeatedly sample from the posterior distribution according to (13). This will be achieved by alternatively sampling parameters conditional on unobservables, and parameters conditional on unobservables. In this specific application, an iteration in the Markov chain is made of three different steps, summarized in the following iteration.

At an iteration s, the inputs of the Markov chain are $\Theta^{(s)}, \xi^{(s)}$ and the history of value functions $\left\{V(\Theta^{(h)})\right\}_{h=s-m}^{s}$ and their associated parameters sets $\left\{\Theta^{(h)}\right\}_{h=s-m}^{s}$ for a given $m \geq 0$. The steps of a typical iterations are

[34] Roberts, Xu, Fan, and Zhang (2012) also uses this type of estimator in a similar context. The main difference being that the unobservables terms are time-invariant in their model while they vary in mine, making the integration issue even more stringent in my setup.

[35] 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. Finally, I also impose the average fixed costs and entry costs parameters (f, fe) to be positive. and the parameter ψ to be larger than zero.

- Sample $\xi^{(s+1)}$ proportionally to $L(D|\xi, \Theta^{(s)}) f(\xi|\Theta^{(s)})$
- Sample $\Theta^{(s+1)}$ proportionally to $L(D|\xi^{(s+1)}, \Theta) f(\xi^{(s+1)}|\Theta) P(\Theta)$
- Update $\{V(\Theta^{(h)}), \Theta^{(h)}\}_{h=s+1-m}^{s+1}$ using $\Theta^{(s+1)}$ and $V(\Theta^{(s+1)})$.

Two important points are worth noticing regarding this algorithm. First of all, the large size of the parameter space requires to update the parameters sequentially rather than simultaneously. In total, 30 parameters will be estimated in the model. Consequently, a Gibbs sampling is used in which different parameters blocks are created and sequentially updated based on the different blocks of the likelihood.[36] Second, the value functions that allow to compute the objects $DV(.)$ and $\mu(.)$ will be obtained on a grid that will be updated throughout the algorithm. The specific values of $DV(.)$ and $\mu(.)$ will then be obtained by interpolation to be evaluated at any point of the state space. I provide extensive details in appendix C about the implementation of the algorithm.

Due to the complexity of the estimation procedure, two parameters will not be estimated and set to specific values from the literature. First of all, I do not estimate the value of β, the discount rate of future periods. This parameter is difficult to identify in dynamic discrete choice models and I therefore set its value to 0.9, following Das, Roberts, and Tybout (2007).[37] Second, I do not estimate the elasticity of substitution of the CES utility function. Estimating the price-elasticity of demand using trade data is not trivial given the absence of product characteristics, which implies unobserved vertical differentiation across good.[38] Therefore, I will use the value obtained by Broda and Weinstein (2006) for the corresponding industry: they estimate an elasticity of 2.2 for the wine industry that I will use and keep constant throughout the algorithm.

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

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 the identification of the parameters. Moreover, the alternative sampling of unobservables and parameters shed light on the separate sources of identifications of each component of the likelihood.

To describe the sources of identification, it is important to distinguish the identification of unobservables and parameters. Let's assume first that the parameters of the model are known. In this situation, the identification of the unobservables mostly come from a variance decomposition of the demand shifters and prices. Indeed, knowing sales 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

[36] Despite the separation of the parameters in different sets, the existence of value functions in the likelihood creates a dependence between most parameters of the parameter set and the different part of the posterior distribution. Therefore, Metropolis-Hastings algorithms are used to sequentially update these different blocks.

[37] Magnac and Thesmar (2002) provides an extensive discussion of identification issues in DDCM.

[38] See Piveteau and Smagghue (2015) for a discussion on the estimation of this elasticity. In theory, prices in other destinations could be used as instrument for the prices. However, this requires to control for the impact of quality on marginal costs, which is part of the model (through the parameter α).

n_{fdt}). Once the product appeal is known, the productivity ϕ_{ft} is identified by price variations across firms. Therefore, the identification of the unobservables mostly comes from a decomposition of observables variables, that is straightforward once the parameters of the model are known. Moreover, the hierarchical structure and the entry decisions will 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 year will help identifying the potential value of the unobservables. Similarly, the entry decisions on foreign destinations will bring additional information on these unobservables: if a firm only exports to one destination at a given year, the fact that it does not export somewhere else will bring some information relatively to the latent value of its appeal or productivity.

Let's now turn to the identification of the parameters of the model, assuming that the unobservables are known. The 30 estimated parameters can be divided in three groups: 17 of them are related to the laws of motion of the unobservables, 6 to the demand and supply equations, and 7 related to the dynamic problem of the firm. Knowing the unobservables of the problem, the identification of the parameters that describes their distribution and law of motions is straightforward. Regarding the parameters that are linked to the demand and pricing functions, their identification is similar to a regression of prices on destination dummies and the appeal of the product, while the parameters of the variance matrix is obtained from the variance of the unexplained variation in prices and sales. Finally, the parameters related to the entry problem of the firm are obtained by comparing potential profits and firms' observed decisions. Based on the characteristics of the firms and destinations, the laws of motion of unobservables, and the parameters of the cost and demand functions, it is possible to construct the potential profit of each firm on each market. Based on these potential profits, the number of exporters will identify the per-period fixed costs, the persistence in exporting the entry costs and the remaining variance in exporting decisions will identify the required variance of these fixed costs shocks.

Consequently, the identification of the unobservables conditional to the parameters, and of the unobservables 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 will describe the posterior distributions of the parameters.

5 Results

I implement my estimation on a set of wine exporters from France. The choice of the industry is based on two criteria. First, wine producers only export wine. Therefore, it is reasonable to assume that the entry decisions on foreign destinations are made at the firm level, and it is possible to aggregate sales and prices at the level of the firm 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 relatively extended set of destinations. In appendix A.2, I describe the specific selection procedure to obtain the estimation sample of 200 firms, and provide statistics to describe this sample.

In order to describe the results of the estimation, I start by describing the fit of the model relative to the exporters' dynamics presented earlier. Then I will present the estimated values of the parameters, and in particular the decrease in entry costs induced by the introduction of the consumer margin. Finally, I will describe the evolution of the consumer margin and the mark-ups charged by firms at different export ages.

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 5 reports the predictions of the model relatively to the data. I also reports the results of the restricted version of the model, that does not contain a consumer margin.

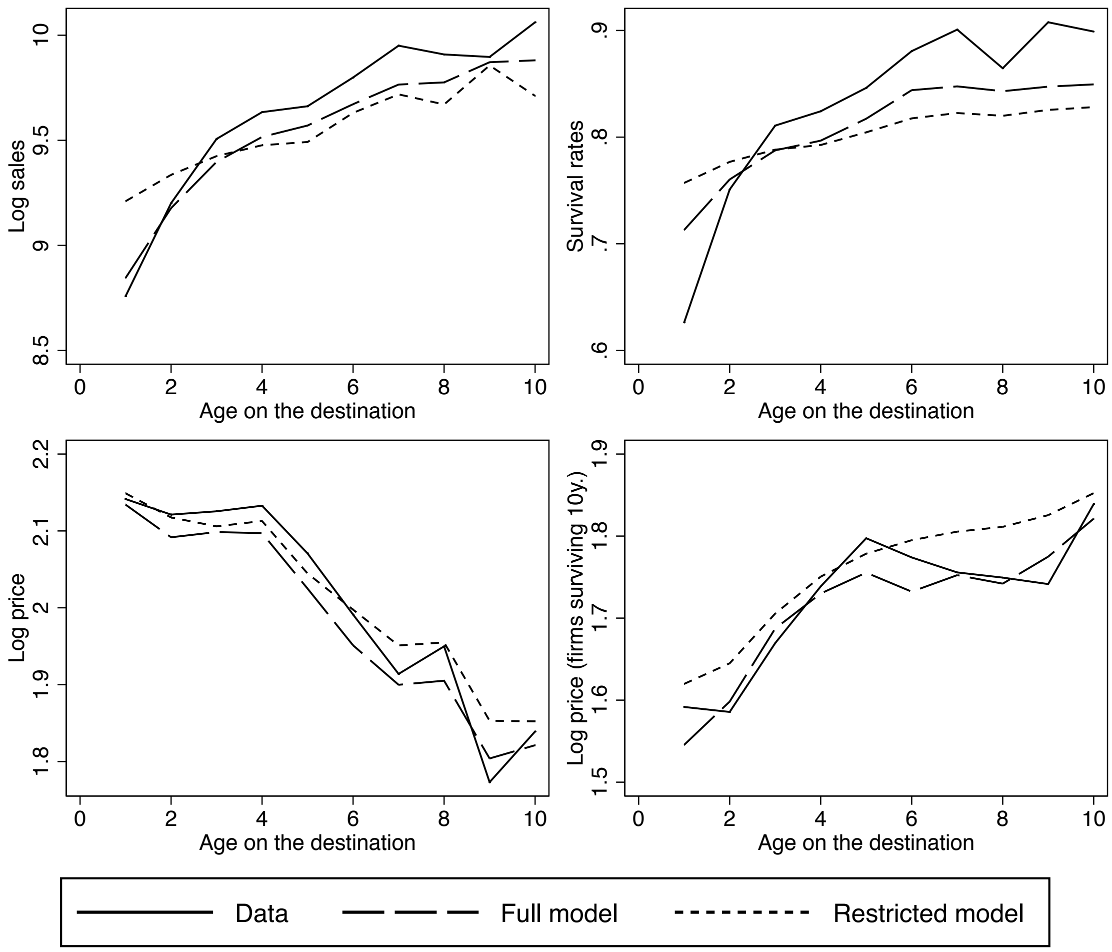

FIGURE 5: Predictions of survival rates, sales and prices across ages.

As reported in figure 5, the full model with consumer accumulation can reproduce most of the growth in sales across ages (top left figure). The ability of the model to capture this growth explains how the model can perform better in terms of survival rates (top right figure): as firm accumulates more consumers on a foreign destination, raising its sales, it also increases its future profit, and therefore its survival rates. However, this growth in sales is not sufficient to fully explain the low survival rates of young exporters, and therefore does not entirely solve the puzzle linked with young exporters dynamics. 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 constant across ages, between 75 and 80 %, which is similar to the average survival rate in the sample. However, the predictions on prices appear quite similar across models (bottom figures). Both of them can reproduce the decrease in prices with ages. When looking at firms surviving 10 years, we can see that the full model can do slightly better in explaining the rise in prices with age. Therefore, the heterogeneous mark-ups obtained from the dynamic problem of the firm seems to help the model in predicting low prices at young age.

After describing the fit of the model, I 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 mean of its posterior distribution, as well as its 90% confidence interval.

First of all, looking at the law of motion of the consumer margin, we note that the initial of consumers at entry (n_0) is relatively small, equal to 3%, which leaves a large potential for firms to grow through the accumulation of consumers. This growth is driven both by the past sales of the firm (η_1), as well as the past shares of consumers (η_2), since the two coefficients are significantly larger than zero. Moreover, we can see that the degree of concavity of this law of motion is significant, with a mean of the posterior distribution of the coefficient ψ equal to 0.44.

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 0.98, 0.93 and 0.93, for respectively the appeal and the productivity of the firm, and the aggregate demand of the destination. Moreover, the appeal appears to have a larger variance across firms ($\frac{0.19}{\sqrt{1-0.98^2}} = 0.95$) than productivity ($\frac{0.09}{\sqrt{1-0.93^2}} = 0.24$). If this is not surprising, given that sales have a larger variances than prices, it is interesting to look at the implied contribution of these two unobservables variables to sales. With a parameter of the cost of appeal α equal to 0.73, it means that an extra unit of appeal has an impact of 11 percents $(1 - 0.74 \times 1.2)$ on sales, which is to compare with an increase of 100 percents from productivity. Consequently, moving from the average appeal to the 5th best percentile increases the sales by 17%, while the same movement for productivity increases sales by 39%.

Finally, because I estimate a structural model of entry, the model is able to deliver euros estimates of the sunk fixed costs of entry as well as the per-period fixed costs paid by an exporter.[39] We see that the obtained fixed costs are relatively low, with the entry cost to an European destination being equal to 33 730 euros.[40] In addition, a firm would have to pay 8 000 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

[39] I separated my destinations in three groups such that each European destination will have similar fixed costs. This does not imply that the firm do not need to pay these costs for each destination it enters. If a firm exports to 5 European destinations, it will have to pay 5 times these fixed costs.

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

TABLE 2: Estimated parameters

Parameter		**Estimate**	**90% Confidence Interval**	
			Lower bound	*Upper bound*
Per-period fixed costs	Europe	7 994	6 761	9 194
(in 2000 euros)	Americas	7 495	6 693	8 304
	Asia/Oceania	8 019	7 080	8 930
Entry fixed costs	Europe	33 730	30 303	37 078
(in 2000 euros)	Americas	23 656	21 092	26 208
	Asia/Oceania	28 619	25 387	31 928
Variance of entry shocks	σ_ν	9 656	8 589	10 620
Law of motion of n	n_0	0.033	0.031	0.034
	$\underline{n}$	0.015	0.014	0.016
	$\eta_1 (10^{-5})$	0.12	0.11	0.14
	η_2	0.27	0.23	0.29
	ψ	0.44	0.00	0.93
Law of motion of appeal	ρ_λ	0.98	0.98	0.98
	σ_λ	0.19	0.18	0.20
Law of motion of productivity	ρ_ψ	0.93	0.91	0.94
	σ_ψ	0.09	0.08	0.09
	μ_ψ	-0.12	-0.14	-0.10
Law of motion of agg. demand	ρ_X	0.93	0.93	0.94
	σ_X	0.09	0.09	0.09
	μ_{X1}	0.98	0.91	1.03
	μ_{X2}	0.88	0.74	0.97
	μ_{X3}	0.89	0.77	0.97
	μ_{X_0}	14.58	14.31	14.83
	σ_{X_0}	0.46	0.32	0.65
Elasticity cost of appeal	α	0.73	0.73	0.74
Cost dummies	γ_2	0.38	0.36	0.39
	γ_3	0.30	0.29	0.30
Variance matrix	Σ_{11}	1.25	1.25	1.26
	Σ_{12}	0.17	0.17	0.17
	Σ_{22}	0.56	0.54	0.57

is 13 000. One of the reasons for these relatively low numbers is the small variance parameter of these fixed costs shocks, whose the average of the posterior distribution is 9 656. This low number reflects the ability of the model to correctly predict the entry and exit of firms, such that a large variance of these fixed costs shocks, that allow the model to rationalize all decisions, is not necessary.

In order to confirm the small magnitudes of these entry fixed costs relatively to the literature, I compare theses parameters with the ones I obtain when estimating the restricted version of the model, that does not have a consumer margin. Results are displayed in table 3.

The comparison between two models highlights that the entry costs, and more generally the fixed costs of exporting, are much larger in the version without consumer margin. For instance, the average entry costs to export to Europe jump from 33 730 to 98 286 euros. Part of this increase comes from the fact that the parameter of variance of the fixed costs goes from 9 656 to

Table 3: Estimated parameters (comparison between models)

Parameter		Full model Estimate	Full model 90% C.I. Lower	Full model 90% C.I. Upper	Restricted model Estimate	Restricted model 90% C.I. Lower	Restricted model 90% C.I. Upper
Per-period fixed costs	Europe	7 994	6 761	9 194	8 521	7 989	9 080
	Americas	7 495	6 693	8 304	14 605	13 429	15 810
	Asia/Oceania	8 019	7 080	8 930	16 133	14 531	17 997
Entry fixed costs	Europe	33 730	30 303	37 078	98 286	87 044	110 368
	Americas	23 656	21 092	26 208	72 073	63 372	81 393
	Asia/Oceania	28 619	25 387	31 928	80 951	71 094	91 913
Elasticity cost of appeal	α	0.73	0.73	0.74	0.39	0.37	0.41
Variance of entry shocks	σ_ν	9 656	8 589	10 620	25 789	23 121	28 703

25 789. This increase reflects the fact that the consumer margin improve the ability of the model to explain entry and exit decisions. But this reduction in average entry costs, when introducing this consumer margin, is not only due to this smaller variance, but also characterize an important change in the relative role played by entry and per-period costs: while the ratio between entry costs and per-period costs is of an order 5 to 10 in the restricted model, it is only 3 to 5 in the full model. This reflects that the introduction of the consumer margin captures an important amount of state dependence, reducing the role played by entry costs in explaining the hysteresis in the export decision. This result will be very important when looking at models' predictions in response to shocks. Estimating large entry costs to export implies that the option value of exporting is very large: the large average entry costs make entering so difficult that firms will not easily decide to exit this export market. I will study these consequences in the next section when comparing the predictions of these models under simulated and observed trade shocks.

Another important difference between these two models comes from the estimates of the cost of appeal. In the full model with consumer margin, appeal is very costly, making high-appeal products barely more profitable than low-appeal ones.[41] However, the model without consumer margin features appeal that has a low impact on costs, with an average estimate of 0.27. This difference is interesting because it describes how the introduction of consumer margin, affects the definition of appeal itself. When appeal is the unique demand shifter, it will capture the role of distribution network for instance and other characteristics that raise the sales of the firm conditional on prices. However, with the introduction of a consumer margin, part of this sales variation will be captured by this new margin, such that what the full model will infer as appeal will be more related to the type of good produced, and its characteristics. As a consequence, the appeal inferred in the full model is closer to what we could refer as product quality, which would explain its larger impact on the marginal costs of production.

[41] In this model, appeal is exogenous and therefore could have a negative impact on sales and profit. This would be the case if $\alpha > \frac{1}{\sigma-1} \approx 0.83$.

5.3 Outcomes of the model

Finally, to conclude the description of the results, I 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 6 provides the distribution of consumer shares for each age of the firm. Remember that when firm enter, they all have an initial share $n_0 \approx 3\%$, which explains why the graphs provides distributions from ages 2 to 10. The main information from this figure comes from the fact that the distribution tends to shift toward the right as age increase. One can see that most of the firms have a small consumer share at age 2: only a small fraction of them are larger than 25%. 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%. However, there is still a large amount of heterogeneity across ages. Some firms are large at ages 2 or 3, and 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 appears to flatten as age increases, rather than translate toward the left. This implies that the process of consumer accumulation does not appear to be identical across firms, and relies very much 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.

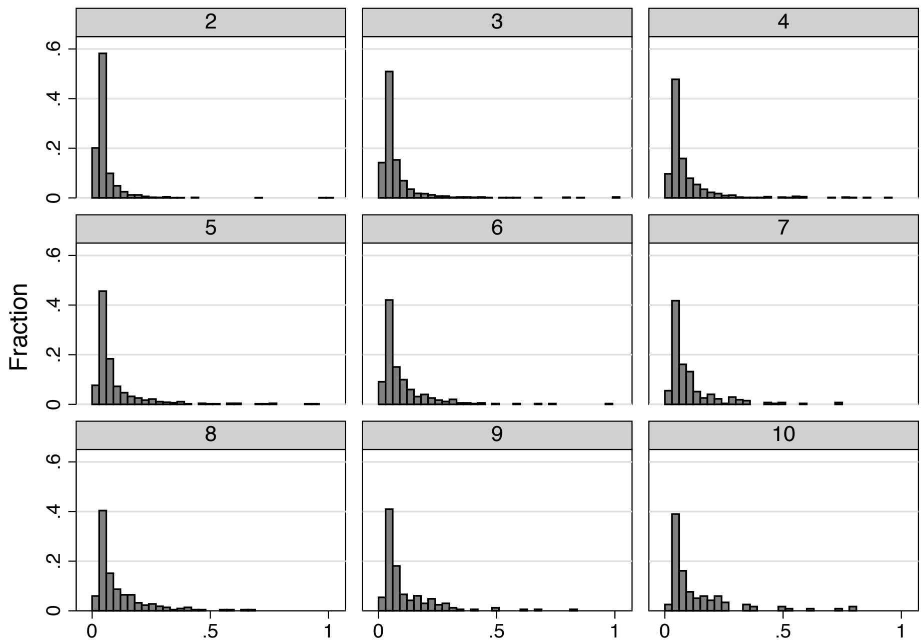

FIGURE 6: Distribution of consumer shares by age

After describing the evolution of the distribution of consumer shares, I turn to the distributions of the mark-ups charged by the firms. These mark-ups were the only tool for the firm to foster accumulation. Figure 7 reports the distributions of mark-ups, separately for each age from 1 to 9. Moreover, I report in red on these histograms, the CES mark-up in the absence of dynamic pricing ($\frac{\sigma}{\sigma-1}$): because of the dynamic benefits of charging low-markups, firms optimally charge a mark-up that is lower than the CES mark-up (as this is implied by the model). One can see that, similarly 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 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, they are likely to not survive in the following years. Therefore, it is optimal to charge low prices because these new consumers increase their probability of survival: indeed, survival rates tend to increase especially in the early years of exports. Finally, we can see that these dynamic incentives are so large, that some firms are willing to make negative profit during the current period, in order to invest in future consumers: a significant number of firms charge a mark-up that is lower than one, implying a price below marginal costs.

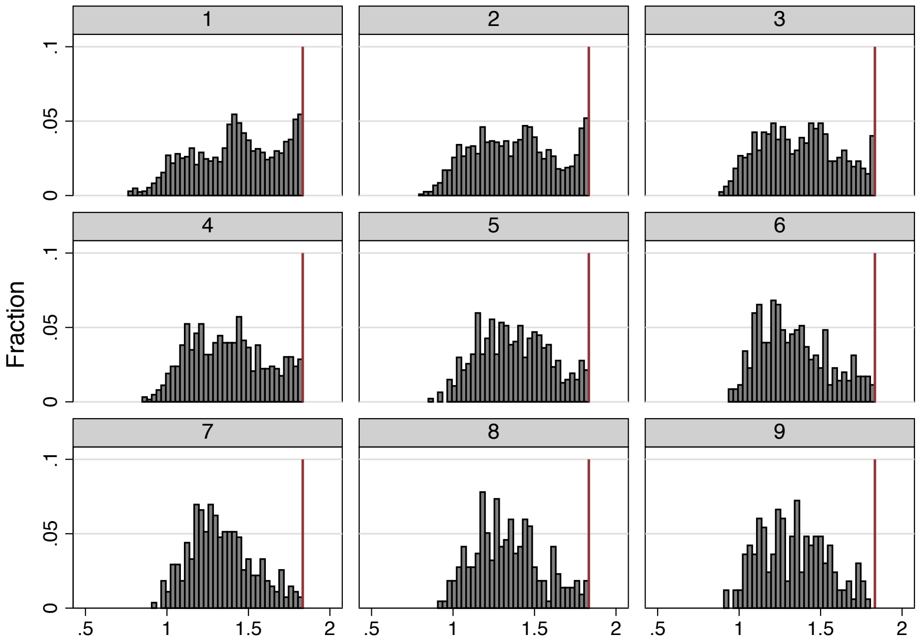

Figure 7: Distribution of mark-ups by age

6 Aggregate implications

In this section, I use simulations and out-of-sample predictions to demonstrate the importance of the model regarding the 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 two important facts regarding aggregate adjustments to trade shocks: first, in response to a positive trade shocks, it will take time for aggregate trade to fully respond, generating a discrepancy between the short and long run trade elasticities. Second, the relative contribution of the extensive margin in this response will be increasing across time, as it has been recently documented in the literature. Finally, I directly test the performance of the model with an out-of-sample predictions exercise. I show that the model can better predict the actual trade response to exchange rate movements that took place during the sample period in the Brazilian market.

6.1 Sluggish trade response

The accumulation of consumers by the firms will generate frictions in growing on foreign markets. As a consequence, the trade response to shocks will be slow at the microeconomic and aggregate level. This pattern, which has been documented in the literature,[42] can explain the discrepancy that exists between the 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, in order to explain variations in trade flows across countries, or trade responses after a trade liberalization episode.[43]

In order to quantitatively evaluate the ability of the model to generate this discrepancy between horizons, I simulate a decrease of 10 points on the tariff applied to export from French firms to the US. I simulate the trajectories of the 200 firms from my sample following this tariff reduction, and compare them to a counterfactual scenario without tariff decrease. I apply this experiment to the full model, as well as the standard model that does not feature consumer accumulation. Figure 8 reports for each model, the log-deviation relatively to the counterfactual scenario without tariff change, of the total trade to the US.

As we can see from figure 8, 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. Moreover, new exporters enter the market such that the trade response is larger than the only sales response to the price decrease. After these first years, no further adjustment occurs. In comparison, the model with consumer margin depicts a slower adjustment to trade as it takes up to 10 years to observe the full effect of the reduction in tariff. The reason for this slow adjustment is that it takes time for existing and new exporters to reach their optimal number of consumers. As a consequence, we see a similar adjustment than the restricted model in the first year, because firms also benefit from lower prices, but this effect is magnified by the increase of the consumer shares of existing firms, as well as the entry of new firms that will grow in the subsequent years. Consequently, the full effect of the tariff reduction will be roughly 3 times the effect recorded after one year. Interestingly, this ratio between long-run and short-run elasticities is roughly consistent with the ratio of elasticities used in the two distinct literatures. As a conclusion, it appears that the model with consumer margin can generate this discrepancy, unlike the standard model that does not feature this margin.

[42]See Alessandria et al. (2013) for instance

[43]See Ruhl (2008) that explains this international elasticity puzzle from the different impacts of permanent and temporary trade shocks.

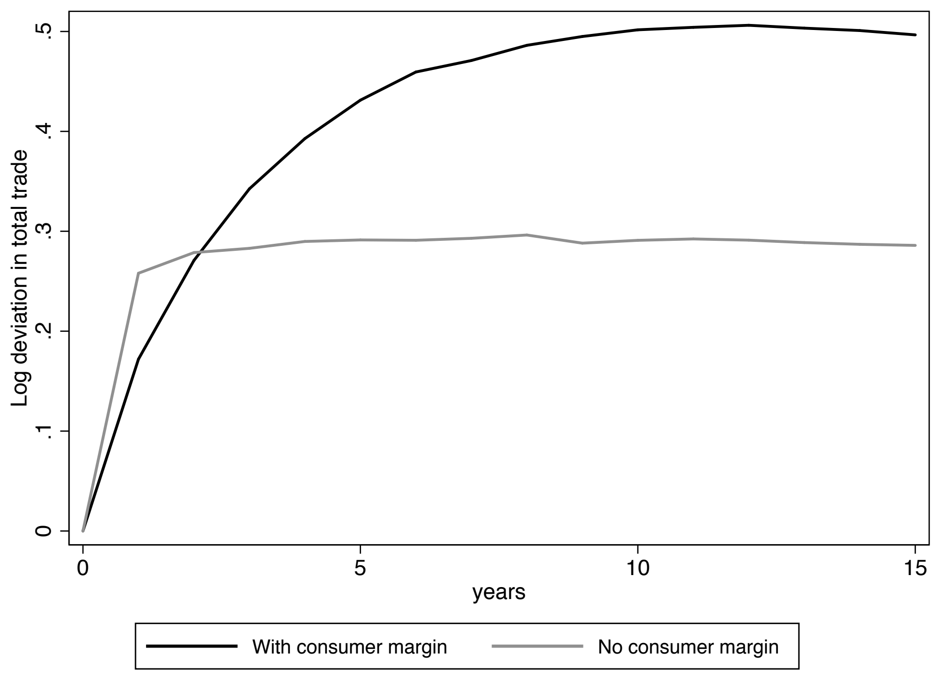

FIGURE 8: Effect of permanent 10 points tariffs decrease.

6.2 Contribution of the extensive margin

A second implication of the model with consumer margin is related to the contribution of the extensive margin to the growth in trade throughout a trade liberalization episode. A number of recent papers documents the increasing contribution of new exporters or new goods at different time horizons: the contribution of the extensive margin is small right after a shock, but tend to increase in the following years to reach a significant contribution in the overall effect. For instance, Kehoe and Ruhl (2013) document this pattern for the contribution of new goods to the trade expansion following the North American Free Trade Agreement (NAFTA). Closer to my empirical application, Alessandria et al. (2013) provide similar evidence when looking at the extensive margin defined at the firm-destination level. In particular, they show that following a devaluation, the contribution of the extensive margin is almost zero in the first quarters after the shock, but can reach 50% of the total trade growth after 5 years.

I explore the predictions of my model, by decomposing the growth of trade following a decrease in tariff. I implement a tariff reduction similar to the previous section, and decompose the total growth in trade following the methodology by Hummels and Klenow (2005): this method allows to measure the contribution of each variable entering the demand function of the firm (intensive margin), and the contribution of new entrants (extensive margin). In my context, I am able to obtain 5 sources of growth: product appeal, consumer margin, prices, aggregate demand that constitute the intensive margin, and the extensive margin. In figure 9 I report the contribution of the aggregate demand (that captures the decrease in tariff), the consumer and the extensive margins along different time horizons.[44]

[44]The other margins being insignificant, I choose to not report them for clarity. The decomposition between all the margins are displayed in figure 19 in appendix D.

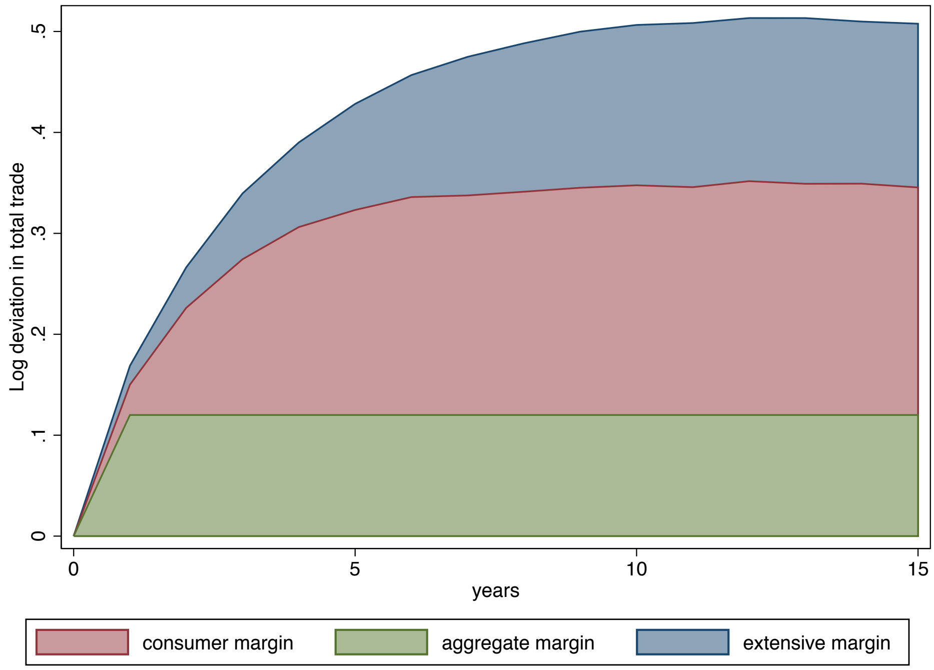

FIGURE 9: Effect of permanent 10 points tariffs decrease.

Figure 9 depicts the increasing contribution of the extensive margin. The first year after the shock, this contribution is very small, around 10% of a small increase in trade. However, as the horizon increases, this contribution is significantly larger, to reach up to 32% of the total growth in trade.[45] There are two important reasons to explain this increasing contribution. First, because of small entry costs, the response of the extensive margin is large: a small decrease in tariff leads to significant entry of new firms on the export market. However, even though the number of these entrants is large, these exporters enter very small, and therefore do no contribute very much to aggregate trade. But as they survive on the market, and increase their stock of consumers, they become large exporters and significantly contribute to the growth in trade triggered by the tariff reduction. Moreover, due to the concavity of the consumer accumulation technology, these new entrants will grow faster than experienced firms, hence increasing their relative contribution across years. We can see that the contribution at the end of the period is around 30%, which is roughly consistent with the numbers provided in Alessandria et al. (2013).[46] In comparison, the model without consumer margin does not feature this growth in the contribution of the extensive margin.[47]

6.3 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 predictions relative to the standard model in an out-of-sample predictions exercise. Because

[45]See figure 20 in appendix D for the relative contribution of each of these margins across time.

[46]They report a contribution of the extensive margin of 30 and 60% after 5 years, respectively in Uruguay and Mexico.

[47]See figure 21 in appendix D for the prediction using the restricted version of the model.

I study the export decisions on a limited set of destinations, I can take advantage of additional destinations, that I haven't 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. In particular, I want to perform this exercise in a market that has recorded important and measurable trade shocks. This will allow me to feed this shock in the model, and compare the predicted response of both models to the actual behaviors of exporters.

I apply this methodology to the Brazilian wine market during my sample period.[48] The choice of the Brazilian market is based on two reasons: first, the Brazilian market 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 first one is 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 an appreciation starting 2003. This depreciation implied a strong increase in the price of French wines in local currency. The second large shock has come from the Argentinian devaluation that took place in 2002. After the abandon of the peso-dollar parity, the Argentinian currency recorded a strong depreciation that led to a strong growth in wine export to Brazil. As a close neighbor and a massive wine producer, this decrease in Argentinian prices implied 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 behaviors, 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 $\mathcal{I}_{fdt-1}$ will be obtained from the predictions of the model, such that only initial conditions are required for these variables. Therefore, with the variable X_{dt} that describes the aggregate demand from Brazil, the model is able to deliver predictions of entry, sales and prices on the Brazilian market for each of the 200 firms I used in the estimation.

TABLE 4: Top market shares

Country	**Average market share**
France	22.1 %
Italy	20.4 %
Chile	19.6 %
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*'.

I will construct this variable X_{dt} for Brazil by using variations in real exchange rates and the

[48] My sample period goes from 1997 to 2010. However, I will stop my predictions in 2007, since the great trade collapse generated a strong decrease in trade that is difficult to feed in the model.

Brazilian GDP. From the demand equation used in the model, X_{dt} is defined as:

$$X_{dt} = \log Y_{dt} - (1 - \sigma) \log P_{dt} + (1 - \sigma) \log(\tau_{dt} e_{dt})$$

in which Y_{dt} is the amount spent by Brazilian consumers in wine, P_{dt} is the price index for wine in Brazil, and τ_{dt} and e_{dt} are transportation costs and exchange rates between French exporters and Brazilian consumers. Therefore, I will proxy variations in $\log Y_{dt}$ by variations in the log GDP of Brazil, and variations in $\log(\tau_{dt} e_{dt})$ using variations in the BRA/FRA exchange rates. Finally, to construct a proxy for the price index, I will use the variations in exchange rates of the main exporters to Brazil as featured in table 4.[49] Based on these data, I can construct variations in $X_{BRA,t}$ from 1997 to 2007.[50] To obtain the values in level of $X_{BRA,t}$, I will set $X_{BRA,t}$ such that the sales of the median prediction equals the realized sales on the market during the year before the shock, 1998. Therefore, the focus of the exercise will be on variations in sales and entry after this year.

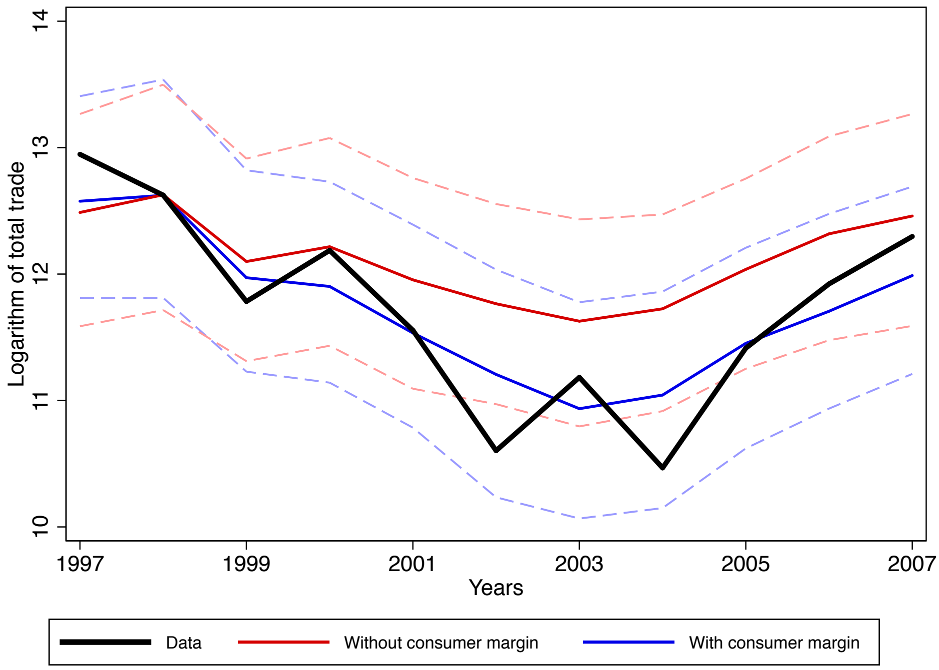

FIGURE 10: Total exports of wine to Brazil from selected firms

The results of these predictions are displayed in figure 10 for the total trade, and figure 11 for the number of exporters. These figures display the realized data, as well as the predictions from the full model with consumer margin and the standard model without consumer margin. Moreover, I report confidence intervals at 90%: each prediction still requires to simulate the non persistent shocks, ε and ν, to infer entry, sales and pricing behaviors, which explains the

[49]These four countries account for 75% of the total wine import of Brazil. The fifth exporter (Portugal) has a market share of less than 2% and therefore is not included in the construction of the price index.

[50]The obtained variations in $X_{BRA,t}$ are displayed in figure 22 in appendix D

variability in the predictions.[51] Figure 10 reports the strong decrease in wine export to Brazil that occurs between 1998 to 2003. This decrease is explained by the Brazilian devaluation in 1999, and the growth in Argentinian export 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. Relatively to the prediction of the models, we can see that the model without consumer margin does not react very much to the changes in exchange rates. This variation in relative prices does reduce sales, but not in the same proportion as in the data. However, the model with consumer margin can predict the large drop in trade, as well as the rebound starting 2004. The main reason between this difference in trade predictions is that the number of exporters does not tend to react very much to exchange rates in the model without consumer margin.

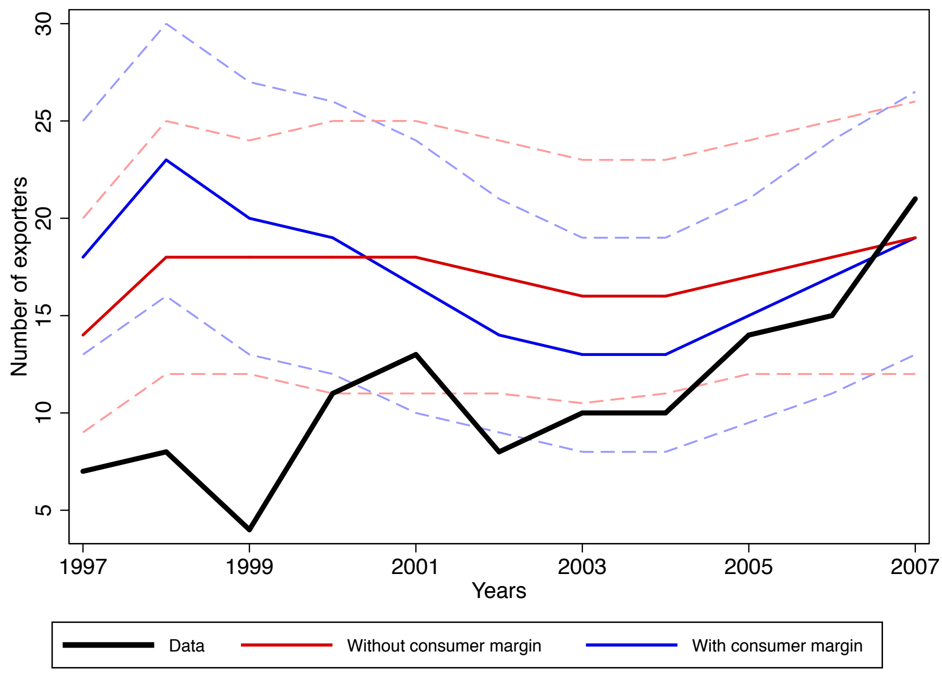

FIGURE 11: Number of wine exporters to Brazil from selected firms

Figure 11 reports the prediction of the number of exporters in the two models. The model with consumer margin, unlike the restricted model, can reproduce the decrease in the number of exporters in 1999 and 2002. This decrease is the reason for the larger variation in total trade shown in the previous figure. The reason for the non exit of exporters in the model without consumer margin comes from the large entry costs of exporting. Therefore, the option value of the exporting activity is so large that no exporters will exit as it will be very hard to reenter in the future. They are willing to lose money temporarily, in order to keep the option value of exporting in the next years. However, in the model with consumer margin and low entry costs, firms are willing to leave the market as the economics condition deteriorates. For similar reasons, as the perspectives on the market improve after 2003, we observe a larger growth rate

[51] For each model, I simulate 500 samples of these shocks, leading to 500 different predictions. I report the median prediction as well as the 5th and 95th percentiles on the figures.

of the number of exporters in the model with consumer margin. However, both models tend to strongly overpredict the number of exporters in the early years of the sample period. Two potential reasons explain this overprediction. First, the model does not account for specific expectations of exporters. Because the law of motion of the aggregate demand term is similar across destinations, the model does not capture the probable low expectations regarding the Brazilian market before the devaluation. Second, part of this overprediction comes from the random nature of the sampling of firms. When looking at aggregate data of the variations in the number of French wine exporters to Brazil, these variations look similar to the observed variations in total trade displayed in figure 10, and to the predictions of the model.

Overall, it appears that the predictions of the model with consumer margin, unlike the standard model, can quantitatively replicate the decrease in total trade during this period. This result mostly comes from the larger response of firms entry and exit, due to the lower level of the entry costs of exporting in this model.

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 a third 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. First, I show that this model can generate a slow response of aggregate trade to shocks. The trade elasticity in the long run is three times larger than the short run, which is consistent with patterns documented in the literature. Second, the model 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 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. This result has important policy implications for countries designing policies 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.*

Arellano, M. and S. Bonhomme (2009): "Robust priors in nonlinear panel data models," *Econometrica*, 77, 489–536.

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 (2015): "Demand learning and firm dynamics: evidence from exporters," *Manuscript.*

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.

Bernard, A. B., R. Massari, J.-D. Reyes, and D. Taglioni (2014): "Exporter dynamics, firm size and growth, and partial year effects," Working Paper 19865, National Bureau of Economic Research.

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

Broda, C. and D. E. Weinstein (2006): "Globalization and the Gains from Variety," *The Quarterly Journal of Economics*, 121, 541–585.

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

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.*

Foster, L., J. C. Haltiwanger, and C. Syverson (2012): "The Slow Growth of New Plants: Learning about Demand?" Working Paper 17853, National Bureau of Economic Research.

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., S. J. Redding, and D. E. Weinstein (Forthcoming): "Quantifying the Sources of Firm Heterogeneity," *The Quarterly Journal of Economics.*

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. (2014): "A structural model of productivity, uncertain demand, and export dynamics," *Manuscript.*

Lincoln, W. F. and A. H. McCallum (2015): "The Rise of Exporting By US Firms," *Manuscript.*

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

Magnac, T. and D. Thesmar (2002): "Identifying dynamic discrete decision processes," *Econometrica*, 70, 801–816.

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 (2014): "Extended Gravity," *Manuscript*.

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.

Piveteau, P. and G. Smagghue (2015): "Estimating firm product quality using trade data," *Manuscript*.

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.

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

Ruhl, K. J. and J. L. Willis (2008): "New exporter dynamics," *Manuscript*.

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. 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

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 displays 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 of all, 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).
  - use weight as counting unit. The reason for this exclusion relies on the use of weight for many categories as the default unit. While this can be a relevant unit for some goods, it is often used for product categories that gather non homogeneous product.
- 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, Massari, Reyes, and Taglioni (2014), a firm will mechanically sell less in average during its first calendar year as exporter. This effect comes from the fact that 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 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. 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.

These procedures 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 destination 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 of implementing 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 wine bottles, whose prices are easy to define, and aggregate at the firm level. For these reasons when defining my sample, I will exclusively use wine producers that do not export any other goods outside of wine. A large share of wine trade is made by wholesalers who export other types of items, and therefore for which the study at the level of the firm is irrelevant. In addition to this homogeneity constraint, my estimation procedure requires to have enough firms exporting to several destinations. As a major exporting industry from France, the wine industry meets both of this conditions: a large number of exporters, exporting a precisely defined good.

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 will 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 will estimate different entry and fixed costs of exporting, as well as different trend in aggregate demand. The list of these destinations can be found in table 5.

TABLE 5: 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 structural sample. The observations related to this destination will be used in the out-of-sample exercise and therefore need to 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 appears 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 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

Similarly 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 exportes. 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 order to do so, I run a regression of the logarithm of prices, on a sets of time, destinations and firm-specific dummies. Formally, I estimate

$$\log p_{fdt} = \alpha_f + \beta_d + \gamma_t + \varepsilon_{fdt}$$

and I define $\log \hat{p}_{fdt} = \hat{\alpha}_f + \hat{\beta}_d + \hat{\gamma}_t$. Therefore I can flag prices that deviate from this predicted prices. In particular, I consider as outlier a price that deviates from a factor 2 of its predicted value ($p_{fdt} > 2\hat{p}_{fdt}$ or $p_{fdt} < 1/2\hat{p}_{fdt}$). As a cleaning procedure I dropped all the observations of a firm which has at least one outlier among its observations.

Finally, a last requirement 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 records at least 15 exporting events. Note that with 14 destinations and 14 years of data, the maximum number of observations by firm is 196. This selection process is could be an issue 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 will tend to go against the theory of consumer accumulation that can accommodate small and short-lived exporters relative to the standard model.

A.2.6 Final sample

Once these cleaning steps implemented, I randomly sample 200 firms among the set of firms available. Moreover, in order to have enough exporters that have an activity in Brazil, and therefore conduct the out-of-sample predictions exercise, I will impose that 100 of these 200 firms have some exporting activity in Brazil during the sample period.

TABLE 6: Description of the sample used in the structural estimation

Statistics:	*pc5*	*median*	*pc95*	*mean*	N
# observations per firm	15	36.5	97.5	44.2	200
av. # destinations per firm-year	1.65	3.64	8.29	4.16	2118
av. # years per firm-destination	2.5	5	9.5	5.29	1626

Table 6 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 15 by the sampling procedure. However, the median firm only records 29 exporting 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 3.11 destinations in an average year.

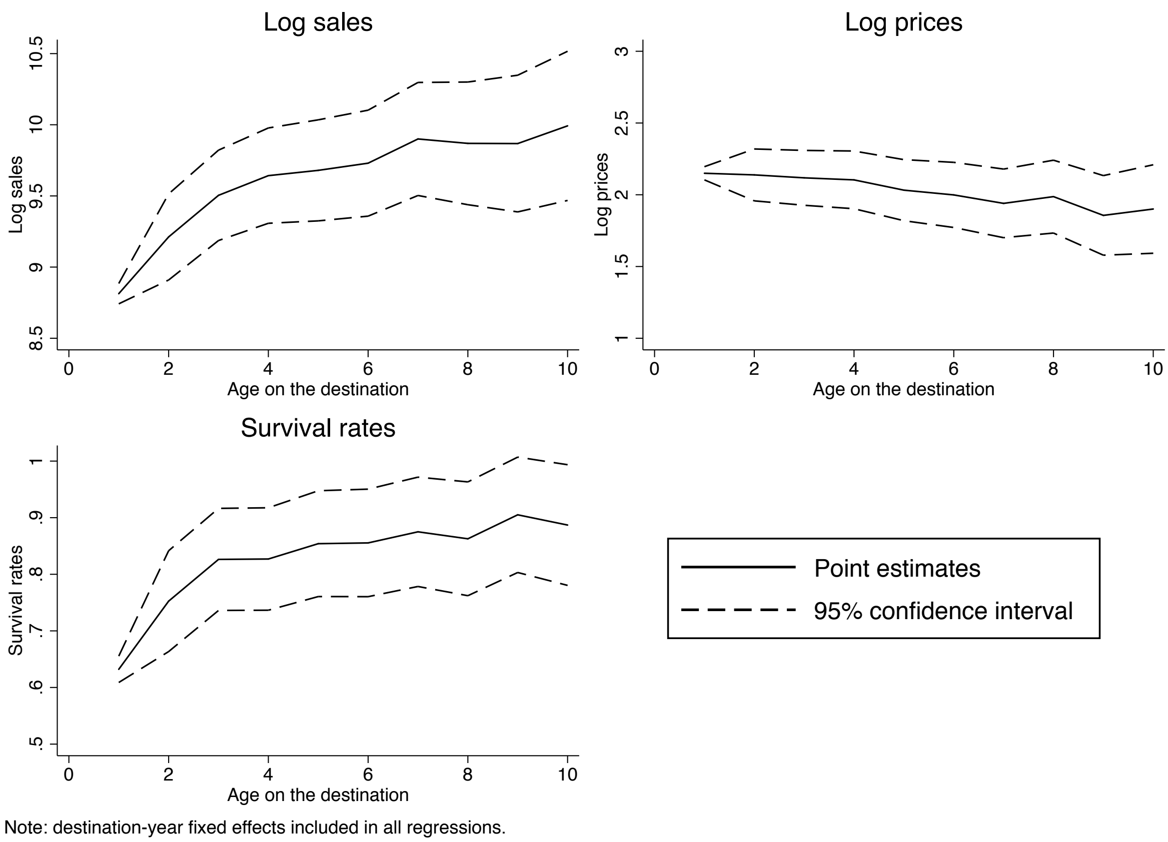

FIGURE 12: Sales, prices and survival rates across ages (Wine producers)

Notes: The figure reports the average log sales, log prices and survival rates of wine producers on 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 on 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 ages dummies I perform in section 2. Figure 12 reports the results of these regressions for sales, prices and survival rates.[52] 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% the first year. Meanwhile, the variations in prices are small and insignificant across ages. However, we can see that the survival rates in the structural sample are larger than the ones displayed in the exhaustive data. While the

[52]Table 7 provides the tables related to these regressions.

survival rate was close to 35% in the full sample, it is around 60% in this restricted sample. The reason comes from the necessary selection I made on the 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 emphasize strong dependence in demand rather than an important role for sunk costs of entry.

Table 7: Age regressions using the structural sample

	No fixed effects			Year x destination fixed effects		
	(1) Log sales	(2) Log prices	(3) Survival rates	(4) Log sales	(5) Log prices	(6) Survival rates
Age 2	0.407***	-0.0199	0.126***	0.366***	-0.0343*	0.122***
	(0.0344)	(0.0165)	(0.0163)	(0.0358)	(0.0161)	(0.0165)
Age 3	0.662***	-0.0254	0.174***	0.627***	-0.0712***	0.171***
	(0.0439)	(0.0221)	(0.0172)	(0.0457)	(0.0214)	(0.0177)
Age 4	0.860***	-0.0295	0.187***	0.849***	-0.0886**	0.188***
	(0.0526)	(0.0270)	(0.0188)	(0.0548)	(0.0270)	(0.0196)
Age 5	0.902***	-0.0200	0.243***	0.898***	-0.0948**	0.231***
	(0.0619)	(0.0336)	(0.0191)	(0.0658)	(0.0334)	(0.0200)
Age 6	0.993***	-0.0339	0.255***	1.006***	-0.111**	0.242***
	(0.0690)	(0.0392)	(0.0204)	(0.0760)	(0.0400)	(0.0216)
Age 7	1.006***	-0.0706	0.246***	1.010***	-0.151**	0.234***
	(0.0791)	(0.0437)	(0.0225)	(0.0886)	(0.0466)	(0.0240)
Age 8	1.053***	-0.0767	0.259***	1.056***	-0.160**	0.249***
	(0.0935)	(0.0497)	(0.0242)	(0.102)	(0.0562)	(0.0266)
Age 9	1.333***	-0.147**	0.318***	1.298***	-0.248***	0.306***
	(0.100)	(0.0519)	(0.0214)	(0.117)	(0.0645)	(0.0234)
Age 10	1.403***	-0.128*	0.309***	1.405***	-0.240***	0.311***
	(0.116)	(0.0568)	(0.0243)	(0.138)	(0.0704)	(0.0280)
Age 11	1.281***	-0.105	0.268***	1.309***	-0.227**	0.274***
	(0.126)	(0.0632)	(0.0352)	(0.158)	(0.0830)	(0.0368)
Age 12	1.455***	-0.105	0.380***	1.576***	-0.252*	0.389***
	(0.170)	(0.0774)	(0.0108)	(0.201)	(0.100)	(0.0225)
Age 13	1.199***	-0.0416	0.199	1.279***	-0.196	0.191
	(0.232)	(0.118)	(0.117)	(0.269)	(0.146)	(0.126)
Age 14	1.608**	-0.429*	.	1.708**	-0.678**	.
	(0.558)	(0.208)	.	(0.589)	(0.254)	.
Constant	8.751***	2.034***	0.620***	8.762***	2.073***	0.623***
	(0.0314)	(0.0214)	(0.0108)	(0.0349)	(0.0216)	(0.0111)
Observations	7525	7525	6821	7525	7525	6821
R^2	0.092	0.002	0.060	0.175	0.172	0.121

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 describe alternative specifications to look at the correlation between sales or prices and age on the export market.

B.1 Additional specifications

Firm-destination-product fixed effects

A natural way to control for heterogeneity across firms that could drive the correlation across ages, is to include firm-destination-product fixed effects such that the regression becomes

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

However, including this set of fixed effects will make impossible the identification of a trend in prices across ages. To understand why, consider first a sample of firms on a give market *pdt*. Because of the market-level fixed effect, their average price is normalized to zero. Now consider this same set of firms a year later. If none of these firms exited, it means that their average price is normalized to zero. More generally, the fact that age is a treatment that is homogenous across firms make the identification of any trend impossible. However, because in the data, some firms will exit the market, it means that this treatment is not entirely symmetrical across firms, such that some identification is possible. But this identification will entirely rely on firms that exit and reenter, with a age that will be one in the future. As a consequence, the inclusion of this set of fixed effects will not control for selection, but instead will make the entry and exit of firms the only source of identification. Figures 13 and 14 report the results of this specification for sales and prices. As we can see, even sales are not increasing with ages with this specification.

Identification across destinations

An alternative way to identify a increase in sales and prices across ages is to compare the similar products sold to different destinations, and therefore having different export experiences. In terms of specifications, it means including a set of firm-product fixed effects such that the variation that identifies the changes with ages is across destinations. However, this specification is also potentially problematic since it compares old destinations, for which the firms has chosen to export first, and young destinations that have been chosen more recently by the firm. Therefore, it is not clear that the ages across these flows are the only differences. To verify this claim, I run the following specification and display the results for sales and prices in figures 15 and 16.

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

We can see that all figures seem to maintain the increasing in trends of sales and prices, even

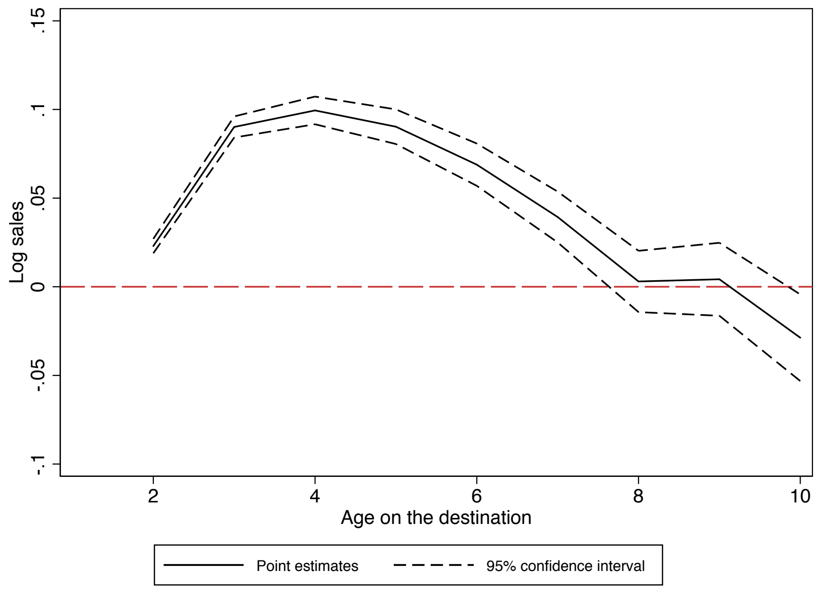

Figure 13: Sales across export ages, within variation

Notes: The figure reports the cumulative growth of sales relatively to age one, of a firm-product category pair on 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 fixed effects. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product-destination level.

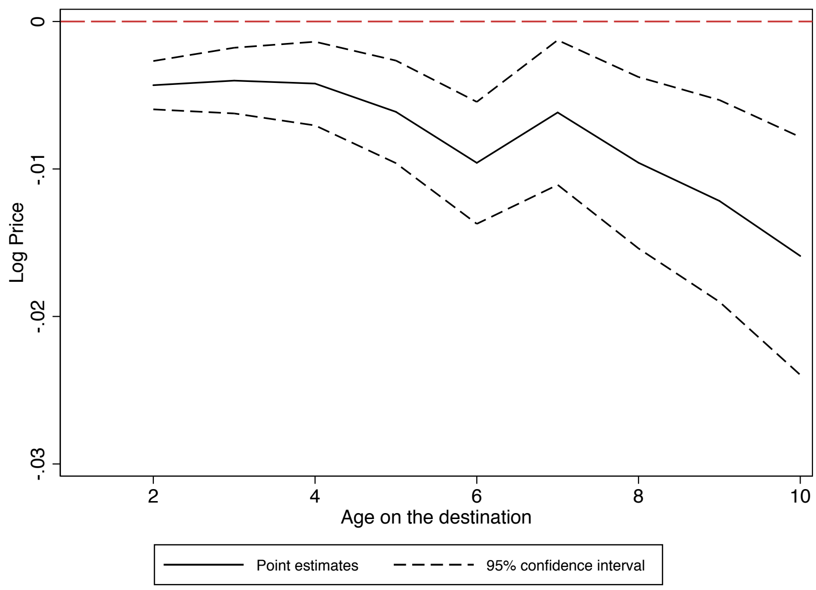

Figure 14: Prices across export ages, within variation

Notes: The figure reports the cumulative growth of prices relatively to age one, of a firm-product category pair on 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 fixed effects. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product-destination level.

though prices regression are not as significant as in the main specification. However, one can see that the endogenous sorting of the destinations seem to play a role in shaping this relationship:

we can see that using a constant set of firms tend to increase the growth in sales. Therefore, it is hard to believe that this specification allow to account for the dynamic selection across ages, but instead generate an endogenous sorting that do not capture the relationship between ages and prices or sales.

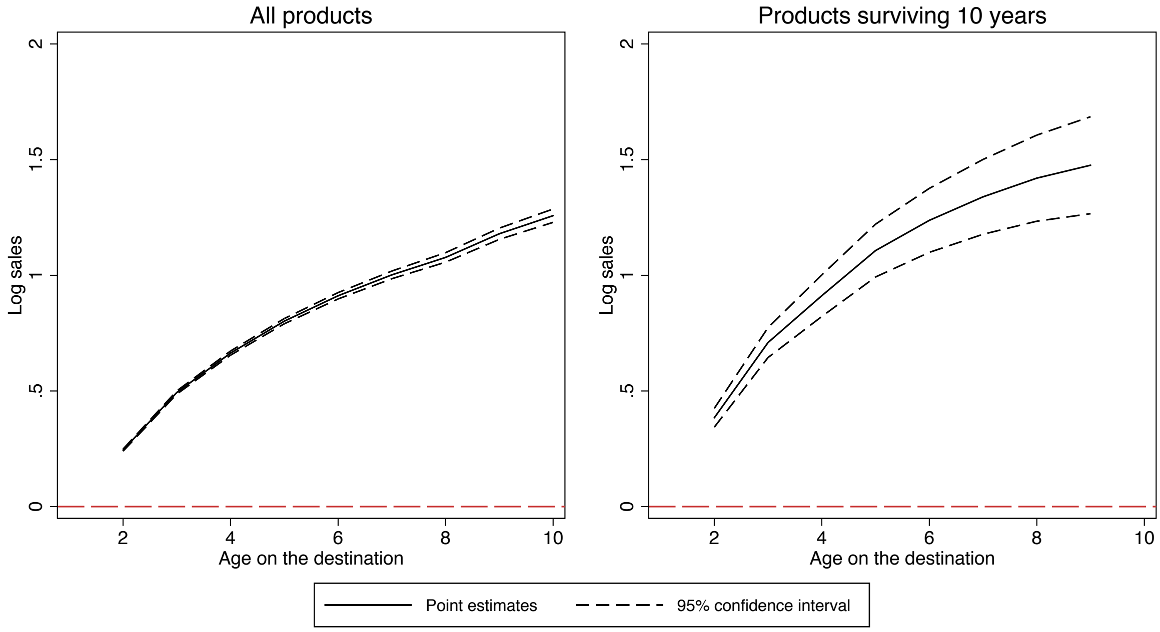

Figure 15: Sales across export ages, across destinations

Notes: The figure reports the cumulative growth of sales relatively to age one, of a firm-product category pair on a destination at different ages. The regression uses logarithm of sales as dependent variable, and includes product category×destination×year and firm×product category fixed effects. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product level.

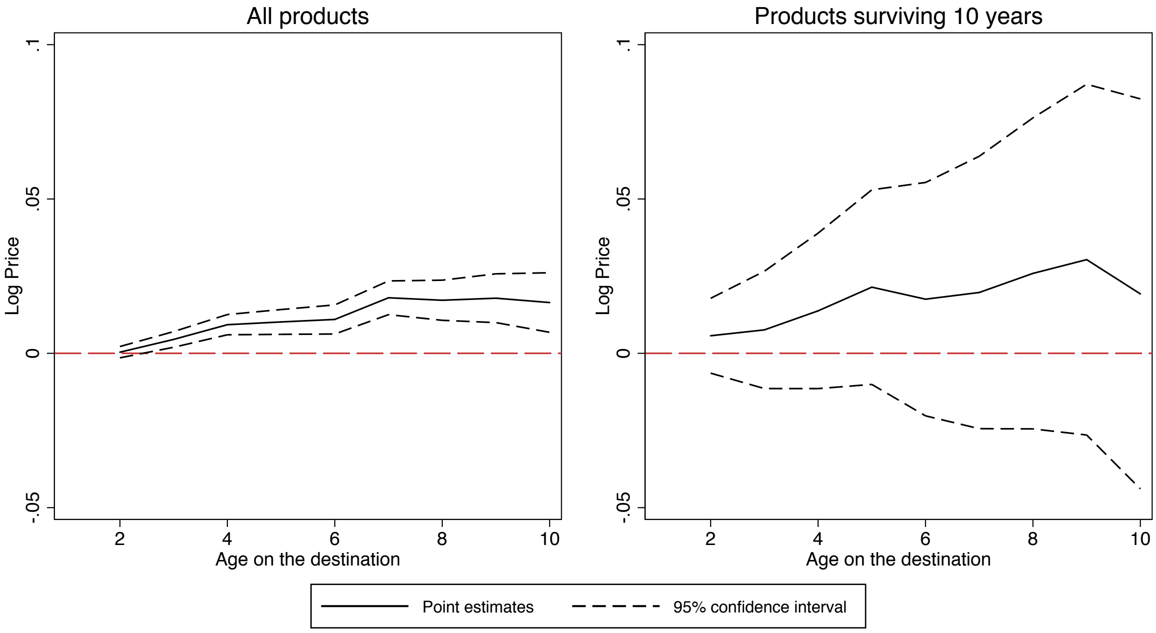

Figure 16: Prices across export ages, across destinations

Notes: The figure reports the cumulative growth of sales relatively to age one, of a firm-product category pair on a destination at different ages. The regression uses logarithm of prices as dependent variable, and includes product category×destination×year and firm×product category fixed effects. The age on a destination is defined as the number of years a firm-product pair has been successively exporting to this country. 95 percent confidence intervals are constructed using standard errors clustered at the firm-product level.

B.2 Tables of results

Table 8: Age regressions (main specification)

	All products			Products surviving 10 years	
	(1) Survival rates	(2) Log sales	(3) Log prices	(4) Log sales	(5) Log prices
Age 2	0.215***	0.550***	0.0207***	0.292***	0.0148
	(0.000675)	(0.00214)	(0.00112)	(0.0162)	(0.0105)
Age 3	0.304***	0.961***	0.0323***	0.528***	0.0304**
	(0.000854)	(0.00317)	(0.00151)	(0.0242)	(0.0112)
Age 4	0.354***	1.240***	0.0470***	0.644***	0.0497***
	(0.00101)	(0.00418)	(0.00189)	(0.0325)	(0.0122)
Age 5	0.380***	1.465***	0.0599***	0.751***	0.0704***
	(0.00118)	(0.00525)	(0.00229)	(0.0410)	(0.0134)
Age 6	0.402***	1.645***	0.0645***	0.795***	0.0795***
	(0.00137)	(0.00652)	(0.00274)	(0.0496)	(0.0143)
Age 7	0.407***	1.808***	0.0771***	0.809***	0.0948***
	(0.00160)	(0.00800)	(0.00330)	(0.0581)	(0.0153)
Age 8	0.419***	1.928***	0.0836***	0.798***	0.114***
	(0.00186)	(0.00973)	(0.00401)	(0.0665)	(0.0163)
Age 9	0.434***	2.051***	0.0855***	0.773***	0.132***
	(0.00215)	(0.0118)	(0.00483)	(0.0752)	(0.0174)
Age 10	0.446***	2.142***	0.0891***	0.637***	0.133***
	(0.00255)	(0.0144)	(0.00574)	(0.0840)	(0.0185)
Constant	0.334***	7.797***	3.799***	9.020***	3.185***
	(0.000290)	(0.00120)	(0.000641)	(0.0431)	(0.0107)
Observations	5311968	5722216	6241358	357751	364700
R^2	0.329	0.439	0.871	0.555	0.918

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 9: Age regressions with alternative specifications

	Firm x product f.e.				Firm x product x dest. f.e.	
	All products		Prod. surviving 10 years			
	(1) Log sales	(2) Log prices	(3) Log sales	(4) Log prices	(5) Log sales	(6) Log prices
Age 2	0.244***	0.000369	0.384***	0.00570	0.0230***	-0.00432***
	(0.00231)	(0.000934)	(0.0207)	(0.00619)	(0.00207)	(0.000836)
Age 3	0.493***	0.00450***	0.709***	0.00760	0.0901***	-0.00401***
	(0.00347)	(0.00130)	(0.0330)	(0.00970)	(0.00304)	(0.00114)
Age 4	0.664***	0.00927***	0.912***	0.0137	0.0995***	-0.00421**
	(0.00459)	(0.00167)	(0.0458)	(0.0128)	(0.00398)	(0.00145)
Age 5	0.802***	0.0102***	1.107***	0.0214	0.0903***	-0.00613***
	(0.00577)	(0.00205)	(0.0582)	(0.0161)	(0.00498)	(0.00177)
Age 6	0.911***	0.0110***	1.237***	0.0175	0.0689***	-0.00959***
	(0.00704)	(0.00240)	(0.0706)	(0.0193)	(0.00609)	(0.00211)
Age 7	1.002***	0.0180***	1.339***	0.0197	0.0393***	-0.00617*
	(0.00845)	(0.00279)	(0.0825)	(0.0225)	(0.00731)	(0.00250)
Age 8	1.077***	0.0172***	1.420***	0.0259	0.00299	-0.00958**
	(0.0105)	(0.00332)	(0.0949)	(0.0257)	(0.00883)	(0.00297)
Age 9	1.180***	0.0179***	1.476***	0.0304	0.00423	-0.0122***
	(0.0124)	(0.00403)	(0.107)	(0.0290)	(0.0105)	(0.00349)
Age 10	1.258***	0.0164***	1.412***	0.0193	-0.0287*	-0.0159***
	(0.0147)	(0.00492)	(0.119)	(0.0322)	(0.0125)	(0.00411)
Constant	7.994***	3.812***	8.631***	3.241***	8.184***	3.817***
	(0.00117)	(0.000445)	(0.0603)	(0.0167)	(0.00109)	(0.000412)
Observations	5722216	6241358	357751	364700	5722216	6241358
R^2	0.716	0.960	0.817	0.979	0.873	0.983

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 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.

C.1 Initial values

I start by describing how the unobservables are obtained, before describing the initial parameters. I start by setting an initial value of 2.2 for σ,[53] that allows me to 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. In order to obtain $\phi_{dt}^{(0)}$, 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 $f^{(0)} = fe^{(0)} = s_v^{(0)} = 1\,000$ for the three different groups of countries.

After setting this initial values, I implement 5000 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 of not accounting for the dynamic problem.

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

C.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

- $\lambda_g \sim N(0, 5\,\mathrm{std}(\lambda_{ft}^{(0)}))$
- $\phi_g \sim N(0, 5\,\mathrm{std}(\phi_{ft}^{(0)}))$
- $X_g \sim N(0, 5\,\mathrm{std}(X_{ft}^{(0)}))$
- $n_g \sim U[\underline{n}^{(0)}\ ;\ 0]$

Note that this grid will be updated every 500 iterations using current unobservables, such that the grid will follow the potential change in the distribution of the unobservables. I will set the size of the grid to be 30 on each dimensions, such that the value function will be iterated at 30^4 different grid points.

[53] I set $\sigma = 2.2$, which is the elasticity obtained by Broda and Weinstein (2006) for the wine industry. Note that I will keep this value constant through the estimation.

C.3 Iteration

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

- the history of value function $\left\{V\left(\Theta^{(s-m+1)}\right), ..., V\left(\Theta^{(s)}\right)\right\}$,
- the set of unobservables $\xi_{fdt}^{(s)} = (\lambda_{ft}^{(s)}, \phi_{ft}^{(s)}, X_{dt}^{(s)})$,
- the history of parameter vectors $\left\{\Theta^{(s-m+1)}, ..., \Theta^{(s)}\right\}$.

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 it defines objects that are used in the other steps. Therefore, I then turn to the sampling of unobservables, and the sampling of parameters.

Update of the value function The value functions is obtained from the Bellman equation, iterated from the previous iterations of the value functions. However, since the value function depends on the set of parameters Θ, I start by finding the nearest neighbor $\Theta^{(h)}$ of $\Theta^{(s+1)}$ in the history $\left\{\Theta^{(s-m+1)}, ..., \Theta^{(s)}\right\}$. Knowing this nearest neighbor $\Theta^{(h)}$, and its associated value function $V(\xi_g, n_g, \Theta^{(h)})$, I can iterate the value function the following way:

$$
\begin{aligned}
V(\xi_g, n_g, \mathcal{I}, \Theta^{(s+1)}) = s_v \log \Bigg[&\exp\left(\frac{1}{s_v} \max_{n' \in n_g} \left\{ E_\varepsilon \pi(\xi_g, n_g, n', \Theta^{(s+1)}) - FC(\mathcal{I}) + EV(\xi_g, n', 1) \right\}\right) + \\
&\exp\left(\frac{1}{s_v} EV(\xi_g, n_0, 0)\right)\Bigg] \qquad (14)\\
\text{with} \qquad EV(\xi_g, n, I) = &\frac{\sum_{\xi \in \xi_g} V(\xi, n, I, \Theta^{(h)})\, P_\xi(\xi \,|\, \xi_g)}{\sum_{\xi \in \xi_g} P_\xi(\xi \,|\, \xi_g)}
\end{aligned}
$$

$P_\xi(.|.)$ being the transition probability of the unobservables at the current parameters. In practice, I can iterate several times the Bellman equation, in order to reduce the error coming from the choice of a nearest neighbor instead of the exact parameter. In this case, I iterate not using the m-th value function anymore, but the current value function and its grid.

In addition to updating the value function, I will define two objects based on the recently updated value functions, that will be used in the sampling of parameters and unobservables. First, I will save the optimal future share of consumer chosen by the firm. This object, evaluated on the grid, will be defined as

$$n_g^* \equiv n'^*(\xi_g, n_g) = \text{argmax}\Big\{E_\varepsilon \pi(\xi_g, n_g, n') + EV(\xi_g, n', 1)\Big\}$$

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

$$DEV(\xi_g, n_g) = EV(\xi_g, n_g^*, 1) - EV(\xi_g, n_0, 0)$$

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

These new value functions are stored in the history of the value functions for use later in the algorithm. The functions $n^*()$ and $DEV()$ will be in comparison used in the next iteration to sample the unobservable.

Sampling of unobservables The marginal density of the unobservables (λ, ϕ or X) is made of three parts:

- the unconditional distribution of the unobservables,
- the entry condition,
- the demand and supply equations.

As an illustration, when looking at a given λ_{ft}, its density, conditional to all the other parameters and unobservables, is

$$\begin{aligned}\lambda_{ft}|... \propto \exp&\left(-\frac{1}{2\sigma_\lambda^2}\left(\lambda_{ft}-\rho_\lambda\lambda_{ft-1}\right)^2-\frac{1}{2\sigma_\lambda^2}\left(\lambda_{ft+1}-\rho_\lambda\lambda_{ft}\right)^2\right)\\ &\times\prod_{d=1}^{D}\left\{\exp(U'_{fdt}\Sigma^{-1}U_{fdt})^{\mathcal{I}_{fdt}}\left[1+\exp\left(\frac{-DV(\xi_{fdt},n_{fdt})+FC(\mathcal{I}_{fdt-1})}{\sigma_\nu}\right)\right]^{-\mathcal{I}_{fdt}}\right. \\ &\qquad\left.\left[1+\exp\left(\frac{DV(\xi_{fdt},n_{fdt})-FC(\mathcal{I}_{fdt-1})}{\sigma_\nu}\right)\right]^{\mathcal{I}_{fdt}-1}\right\}\end{aligned} \tag{15}$$

with

$$U_{fdt}=\begin{pmatrix}\log s_{fdt}-\log n_{fdt}-\lambda_{ft}-X_{dt}+\sigma\log p_{fdt}\\ \log p_{fdt}+\phi_{ft}-\alpha\lambda_{ft}-\log\mu(\xi_{fdt},n_{fdt})\end{pmatrix}$$

I use a Metropolis-Hastings algorithm to sample from this distribution. For each period t, from t=0 to t=T, I draw a set of unobservables λ^*_{ft} from their hierarchical distributions (first line of the formula (15)). Then these new draws are accepted, firm by firm, based on the evaluation of the multivariate normal and exporting probabilities (second and third line from (15)).

The complexity comes from evaluating the functions $DV()$ and $\mu()$ at the proposed unobservables ξ^*. In order to do so, I follow these steps:

- Obtain the targeted n^* for each observation, from interpolation of $n^*()$: $n^*_{fdt}=n^*(\xi_{fdt},n_{fdt})$.
- Compute the contemporaneous profit analytically: $\pi_{fdt}=\pi(\xi_{fdt},n_{fdt},n^*_{fdt})$.
- Evaluate the difference in expected value functions from interpolation $DEV_{fdt}=DEV(\xi_{fdt},n_{fdt})$ to obtain $DV_{fdt}=\pi_{fdt}+\beta DEV_{fdt}-FC(\mathcal{I}_{fdt})$.
- From the first order condition, I obtain a analytic formula for μ: $\mu_{fdt}=\frac{\partial\pi(\xi_{fdt},n_{fdt},s(n^*_{fdt}))}{\partial s(n^*_{fdt})}$

With the values in hand, it is then straightforward to compare firm by firm the conditional densities using λ^* and $\lambda^{(s)}_{ft}$. Once this procedure has been applied for all periods from t=0 to t=T, the same sampling is applied to ϕ_{ft} and X_{dt}, allowing us to obtain a new set of unobservables $\xi^{(s+1)}_{fdt}$.

Sampling of parameters The sampling of parameters is somewhat similar to the unobservables. However, the main difference is that the functions $DEV()$ and $\mu()$ need to be reevaluated for a new Θ, rather than for new unobservables. Consequently, for all the parameters, a Metropolis-Hastings sampler needs to be used. As a second consequence, it is necessary to iterate the value functions for this new parameter Θ in a way similar to the update of the value functions.

Formally, the sampling of a given block of parameter Θ takes the following steps:

- A new parameter Θ^* is drawn using proposal functions.
- The nearest neighbor of Θ^* is found in the history $\{\Theta^{(s-m+1)}, ..., \Theta^{(s)}\}$.
- The value function $V(\xi_g, n_g, I, \Theta^*)$ is obtained from equation (14) and the functions $DEV(\xi_g, n_g)$ and $\mu(\xi_g, n_g)$ are obtained.
- I obtain by interpolation DV_{fdt} and μ_{fdt} as in the step updating the unobservables, allowing us to compute the likelihood function.
- $\Theta^{(s+1)}$ is set to be Θ^* with probability $\max\left\{1, \frac{\prod_t \prod_d \prod_f L_{fdt}(D, \xi_{fdt}^{(s+1)}; \Theta^*)}{\prod_t \prod_d \prod_f L_{fdt}(D, \xi_{fdt}^{(s+1)}; \Theta^{(s)})}\right\}$.

In order to make the update of the parameters more tractable, I divide my set of parameters in blocks, as it is usually done when the set of parameters is large. The blocks of parameters and their proposal functions are the following:

- α, and γ_d using a random walk proposal functions which targets an acceptance rate of 0.25.
- η_1, η_2, n_0, $\underline{n}$ and ψ using a random walk proposal functions which targets an acceptance rate of 0.25.
- Σ using a Wishart distribution from the previous Σ parameters, which targets an acceptance rate of 0.3.
- ρ_ϕ, σ_ϕ, μ_ϕ using a random walk proposal functions which targets an acceptance rate of 0.25. A similar step is implemented for X and λ.
- f and fe, using a random walk proposal functions which targets an acceptance rate of 0.2.
- s_ν using a random walk proposal functions which targets an acceptance rate of 0.4.

C.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 that 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, 15 years and 15 destinations and I run 80 000 iterations of my algorithm, discarding the first 40 000, as I do in the estimation procedure. I report in figures 17 and 18 the Markov chains and the posterior distributions for the fixed costs of exporting, as well as the true value of the parameters displayed by the red lines. As displayed on these figures, the estimation provides confidence interval that are consistent with the true value of the parameters.

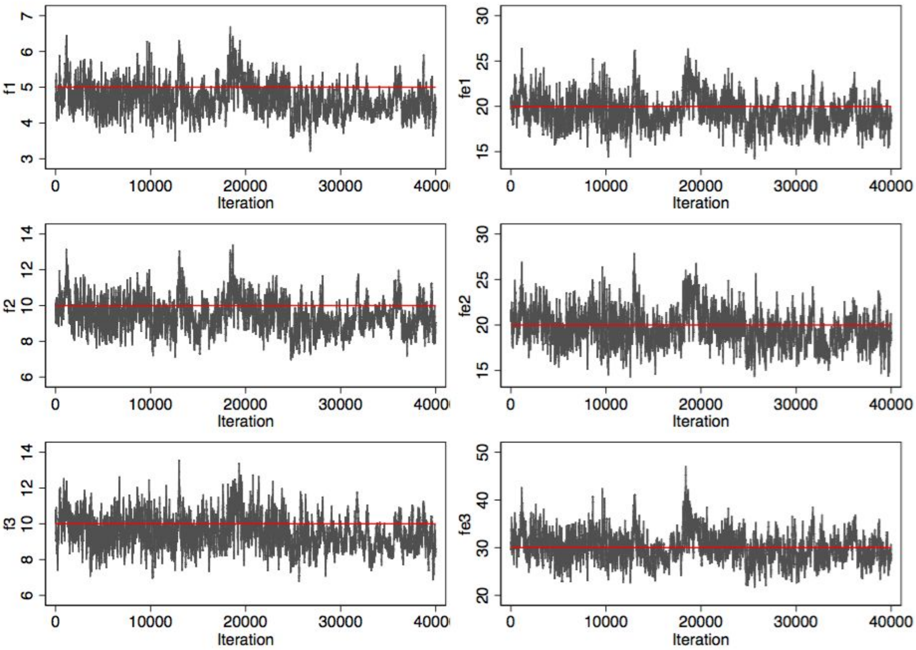

FIGURE 17: Markov Chains for fixed costs on simulated data.

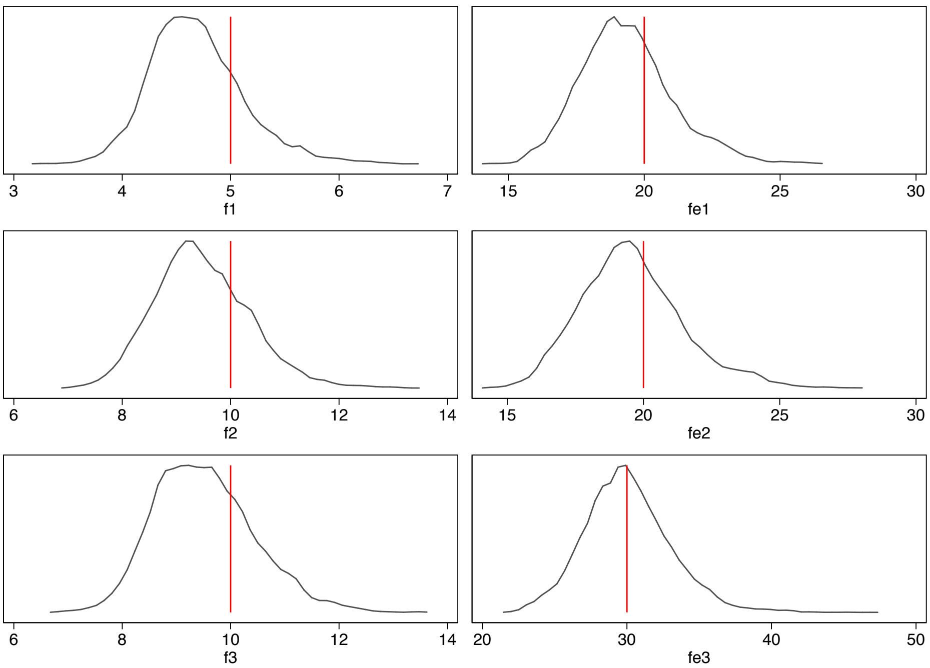

FIGURE 18: Posterior distributions for fixed costs on simulated data.

D Additional figures

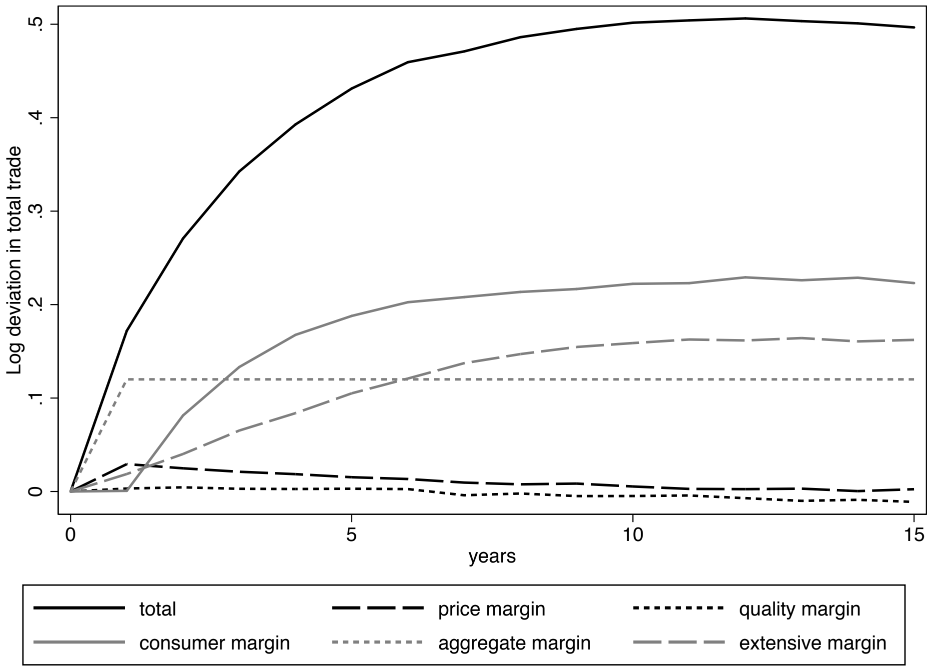

FIGURE 19: Effect of permanent 10 points tariffs decrease (All margins).

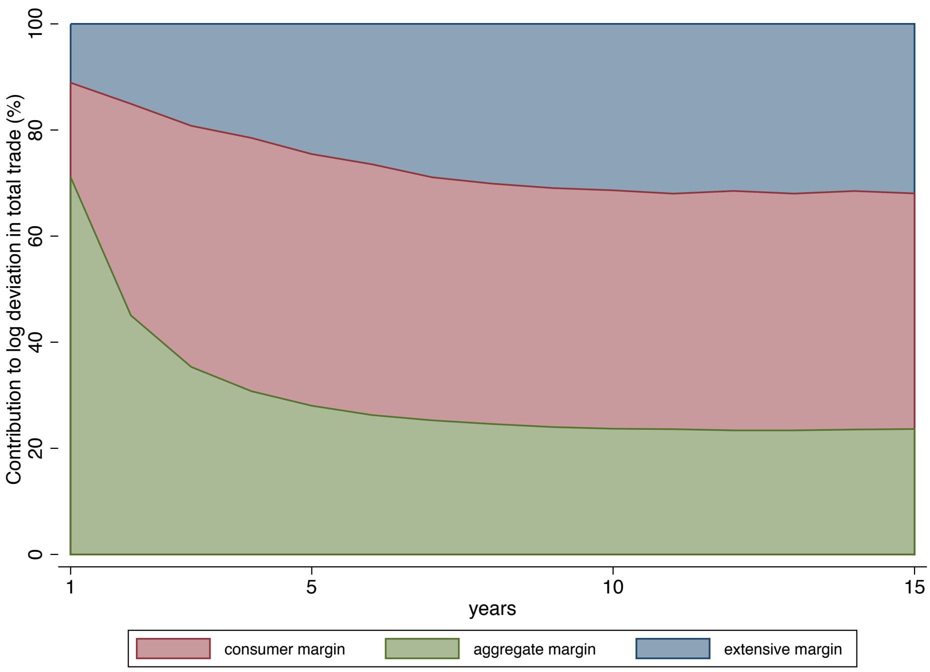

FIGURE 20: Contribution of different margins to trade expansion.

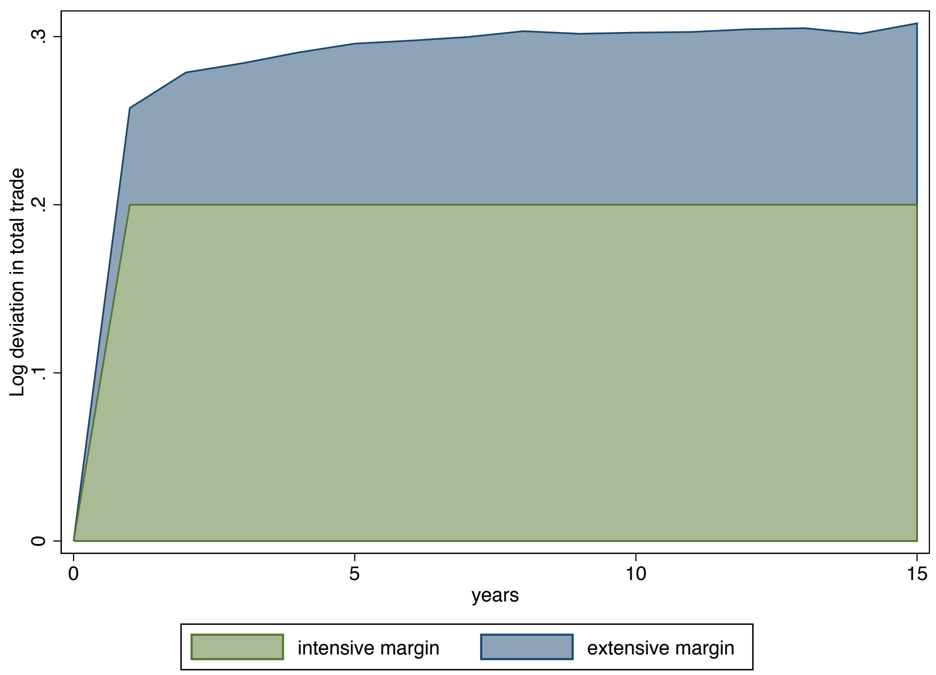

FIGURE 21: Effect of permanent 10 points tariffs decrease (Restricted model).

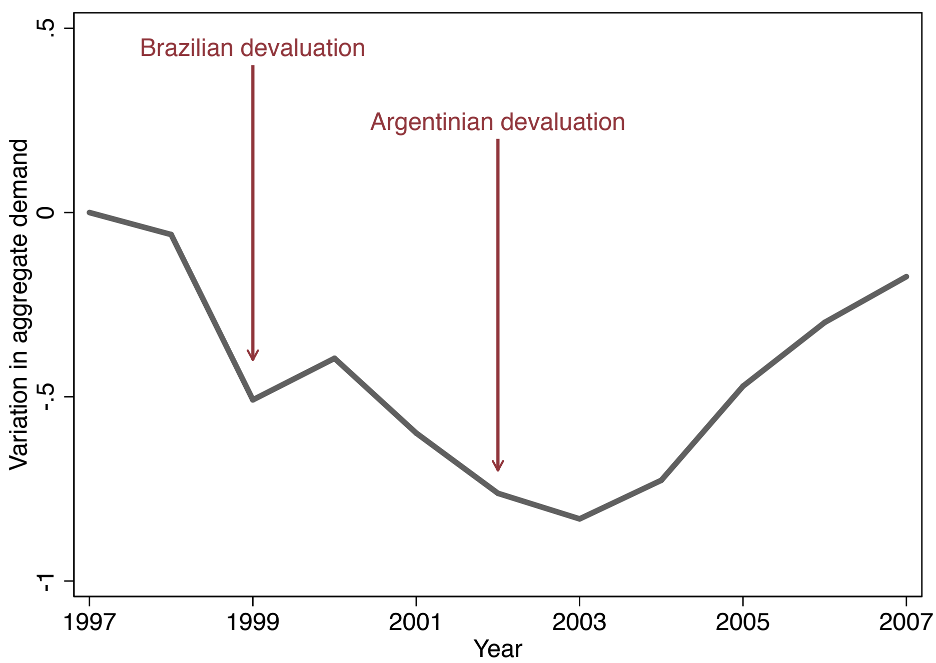

FIGURE 22: Computed variations in aggregate demand for French wine from Brazil.