WORKING PAPERS

N° 17-871

December 2017

# "Demand learning and firm dynamics: evidence from exporters"

Nicolas Berman, Vincent Rebeyrol and Vincent Vicard

Toulouse School of Economics

# Demand learning and firm dynamics: evidence from exporters*

Nicolas Berman† Vincent Rebeyrol‡ Vincent Vicard§

First version: November 2014
This version: December 5, 2017

**Abstract:** This paper provides direct evidence that learning about demand is an important driver of firms' dynamics. We present a model of Bayesian learning in which firms are uncertain about their idiosyncratic demand in each of the markets they serve, and update their beliefs as noisy information arrives. Firms are predicted to update more their beliefs to a given demand shock, the younger they are. We test and empirically confirm this prediction, using the structure of the model together with exporter-level data to identify idiosyncratic demand shocks and the firms' beliefs about future demand. Consistent with the theory, we also find that the learning process is weaker in more uncertain environments.
*Keywords*: firm growth, belief updating, demand, exports, uncertainty.
*JEL classification*: D83, F14, L11.

*We thank the editor Emir Kamenica and four referees, as well as Richard Baldwin, Giuseppe Berlingieri, Patrick Blanchenay, Thomas Chaney, Simon Fuchs, Fabien Gensbittel, Joe Francois, Sacha Kapoor, Omar Licandro, Erzo Luttmer, Thierry Mayer, Marti Mestieri, Rahul Mukherjee, Franck Portier, Jim Tybout and seminar participants at the Geneva Trade and Development Workshop, EBRD, Aix-Marseille University, Universita Cattolica Milano, Nottingham University, Universidad de Navarra, Tilburg University, Banque de France, the VI$^{th}$ Villars Research Workshop on International Trade, LETC, DEGIT conference, SED conference and Compnet conference at ECB for very useful discussions and comments. The authors gratefully acknowledge financial support from Fondation Banque de France. Nicolas Berman thanks the A*Midex for financial support (grant ANR-11-IDEX-0001-02 funded by the French government "Investissement d'Avenir" program). The opinions expressed in this paper are those of the authors and do not reflect the views of the Banque de France. This paper features an online appendix containing additional results and available on the authors' webpages.

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS, Graduate Institute Geneva and CEPR. E-mail: nicolas.berman@univ-amu.fr.

‡Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France. E-mail: vincent.rebeyrol@ut-capitole.fr

§CEPII and Banque de France. E-mail: vincent.vicard@cepii.fr.

# 1 Introduction

Why do some firms grow faster than others? While some producers rapidly expand after entry, many others do not survive the first few years. After some time however, those surviving firms account for a large share of sales on both domestic or foreign markets (Haltiwanger et al., 2013; Bernard et al., 2009; Eaton et al., 2008). In the case of French firms, those that did not serve foreign markets a decade earlier account for 53.5 percent of total foreign sales, of which 40 percent comes from post-entry growth.[1] Understanding the sources of heterogeneity in post-entry firm dynamics – survival and growth – is therefore crucial to explain the dynamics of aggregate sales and firm size distribution.

Firm dynamics are characterized by a number of systematic patterns, which have been documented by a large body of empirical literature. New firms start small and have higher exit rates. For those that survive, the average growth of their sales declines with their age.[2] These facts can be rationalized by several theories, relying on different underlying mechanisms such as stochastic productivity growth, endogenous R&D investment, financial constraints, adjustment costs, demand accumulation or demand learning. Yet, empirically, disentangling the role of these specific channels has proven difficult, as it requires identifying separately the contributions of idiosyncratic demand and productivity to the variations of firms sales. For this reason, the literature has followed an indirect approach: it has studied which models are able to replicate the behavior of observables such as sales growth and exit. In contrast, this paper directly tests for the existence of demand learning by identifying firms' beliefs about demand and the signals they receive, and shows that it is an important driver of post-entry firm dynamics.

We first document two novel stylized facts using detailed data from the French customs containing information on firms' sales by destination and 6-digit product between 1994 and 2005. Throughout the paper, we refer to a product-destination pair as a *market*, and define age as the tenure of a firm in a specific *market*. We show that existing results about aggregate firm behavior carry over at the firm-market level. More precisely, sales growth, exit rates and the variance of sales growth within cohort all decrease with the age of the firm in its market. Importantly, these patterns are still present after controlling for firm-market size or conditioning on firm-product-year fixed effects. In addition, we find that the market-specific growth paths after entry are highly heterogeneous across firms: while entrants grow on average in their first years, a significant share of survivors exhibit negative post-entry growth in the markets they serve. For instance, around 40% of the firms that enter a market in 1996 and stay until 2005 sell *less* at the end of the period than in their second year.

We then present a standard model with Bayesian demand learning in the spirit of

[1]These numbers are based on the 1996-2005 period – see online appendix, Section B.

[2]See Evans (1987), Dunne et al. (1989), Cabral and Mata (2003), Haltiwanger et al. (2013) among many others. Eaton et al. (2008), Buono and Fadinger (2012), Berthou and Vicard (2015), or Ruhl and Willis (2017) show that these dynamics are also observed for exporters.

Jovanovic (1982) that can rationalize these facts.[3] Firms operate under monopolistic competition and face CES demand, but at the same time are uncertain about their idiosyncratic demand in each market, and learn as noisy information arrives in each period. These signals determine the firms' posterior beliefs about demand, on which they base their quantity decision. A higher than expected signal leads younger firms to update more their beliefs than older ones, which implies that the growth rates of young firms are more volatile, even conditional on their size. The model also predicts that market-specific uncertainty limits the extent of belief updating and the impact of age on the updating process. The main contribution of this paper is to test these core predictions, which are specific to the passive learning mechanism.[4]

To do so, we derive from the theory a methodology which allows to separately identify the firms' beliefs and the demand shocks (the signals) they face in each period, in each of the markets they serve. First, we purge market-specific conditions and firm-specific supply side dynamics (e.g. productivity) from quantities and prices. This is made possible by a unique feature of international trade data, in which we can observe the values and quantities sold of a given product by a given firm in different markets. This is key as it enables to cleanly separate productivity from demand variations. In addition, observing different firms selling the same product in the same destination allows to control for aggregate market-specific conditions. Second, we use the fact that, in the model, quantity decisions only depend on the firms' beliefs while prices also depend on the realized demand shocks. This allows to separate out the firms' beliefs from the demand signals. Hence, while requiring few, standard assumptions, our methodology allows to directly test predictions that relate the evolution of firms' beliefs to firm age and demand signals.

We find strong support for the core predictions of the model. Belief updating following demand shocks is stronger for younger firms, with age being defined at the firm-product-destination (i.e. firm-market) level. Further, using a theory-based measure of market-specific uncertainty, we find that the learning process is significantly weakened and less dependent on age in more uncertain environments. We provide several robustness exercises to show that these results are not driven by our main modeling assumptions. Our findings survive after accounting for potential endogenous selection bias, and are extremely stable across alternative samples, specifications and changes in variables' measurement. We also discuss the implications of relaxing several important assumptions of the model related to the timing of price and quantity adjustment, market structure and firms' productivity. We show that even after relaxing these assumptions, our results can still be interpreted as evidence of belief updating. Some of these extensions however require that we control for firm-market size in our estimations, which leaves our results unchanged.

[3] In Jovanovic (1982), firms actually learn about their cost parameter. While the learning mechanism is the same, we apply it to demand, as in Timoshenko (2015).

[4] We additionally show in the online appendix that exit behavior is also consistent with the learning model: the exit rate decreases with firms' beliefs and the demand shocks the firms face, and demand shocks trigger more exit in younger cohorts.

The literature has proposed a number of potential supply or demand side drivers of firm dynamics. But, learning apart, they cannot explain our main result of a smaller quantity adjustment to past demand shocks for older firms. Suppose indeed that firms have full information about demand, except about a stochastic shock each period. If these shocks are *iid*, there is no reason for the firm to adjust quantities the period after, as these shocks do not convey any information and have no relevance beyond the current period. If instead shocks are persistent, there is no reason for older firms to react less to a shock of a given size. Alternative mechanisms are also difficult to reconcile with our stylized facts.

On the supply side, several papers attempt to explain the heterogeneity in firm size with productivity variations only (through stochastic shocks or endogenous decisions).[5] By construction, they are not able to generate an age dependence of firm growth, *conditional* on size. In contrast, models introducing additional sources of heterogeneity, such as financial constraints or capital adjustments costs, are able to generate this age dependence.[6] Yet, since these sources of heterogeneity apply to the firm as a whole, they cannot deliver the heterogeneous firm-market specific dynamics that we find in the data.

Beyond learning, some demand side mechanisms could be affecting firm growth at the market level. Various processes giving rise to demand accumulation have been proposed. Firms could engage in market-specific investments (e.g. Ericson and Pakes, 1995, Luttmer, 2011, Eaton et al., 2014, Fitzgerald et al., 2016), price low in their first years to build a consumer base (Foster et al., 2016, Gourio and Rudanko, 2014, Piveteau, 2016), or face demand that evolves exogenously over time (Ruhl and Willis, 2017). Among the most recent contributions, Ruhl and Willis (2017) use a model with stochastic entry costs and gradual increase in demand to match the average growth and exit rates of Colombian exporters. Arkolakis (2016) shows that a combination of idiosyncratic productivity shocks and market penetration costs is able to reproduce some important patterns of the distribution of US and Brazilian exporters' growth. Since these models include some mean reversion effects, they can generate an age dependence of firm growth conditional on size; but they fall short at predicting the decline in the variance of growth rates with age, conditional on size.[7] On the other hand, we show that our estimates of firms' beliefs reproduce well this observed decline in the variance of sales growth.

The last part of our paper discusses whether alternative demand based theories, possibly on top of a learning effect, could be driving our findings. We show in particular that

[5] See for instance Hopenhayn (1992), Luttmer (2007), Impullitti et al. (2013) for models with stochastic shocks to productivity, Klette and Kortum (2004) or Rossi-Hansberg and Wright (2007) for theories of endogenous productivity growth.

[6] See for example Cooley and Quadrini (2001) or Clementi and Hopenhayn (2006) for financial constraints, and Clementi and Palazzi (2016) for adjustments costs.

[7] For example, Arkolakis (2016) assumes an exogenous Ornstein-Uhlenbeck process for productivity, which generates an age dependence of firm growth at the market level, conditional on size. But this set-up cannot explain the decline in the variance of growth within cohorts at the market level, as Ornstein-Uhlenbeck processes have a constant variance. What would be needed is a process that implies both smaller shocks over time *and* a smaller variance of these shocks. This is not a standard feature of the most common stochastic processes.

theories of demand accumulation would have serious difficulties matching the profiles of prices and quantities that we find. Indeed, in our data, once purged from their productivity component, firm-market-specific prices are (slightly) decreasing with age. Such a pattern contradicts models of active demand accumulation through pricing decisions or models featuring learning in which firms set prices rather than quantities. It is however consistent with the passive demand learning model, in which survivors tend to have received relatively more "good news" than exiters, leading them to adjust their prices upwards to take advantage of this unexpectedly high demand. As firms get better informed over time, their prices converge to their optimal pricing rule. However, once composition effects are controlled for, prices – in the model and in the data – are constant as firms have equal probabilities to update upwards or downwards. Similarly, quantities should increase over time, but in the learning model this is mostly due to selection. This prediction is confirmed empirically: when accounting for composition effects triggered by selection, we find that quantities within firms-markets exhibit a very limited positive growth, observed only in the first years. This matches well our second stylized fact: a substantial part of survivors shrinks in size due to their "over-optimistic" beliefs at entry.[8]

Overall, these results do not preclude alternative mechanisms to be jointly at work, but they clearly suggest that the patterns we identify in our data are unlikely to be driven by demand accumulation processes. Demand learning appears to be an important determinant of the micro-dynamics of firms in narrowly defined markets, which is key as more than half of the variance of sales growth in our sample is due to firm-market factors. This supports the view of several recent works arguing that demand learning models reproduce well some important characteristics of the dynamics of firms and exporters.[9] Compared to these papers, we follow a different strategy as we propose a direct test of the updating process, which lies at the core of the learning mechanism. Our empirical methodology is close in spirit to Foster et al. (2016, 2008), in that they also separate idiosyncratic demand shocks from firms' productivity, but our paper differs in several ways. In particular, we do not need to measure productivity or other firm-specific determinants of sales to identify demand shocks.

Finally, we assume that the actual sales of a firm in a given product-destination market are the only source of information about demand. In other words, we assume away information spillovers. A firm's belief in a given market might well be affected by its beliefs in other destinations (Albornoz et al., 2012), or about other products in the same destination (Timoshenko, 2015). These effects might be stronger for similar destinations and products (Morales et al., 2014; Defever et al., 2015; Lawless, 2009). The behavior of other firms serving the same market might also play a role (Fernandes and Tang, 2014). Studying the relative importance of these various potential sources of information is an interesting and

[8]We also perform a test initially proposed by Pakes and Ericson (1998), in which we regress current firm beliefs on immediate past beliefs and initial beliefs. Consistent with a passive learning model we find that initial beliefs are useful to forecast future firms' beliefs throughout their life.

[9]See for instance Albornoz et al. (2012), Timoshenko (2015), Fernandes and Tang (2014) or Eaton et al. (2014).

vast question in itself, that we indeed plan to study in the future, but which is beyond the scope of this paper.

The paper proceeds as follows. In the next section, we describe our data, document new stylized facts about firms' post-entry dynamics, and discuss them in light of existing theories. In section 3 we present the model and our identification strategy. Section 4 contains our main results and section 5 various robustness exercises. Section 6 discusses whether our results could be explained by alternative demand-based mechanisms. The last section concludes.

# 2 Firm dynamics on foreign markets and export growth

## 2.1 Data

We use detailed firm-level data by product and destination country provided by the French Customs. The unit of observation is an export flow by a firm $i$ of a product $k$ to a destination $j$ in year $t$. A product is defined at the 6-digit level (HS6). The data cover the period from 1994 to 2005, and contain information about both the value and quantity exported by firms, which will allow us to compute firm-destination-product specific unit values that we will use as a proxy for prices in the second part of the paper. Section A.1 of the online appendix provides more details on the source data.

Two important notes on the terminology we use throughout the paper. First, what we call a *market* is a product×destination combination. Second, *age* is defined by market. Our baseline definition of age is the number of years of presence since the last entry of a firm in a product-destination. Age is reset to zero whenever the firm exits for at least a year from a specific market. What we call age is therefore equivalent to market-specific tenure. Section I.2 in the online appendix discusses alternative measures. Note that in all the empirical analysis, to ensure the consistency of our measures of age, we drop firm-product-destination triplets already present in 1994 and 1995, as these years are used to define entry.

Finally, a *cohort* of new exporters in a product-destination market includes all firms starting to export in year $t$ but that were not exporting in year $t-1$, and we are able to track all firms belonging to a cohort over time.

Our final dataset covers the sales of 3,844 HS6 product categories to 179 destinations by 77,076 firms over the period 1996-2005. All these firms entered at least one market over the period.

## 2.2 Stylized facts

In this section we provide two novel stylized facts on the post-entry dynamics of firms at the product-destination level. The first is that growth rates and their variance within cohorts decline sharply with age, within *firm-markets and conditional on size*. The second is that among survivors, growth paths are highly heterogeneous, with a large number of firms exhibiting negative growth rates. We will argue that both facts are difficult to reconcile with most theories of firm dynamics apart from the passive learning model.

Before explaining these facts in more details, note that our data exhibits patterns that are in line with those found by the literature. Consistent with the results of Eaton et al. (2008) on Colombian data (see also Haltiwanger et al., 2013 and Bernard et al., 2009), we find that new firms-markets contribute disproportionately to aggregate trade growth: new flows account for only 12.3 percent of total export value after a year, but this share reaches 53.5 percent after a decade. Moreover, regressing firm-market sales growth on various sets of fixed effects, we find that market-time and firm-product-time factors only account for 44 percent of the variance of sales growth, a result which echoes the findings of Eaton et al. (2011) or Munch and Nguyen (2014). In other words, firm-market factors are key to explain growth dynamics. The online appendix, section A, provides further discussion of these results.

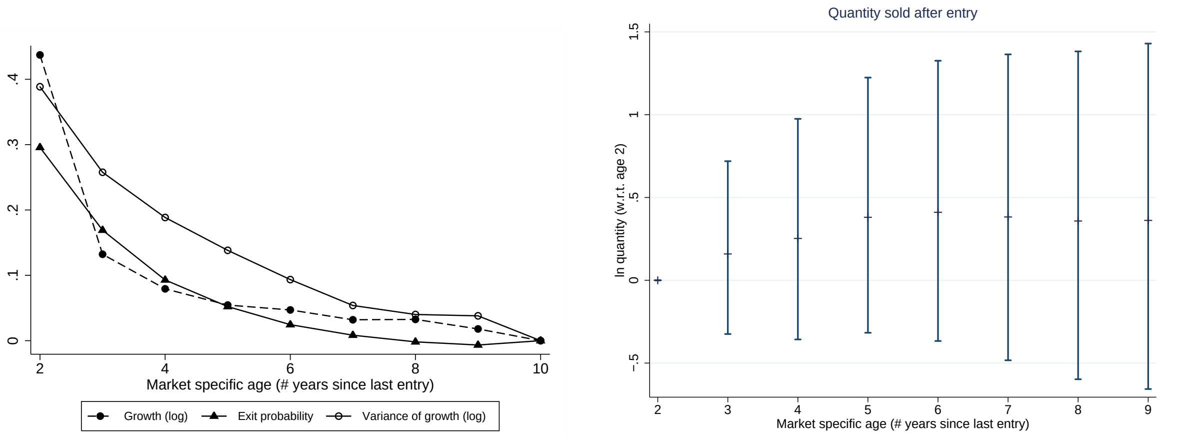
**Fact #1: Firm-market growth and its variance decline with age, conditional on size.** Contrary to most existing papers that have documented facts about the aggregate dynamics of firms or exporters, our data allows to study growth and survival in each market served by the firms. We consider three components of firm-market post-entry dynamics: sales growth, exit rate, and the variance of sales growth within cohort. Figure 1.a plots the coefficients obtained by regressing these different variables on age dummies, controlling for sector and time dummies and, more importantly, for bins of firm size. The full set of results is shown and further discussed in the online appendix, section A.2. All three sets of coefficients sharply decrease with age, with age being defined as firm-product-destination specific tenure. Both the growth rates of firms-markets and the variance of these growth rates within cohort is about 40 percent higher in the second year than after ten years. Importantly, we still find that sales growth declines with age when we include in our regressions firm$\times$product$\times$year fixed effects which control for any unobserved supply side factors (like financial constraints) which are common to all markets within a firm (see online appendix, Table A.3, column 2).

**Fact #2: Post-entry growth dynamics are heterogeneous across survivors.** Our second stylized fact appears in Figure 1.b, where we plot the log of quantities sold by firms entering a given market in 1996 and staying the entire period (until 2005).[10] Quantities are normalized to one in year two.[11] The horizontal lines depict the first quartile, the median and last quartile at each age. Survivors grow after entry consistently with existing evidence (Eaton et al., 2008, 2014; Foster et al., 2016; Ruhl and Willis, 2017; Fitzgerald

[10] A similar pattern is obtained with different dates of entry, or using values instead of quantities. See Figures A.1 and A.2 in section A of the online appendix.

[11] We do not consider the first year because of its potential incompleteness when measured over a calendar year (Berthou and Vicard, 2015). Similarly, we plot the statistics up to 9 years and not 10 because we want to look at flows that will still be present the year after (and 10-year-old flows can only observed in 2005, which is the last year of our sample). The online appendix section I.3 discusses this point.

Figure 1: Stylized facts




(a) Impact of age conditional on size (b) Quantity dynamics for surviving firms

Note: Figure (a) plots the coefficients obtained from of a regression of the log change of firm sales (respectively variance of firms' sales and exit) on age bins, firm size and year and sector dummies (see table A.3 in online appendix section A for the full set of results). All coefficients are relative to the omitted category, age of ten years. The variance of firms' sales growth is measured within cohorts of firms on a product-destination market. Similar patterns are obtained when controlling for country-and-sector fixed effects. Figure (b) plots statistics about market-specific firm quantities with respect to age for the cohorts of firms which entered in 1996. Quantities are normalized to 1 in age 2. The upper and lower limits of the lines represent the first and last quartiles of the variable, with the median in between.

et al., 2016), a pattern that has motivated theories of demand accumulation. But Figure 1.b makes it clear that growth paths are greatly heterogeneous and that a significant share of firm-markets experience negative growth. More precisely, around 40 percent of the firms shown in this figure sell actually *less* at the end of the period than in their second year.

As mentioned in the introduction, the set of facts shown in Figures 1.a and 1.b is difficult to rationalize using existing theories that do not incorporate learning. Models featuring solely supply side dynamics that are firm or firm-product specific (productivity, financial constraints, capital adjustment costs) cannot help understanding the behavior of firms-products across destinations. Theories introducing both supply and demand mechanisms are better designed to explain a heterogeneity across destinations, but they typically fail to generate the dependence of the variance of growth rate to firm age that we observe in the data. Finally, in models of firm dynamics with demand accumulation, survivors tend to be those that have been able to accumulate demand. This allows to fit the average growth path of new firms/exporters observed in the data but does not necessarily provide a framework to think about heterogeneous outcomes across firms. On the other hand, the passive learning model naturally generates these patterns. The decline in the variance of growth rates with cohort age is caused by the larger updating of younger firms. The decline in growth rates is mostly driven by selection: firms that decline the most in size exit the market, which implies that the distribution of growth rates is truncated from below. Together with their larger variance, this implies larger growth rates for younger firms, conditional on survival. It should be noted that the passive learning model is also consistent with larger unconditional growth rates.[12] Finally, the high heterogeneity in firms' growth paths after some years comes from the fact that initial prior beliefs may not be accurate, leading some firms to shrink in size over time.

# 3 A model of firm growth with demand learning

We consider a standard model of international trade with Dixit-Stiglitz monopolistic competition and demand learning in the spirit of Jovanovic (1982). As earlier, we index firms by $i$, destination markets by $j$, products by $k$ and time by $t$.

## 3.1 Economic environment

**Demand**. Consumers in country $j$ maximize utility derived from the consumption of goods from $K$ sectors. Each sector is composed of a continuum of differentiated varieties of product $k$:

$$\begin{aligned} U_j &= \mathbb{E}\sum_{t=0}^{+\infty} \beta^t \ln\left(C_{jt}\right) \\ \text{with } C_{jt} &= \prod_{k=0}^{K}\left(\int_{\Omega_{kt}} \left(e^{a_{ijkt}}\right)^{\frac{1}{\sigma_k}} c_{kt}(\omega)^{\frac{\sigma_k-1}{\sigma_k}} d\omega\right)^{\frac{\mu_k\sigma_k}{(\sigma_k-1)}} \end{aligned}$$

with $\beta$ the discount factor, $\Omega_{kt}$ the set of varieties of product $k$ available at time $t$, $c_{kt}$ is the consumption level of each variety, and $\sum_k \mu_k = 1$. Demand in market $j$ at time $t$ for a variety of product $k$ supplied by firm $i$ is given by:

$$q_{ijkt} = e^{a_{ijkt}} p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} \qquad \text{where} \qquad P_{jkt}^{1-\sigma_k} = \int_{\Omega_{kt}} e^{a_{ijkt}} p_{ijkt}^{1-\sigma_k} d\omega \tag{1}$$

where $\sigma_k$ is the (sector-specific) elasticity of substitution, $Y_{jt}$ is total expenditure and $P_{jkt}$ is the ideal price index of destination $j$ in sector $k$, during year $t$. The demand parameter $a_{ijkt}$ is given by $a_{ijkt} = \overline{a}_{ijk} + \varepsilon_{ijkt}$, with $\varepsilon_{ijkt}$ a white noise. $\overline{a}_{ijk}$ is an idiosyncratic constant parameter and is unknown to the firm.

**Production**. Each period, firms make quantity decisions for their product(s), before observing demand in each market served, i.e. before observing $a_{ijkt}$. The unit cost function is linear in the marginal cost and there is a per-period fixed cost $F_{ijk}$ to be paid for each product-destination pair. Labor $L$ is the only factor of production. Current input prices are taken as given (firms are small) and there is no wedge between the buying and selling price of the input (i.e. perfect reversibility in the hiring decision). Hence, the quantity decision is a static decision.

[12] This is however generated by functional form assumptions. See the online appendix G for details.

We do not make any assumption on the evolution of firm productivity. Productivity may also be subject to learning, in which case the firm would base its quantity decision on its beliefs about its costs. As we will not back out learning from firms' productivity, we do not add expectation terms here to save on notations. We only need to assume that unit costs at the firm-product level are *not* destination specific – we come back to this assumption in section 3.3. Per period profits in market $j$ from product $k$ write:

$$\pi_{ijkt} = q_{ijkt}p_{ijkt} - \frac{w_{it}}{\varphi_{ikt}}q_{ijkt} - F_{ijk} \tag{2}$$

where $w_{it}$ is the wage rate in the origin country, $\varphi_{ikt}$ is the product-time specific productivity of firm $i$.

**Learning.** Firm $i$ is uncertain about the parameter $\overline{a}_{ijk}$. Before observing any signal, its prior beliefs about $\overline{a}_{ijk}$ are normally distributed with mean $\theta_{ijk0}$ and variance $\sigma^2_{jk0}$. Different firms may well have different initial beliefs prior to entry (i.e. different $\theta_{ijk0}$). $\theta_{ijk0}$ is drawn from a normal distribution with mean $\overline{a}_{ijk}$ and variance $\sigma^2_{jk0}$: prior beliefs may not be accurate, but are unbiased on average.[13] The firm observes $t$ independent signals about $\overline{a}_{ijk}$: $a_{ijkt} = \overline{a}_{ijk} + \varepsilon_{ijkt}$, where each $\varepsilon_{ijkt}$ is normal with (known) mean 0 and variance $\sigma^2_\varepsilon$. According to Bayes' rule, the firm's posterior beliefs about $\overline{a}_{ijk}$ after $t$ signals are normally distributed with mean $\widetilde{\theta}_{ijkt}$ and variance $\widetilde{\sigma}^2_{ijkt}$, where:

$$\widetilde{\theta}_{ijkt} = \theta_{ijk0}\frac{\frac{1}{\sigma^2_{jk0}}}{\frac{1}{\sigma^2_{jk0}} + \frac{t}{\sigma^2_\epsilon}} + \overline{a}_{ijkt}\frac{\frac{t}{\sigma^2_\epsilon}}{\frac{1}{\sigma^2_{jk0}} + \frac{t}{\sigma^2_\epsilon}} \tag{3}$$

$$\widetilde{\sigma}^2_{ijkt} = \frac{1}{\frac{1}{\sigma^2_{jk0}} + \frac{t}{\sigma^2_\epsilon}} \tag{4}$$

and $\overline{a}_{ijkt}$ is the average signal value, $\overline{a}_{ijkt} = \left(\frac{1}{t}\sum_t a_{ijkt}\right)$. Note that contrary to $\widetilde{\theta}_{ijkt}$, the posterior variance $\widetilde{\sigma}^2_{ijkt}$ does not depend on the realizations of the signals and decreases only with the number of signals (i.e. learning reduces uncertainty). Hence, the posterior variance is always smaller than the prior variance, $\widetilde{\sigma}^2_{ijkt} < \widetilde{\sigma}^2_{ijkt-1}$.

In the following, it will be useful to formulate the Bayesian updating recursively. Denoting $\Delta\widetilde{\theta}_{ijkt} = \widetilde{\theta}_{ijkt} - \widetilde{\theta}_{ijkt-1}$, we have:

$$\Delta\widetilde{\theta}_{ijkt} = g_t\left(a_{ijkt} - \widetilde{\theta}_{ijkt-1}\right) \text{ with } g_t = \frac{1}{\frac{\sigma^2_\epsilon}{\sigma^2_{jk0}} + t}. \tag{5}$$

Intuitively, observing a higher-than-expected signal, $a_{ijkt} > \widetilde{\theta}_{ijkt-1}$ leads the agent to revise the expectation upward, $\widetilde{\theta}_{ijkt} > \widetilde{\theta}_{ijkt-1}$, and vice versa. This revision is large when $g_t$ is large, which happens when $t$ is small, i.e. when the firm is "young" in market $jk$.

[13] We could further assume, leaving our results fully unchanged, that the variance of the prior beliefs is firm specific, i.e. $\sigma^2_{ijk0}$. We would need to assume in that case that this firm specific variance is independent from firm characteristics.

## 3.2 Firm size and belief updating

Firms maximize expected profits, subject to demand. Labelling $G_{t-1}(a_{ijkt})$ the prior distribution of $a_{ijkt}$ at the beginning of period $t$ (i.e. the posterior distribution after having observed $t-1$ signals), firm $i$ maximizes:

$$\max_q \int \pi_{ijkt} dG_{t-1}(a_{ijkt}) \qquad \text{s.t.} \qquad p_{ijkt} = \left( \frac{\mu_k Y_{jt} e^{a_{ijkt}}}{q_{ijkt} P_{jkt}^{1-\sigma_k}} \right)^{\frac{1}{\sigma_k}}. \tag{6}$$

Here, we assume for simplicity that aggregate market conditions at time $t$, i.e. $\mu_k Y_{jt}/P_{jkt}^{1-\sigma_k}$, are observed by firms before making their quantity decision. This leads to the following optimal quantities and prices (see appendix):

$$q^*_{ijkt} = \left( \frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}} \right)^{-\sigma_k} \left( \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} \right) \left( \mathbb{E}_{t-1}\left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] \right)^{\sigma_k} \tag{7}$$

$$p^*_{ijkt} = \left( \frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}} \right) \left( \frac{e^{\frac{a_{ijkt}}{\sigma_k}}}{\mathbb{E}_{t-1}\left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right]} \right) \tag{8}$$

with $\mathbb{E}_{t-1}[e^{\frac{a_{ijkt}}{\sigma_k}}] = \int e^{\frac{a_{ijkt}}{\sigma_k}} dG_{t-1}(a_{ijkt})$. As firm $i$ makes a quantity decision before observing demand for its product, $q^*_{ijkt}$ depends on expected demand, not on demand realization, contrary to $p^*_{ijkt}$.

The literature has typically computed correlations between firm age and firm growth rates, and attributed negative ones as potential evidence for a learning mechanism. Indeed the fact that younger firms adjust more their beliefs leads growth rate to decrease with age in absolute value. But of course, as is clear from equations (7) and (8), firm size and therefore firm growth (would it be measured in terms of employment or sales) also depend on the evolution of market-specific conditions and firm productivity, which could be correlated with firm age. Directly testing for the presence of demand learning thus requires either making assumptions about the dynamics of aggregate market conditions and firm productivity or finding a way to account for them. Our methodology follows the second route.

Let us now decompose optimal quantities and prices into three components. They first depend on unit costs, which are a function of wages in country $i$ and firm-product specific productivity $\varphi_{ikt}$. This first component is $ikt$-specific, i.e. is independent of the destination served; we label it $C_{ikt}$. Second, they depend on aggregate market conditions, which are common to all firms selling product $k$ to destination $j$. We label this component $C_{jkt}$. Finally, they depend on the firm $i$ beliefs about expected demand in $j$ for its product $k$ and on the demand shock at time $t$. This last composite term – labelled $Z_{ijkt}$ – is the only one to be impacted by firm learning about its demand in a specific destination market: it is $ijkt$-specific. We can now rewrite the above expressions for quantities and prices as:

$$q^*_{ijkt} = C^q_{ikt} C^q_{jkt} Z^q_{ijkt} \tag{9}$$
$$p^*_{ijkt} = C^p_{ikt} Z^p_{ijkt}. \tag{10}$$

The impact of demand learning is fully included in the $Z^q_{ijkt}$ and $Z^p_{ijkt}$ terms. These terms can be understood as the quantity and price of firm $i$ for product $k$ on market $j$ at time $t$, purged from firm unit costs and aggregate market conditions, and may be very different from the actual firm size and firm price. From a methodological point of view, any prediction about firm demand learning should be based on these $Z_{ijkt}$ terms rather than the actual $q^*_{ijkt}$ and $p^*_{ijkt}$. This also means that we will not look at the dynamics of firm size (at least per se), but directly at the dynamics of the firms' beliefs about demand. Their growth rate can be expressed as:[14]

$$\Delta \ln \mathbb{E}_t \left[ e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] = \frac{g_t}{\sigma_k} \left( a_{ijkt} - \widetilde{\theta}_{ijkt-1} \right) - \frac{g_t}{\sigma_k} \frac{\widetilde{\sigma}^2_{ijkt-1}}{2\sigma_k}. \tag{11}$$

At the beginning of period $t$, firms make quantity decisions based on their beliefs about local demand for their product ($\widetilde{\theta}_{ijkt-1}$). Then, demand is realized ($a_{ijkt}$) and firms update their beliefs. A higher than expected demand leads the firm to update upwards its belief. The opposite is true for a lower than expected demand. Importantly, as is clear from equation (11), this upward or downward updating is larger for younger firms. It follows our main prediction:

**Prediction # 1** (updating and age): *A given difference between realized and expected demand leads to a larger updating of the belief, the younger the firm is.*

It is also interesting to note that larger uncertainty (i.e. a higher $\sigma^2_\epsilon$) reduces the extent of belief updating and the effect of age on belief updating. This is because a signal is less informative when uncertainty is higher. Put differently, the information contained in the realized price will be noisier when $\sigma^2_\varepsilon$ is large, in which case firms will adjust less their beliefs in the next period. This is our second prediction:

**Prediction # 2** (updating and uncertainty): *A higher level of market uncertainty reduces the extent of beliefs updating, and the effect of age on belief updating.*

In the next section, we derive our methodology to isolate the $Z^q_{ijkt}$ and $Z^p_{ijkt}$ terms and distinguish the beliefs from the demand shock component.

## 3.3 Identification and measurement

**Identifying beliefs.** In order to isolate $Z^q_{ijkt}$ and $Z^p_{ijkt}$, we need to purge supply side and market specific factors from actual quantities and prices. This is achieved by estimating

[14] Detailed derivations and proofs of all our propositions are relegated to the appendix.

the following quantity and price equations in logs:[15]

$$\ln q_{ijkt} = \mathbf{FE}_{ikt} + \mathbf{FE}_{jkt} + \varepsilon^q_{ijkt} \tag{12}$$
$$\ln p_{ijkt} = \mathbf{FE}_{ikt} + \varepsilon^p_{ijkt} \tag{13}$$

where $k$ is a 6-digit product and $t$ is a year. $\mathbf{FE}_{ikt}$ and $\mathbf{FE}_{jkt}$ represent respectively firm-product-year and destination-product-year fixed effects. Note that we do not have direct price data, so we rely on unit values, defined as $S_{ijkt}/q_{ijkt}$, where $S_{ijkt}$ denote firms sales, to proxy them. In our baseline estimations, we stick to the model and estimate the price equation without the $jkt$ fixed effects, as implied by the CES assumption. In section 5.1 we discuss the implications of relaxing the CES assumption, one of them being that we need to control for market-specific conditions in the price equation.

The estimates of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ are estimates of the $Z_{ijkt}$ terms. Using (7) and (8), we get:

$$\varepsilon^q_{ijkt} = \ln Z^q_{ijkt} = \sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] \tag{14}$$
$$\varepsilon^p_{ijkt} = \ln Z^p_{ijkt} = \frac{1}{\sigma_k} a_{ijkt} - \ln \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]. \tag{15}$$

This identification strategy is possible to implement because we are able to observe the sales of the same product by the same firm in different destination markets, which allows purging market-specific firm dynamics from the evolution of firm productivity through the inclusion of $\mathbf{FE}_{ikt}$.[16] As we account for all time-varying, market- and firm-product-specific determinants of quantities and prices, our approach could accommodate any underlying dynamic process for the $ikt$ and $jkt$ terms. This includes processes driving the evolution of firm productivity, but also any other time-varying, firm-specific factors that might affect firm dynamics, such as financial constraints, as well as variations in market-specific trade costs.

Consistently estimating the residuals of (12) and (13) however requires some identification assumptions. In particular, $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ need to be orthogonal to firm characteristics $\{w_{it}, \varphi_{ikt}\}$, and $\varepsilon^q_{ijkt}$ must also be orthogonal to market conditions $\{Y_{jt}, P_{jkt}\}$. This implies that beliefs do not vary systematically with productivity, or, in other words, that initial beliefs must be unbiased also along the firm productivity dimension. This rules out the possibility that firms engage in overall productivity-enhancing investments because they have higher beliefs in a given market. Note however that our identification strategy does not preclude firms to modify a *market-specific* productivity component in response to changes

[15] We use the Stata routine `reghdfe` developed by Sergio Correia, based on Guimaraes and Portugal (2010).

[16] The reason why we do not model learning about productivity appears more clearly in equations (14) and (15). Identifying demand variations is possible because we are able to control for productivity through the inclusion of $ikt$ fixed effects. On the other hand, we cannot distinguish productivity variations from global demand shocks faced by firms in all the markets, as these would be mixed with unit costs in the $\mathbf{FE}_{ikt}$.

in their information set. In section 5.1, we thus allow productivity to differ across destinations for a given firm-product. The condition on $\varepsilon^p_{ijkt}$ also implies that demand signals $a_{ijkt}$ must be orthogonal to firms' overall costs $\{w_{it}, \varphi_{ikt}\}$. Put differently, we make the standard assumption that firms with high productivity do not enjoy higher market specific demand beyond the effect of their productivity on demand through lower prices.

These orthogonality restrictions also reflect our assumption that beliefs are market-specific, i.e. that firms do not adjust their beliefs to information arriving from other markets. As mentioned in the introduction, in theory there could be spillovers taking many different forms: beliefs could depend on the experience accumulated by the firm in selling the same product to other destinations, including the domestic market. They could also vary with the information obtained when selling other products in the same market. Studying such informational spillovers is beyond the scope of this paper. Yet, we are confident that the information we capture is indeed market-specific. The reason is that our identification strategy *de facto* constrains the set of possible determinants of beliefs. For instance, if these are partly determined by past domestic market experience for the same product, or by past experience in other markets for the same product, then the $ikt$ fixed effects will account for them. In other words, $\varepsilon^q_{ijkt}$ captures the firms' beliefs net of the effect of experience in other markets at time $t$.[17]

Finally, a note on our interpretation of the residuals (14) and (15). Following the model, we consider that these residuals reflect the demand-side components of prices and quantities. Our identification assumption is that, within a given firm, costs can differ across products but not across products *and* destinations. Note however that we allow variations in costs across markets for a given product. These include in particular trade costs and potential differences in demand for quality and are captured by $\mathbf{FE}_{jkt}$. In section 5.1 we allow productivity to be market-specific and show that we can still consistently estimate the demand shocks. We do not, however, allow firms to learn about market-specific costs. As discussed in section 6, the evidence we find on the profiles of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ is more consistent with firms learning about demand than about costs, but we cannot exclude that firms are learning about demand shifters such as market-specific trade costs. Such a learning process would be isomorphic to learning about demand. We favor the traditional demand learning formulation, yet what we call demand learning could be encompassing learning about demand-shifters.

**Identifying demand shocks.** Testing prediction 1 requires getting estimates of the demand signals $a_{ijkt}$. Because the firm takes its quantity decision before observing the

[17] Our methodology does not, on the other hand, takes into account the possibility that beliefs depend on the information gathered by the firm while selling other products in the same destination. This would require including $ijt$ fixed effects in equations (12) and (13). We have tried to include these and our estimates were largely unaffected (see section I.5 in the online appendix). This lends support to our assumption that information is indeed mostly product-market specific: if shocks and beliefs were correlated across products within destinations, the firms' response to a demand shock would partly reflect its belief updating behavior on other products, and including $ijt$ fixed effect should dampen the extent of estimated belief updating.

demand realization, $\ln Z^q_{ijkt}$ depends on the firms' beliefs about demand only, while $\ln Z^p_{ijkt}$ is adjusted for the demand shock (an assumption that we discuss in section 5.1). Thus, the residual $\varepsilon^q_{ijkt}$ provides a direct estimate of the firms' beliefs. We only need to correct for $\sigma_k$. In order to back out the demand shock and get an estimate of $\sigma_k$, we regress $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$. Using (15) and (14), we get:

$$\left(\frac{1}{\sigma_k}a_{ijkt} - \ln \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right) = \beta\left(\sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right) + \lambda_{ijk} + v_{ijkt}. \tag{16}$$

We need to include firm-product-destination fixed effects $\lambda_{ijk}$ to account for the fact that $a_{ijkt} = \overline{a}_{ijk} + \varepsilon_{ijkt}$. Omitting these fixed effects would generate inconsistent estimates of $\beta$ as both $v_{ijkt}$ and the firm beliefs $\mathbb{E}_{t-1}\left[\exp(\frac{a_{ijkt}}{\sigma_k})\right]$ would depend on $\overline{a}_{ijk}$, which would violate the zero conditional mean assumption.[18] Including $\lambda_{ijk}$ allows to take out $\overline{a}_{ijk}$ from the error term $v_{ijkt}$ and recover consistent estimates of $\beta$. We estimate (17) by 6-digit product to allow $\sigma_k$ to differ across products and obtain:[19]

$$\widehat{\beta} = -\frac{1}{\sigma_k} \quad ; \quad \widehat{\lambda_{ijk}} + \widehat{v_{ijkt}} \equiv \widehat{a_{ijkt}} = \frac{1}{\sigma_k}a_{ijkt} \quad ; \quad \widehat{v}_{ijkt} = \frac{1}{\sigma_k}\varepsilon_{ijkt}. \tag{17}$$

Note that the level of uncertainty can be directly inferred from our estimates of demand signals. We define market-specific uncertainty as the standard deviation of $a_{ijkt}$, computed by product-and-destination, over our data period.

The last variable we need to test our predictions is market-specific firm age, which has been defined in section 2. Age is either constructed as a single discrete variable or as a set of dummies, to allow the learning processes to be non-linear.

**Testing prediction #1.** We can now derive our testable equation. Equation (11) cannot be tested directly as we do not observe $\widetilde{\theta}_{ijkt-1}$ but only $\varepsilon^q_{ijkt}$. We make use of (11), (5) and (14), to get the following specification (see the appendix):

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t\frac{\sigma^2_\varepsilon}{2\sigma_k}. \tag{18}$$

This equation is equivalent to (11), except that it can be tested: our estimates of $\varepsilon^q_{ijkt}$ comes from (12), and $a_{ijkt}$ is computed from equation (17) as the product of $\widehat{a_{ijkt}}$ times $\sigma_k$. $g_t$ is an inverse function of market-specific age (equation (5)). We estimate:

$$\Delta\varepsilon^q_{ijkt+1} = \sum_{g=2}^{G}\alpha_g(a_{ijkt} - \varepsilon^q_{ijkt}) \times \text{AGE}^g_{ijkt} + \sum_{g=1}^{G}\beta_g\text{AGE}^g_{ijkt} + u_{ijkt} \tag{19}$$

[18] We thank an anonymous referee for pointing out this issue.

[19] Whenever our estimates of $\beta$ are statistically insignificant or imply values of $\sigma_k$ which are lower than 1, we replace $\widehat{v}$ by a missing value and do not consider the observation in the estimations. Note that our results are insensitive to such cleaning of the data. $\sigma_k$ is lower than 1 for only 0.01% of observations, and insignificant $\beta$ coefficients (at the 5% level) are obtained for 1.6% of observations. See the upper panel of Table A.4 in the online appendix. We also perform a robustness exercise where equation (17) is estimated at the 4-digit instead of 6-digit level to end up with more observations by product and more efficient estimations.

where $AGE^g_{ijkt}$ are dummies taking the value 1 for each age category $g = 2, ...10$ representing the number of years of presence in the export market (e.g. $g = 2$ in the second year of presence). Standard errors are robust to heteroscedasticity and clustered by firm (or, alternatively, bootstrapped). We expect $\alpha_g$ to be positive on average, and $\beta_g$ to be decreasing with age. Our main prediction is that $\alpha_g$ decreases with age $g$. Note that equation (18) predicts that $\alpha_g = g_t = \frac{1}{\sigma^2_\varepsilon/\sigma^2_{jk0}+t}$ with $g_t$ measuring the speed of learning. Hence, the evolution of the $\alpha_g$ coefficients with firm age allow to assess how firms learn about their demand parameter.

Our test of the passive learning mechanism therefore builds on the evidence that firms adjust their quantities to past demand shocks *and* that such a reaction gets smaller as firms grow older in a market. This decline of the quantity reaction to past demand shocks is a distinctive feature of the learning process. If firms had full information about demand, stochastic *iid* shocks should not generate any quantity reaction beyond the current period as these shocks would not provide any information. In that case, the coefficients $\alpha_g$ should be equal to zero. If instead shocks were persistent, firms would always adjust their next period quantities in the same way: the $\alpha_g$ coefficients would be positive but constant over time.

# 4 Main results

In this section, we start by providing some descriptive statistics about our final sample, before discussing the results obtained when testing prediction 1. We then study how market uncertainty affect the characteristics of the learning process.

## 4.1 Sample statistics

Table 1 contains some descriptive statistics about our final sample. Firms are typically young in the markets they serve: the average age is comprised between 3.5 and 3.8 years depending on the definition (note that since we focus on $\Delta\varepsilon^q_{ijkt}$ in the following, firms that exit during the first year are dropped and 2 is the minimum value that our age variable can take). This is evidence of the low survival rates observed during the first years a firm serves a particular market (Figure 1.a). Over the period, the firm-market specific beliefs have been characterized by a positive average growth, while $\Delta\varepsilon^p_{ijkt}$ is slightly negative on average.

Our methodology generates reasonable estimates of $\sigma_k$: we get a median value of 5.1 and an average of 6.2 in our final sample. These numbers are comparable to the ones found by the literature, using very different methodologies and data.[20] Our estimates of $\sigma_k$ also follow expected patterns: considering Rauch (1999)'s classification, the median (resp. mean) across products is 5.2 (resp. 6.1) for differentiated goods, 7.3 (resp. 8.6) for referenced priced goods and 8.9 (resp. 10.1) for goods classified as homogenous. These

[20]See Imbs and Mejean (2015) for a detailed literature review.

Table 1: Sample statistics

| | Obs. | Mean | S.D. | Q1 | Median | Q3 |
|---|---|---|---|---|---|---|
| $\ln q_{ijkt}$ | 4382989 | 6.237 | 2.795 | 4.277 | 6.004 | 8.001 |
| $\ln p_{ijkt}$ | 4382989 | 3.115 | 1.969 | 1.808 | 3.058 | 4.358 |
| $\Delta\varepsilon^q_{ijkt}$ | 1854141 | 0.030 | 1.200 | -0.631 | 0.026 | 0.687 |
| $\Delta\varepsilon^p_{ijkt}$ | 1854141 | -0.002 | 0.672 | -0.224 | -0.000 | 0.221 |
| $a_{ijkt} - \varepsilon^q_{ijk.t-1}$ | 1854141 | -0.052 | 3.425 | -1.340 | -0.041 | 1.180 |
| $a_{ijkt}$ | 1854141 | -0.003 | 0.562 | -0.261 | 0.001 | 0.256 |
| $\sigma_k$ | 1854141 | 6.205 | 4.783 | 3.566 | 5.089 | 6.593 |
| $\text{sd}(a_{ijkt})$ | 1848126 | 2.603 | 1.493 | 1.895 | 2.267 | 2.802 |
| $\text{Age}^1_{ijkt}$ | 1854141 | 3.505 | 1.800 | 2 | 3 | 4 |
| $\text{Age}^2_{ijkt}$ | 1854141 | 3.671 | 1.851 | 2 | 3 | 5 |
| $\text{Age}^3_{ijkt}$ | 1854141 | 3.759 | 1.851 | 2 | 3 | 5 |

Source: Authors' computations from French Customs data. $\ln q_{ijkt}$ and $\ln p_{ijkt}$ are the logs of quantities and prices sold by a firm $i$ in a market $jk$ a given year $t$. $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ are respectively the belief of the firm about future demand from equation (14) and the residuals of the price equation from equation (15). $\text{Age}^1_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). $\text{Age}^2_{ijkt}$: reset after 2 years of exit; $\text{Age}^3_{ijkt}$: years of exporting since first entry (never reset to zero). $a_{ijkt}$ is our estimate of the demand shock from equation (17). $\sigma_k$: elasticity of substitution from equation (17). $\text{sd}(a_{ijkt})$ is the standard deviation of $a_{ijkt}$, computed by market (product-destination).

means and medians of $\sigma_k$ are statistically different across the three groups.[21]

## 4.2 Baseline results

The results obtained when estimating equation (19) are provided in Table 2. The first column considers separately the effect of demand shocks and age on changes in firms' beliefs. Columns (2) to (7) study how age affects the reaction of beliefs to demand shocks. Columns (3), (5) and (7) are equivalent to columns (2), (4) and (6) with standard errors being bootstrapped rather than clustered by firm, to account for the fact that the right hand side variables have been estimated.

As predicted, firms update their beliefs positively when they face a higher than expected demand, and the growth in beliefs declines with age on average (column (1)). More importantly, we find support for our key prediction: belief updating following a demand shock is significantly stronger when firms are young (columns (2)-(7)). Including age linearly (column (2) and (3)) or through bins (columns (4) to (7)) leads to the same conclusion. Bootstrapping the standard errors also leaves the results unaffected.

After a decade of presence in the market, the magnitude of belief updating following a given demand shock is 30 percent smaller than after entry. In columns (6) and (7), we find that, when compared to the benchmark category – age of ten years –, the coefficients of the first four years (first six years with bootstrapped standard errors) are significantly

[21] See section B of the online appendix for details. Note that these numbers are slightly higher than the means and medians displayed in Table 1 because they are computed across products, while the statistics in Table 1 are based on our final sample, i.e. also reflect the number of French firms selling each product.

Table 2: Prediction 1: demand shocks and beliefs updating

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. var. | | | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.064$^a$ | 0.074$^a$ | 0.074$^a$ | | | 0.047$^a$ | 0.047$^a$ |
| | (0.001) | (0.001) | (0.001) | | | (0.007) | (0.006) |
| $\times$ Age$_{ijkt}$ | | -0.003$^a$ | -0.003$^a$ | | | | |
| | | (0.000) | (0.000) | | | | |
| $\times$ Age$_{ijkt} = 2$ | | | | 0.069$^a$ | 0.069$^a$ | 0.022$^a$ | 0.022$^a$ |
| | | | | (0.001) | (0.001) | (0.007) | (0.005) |
| $\times$ Age$_{ijkt} = 3$ | | | | 0.064$^a$ | 0.064$^a$ | 0.017$^b$ | 0.017$^a$ |
| | | | | (0.001) | (0.001) | (0.007) | (0.005) |
| $\times$ Age$_{ijkt} = 4$ | | | | 0.060$^a$ | 0.060$^a$ | 0.013$^c$ | 0.013$^b$ |
| | | | | (0.002) | (0.001) | (0.007) | (0.005) |
| $\times$ Age$_{ijkt} = 5$ | | | | 0.057$^a$ | 0.057$^a$ | 0.010 | 0.010$^b$ |
| | | | | (0.002) | (0.002) | (0.007) | (0.005) |
| $\times$ Age$_{ijkt} = 6$ | | | | 0.058$^a$ | 0.058$^a$ | 0.011 | 0.011$^b$ |
| | | | | (0.002) | (0.002) | (0.007) | (0.005) |
| $\times$ Age$_{ijkt} = 7$ | | | | 0.054$^a$ | 0.054$^a$ | 0.007 | 0.007 |
| | | | | (0.002) | (0.002) | (0.006) | (0.005) |
| $\times$ Age$_{ijkt} = 8$ | | | | 0.052$^a$ | 0.052$^a$ | 0.004 | 0.004 |
| | | | | (0.003) | (0.002) | (0.006) | (0.005) |
| $\times$ Age$_{ijkt} = 9$ | | | | 0.052$^a$ | 0.052$^a$ | 0.005 | 0.005 |
| | | | | (0.004) | (0.003) | (0.006) | (0.005) |
| $\times$ Age$_{ijkt} = 10$ | | | | 0.047$^a$ | 0.047$^a$ | | |
| | | | | (0.007) | (0.005) | | |
| Age$_{ijkt}$ | -0.033$^a$ | -0.033$^a$ | -0.033$^a$ | | | | |
| | (0.001) | (0.001) | (0.000) | | | | |
| Observations | 1854141 | 1854141 | 1854141 | 1854141 | 1854141 | 1854141 | 1854141 |

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3), (5) and (7)). $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (4) to (7) but coefficients not reported. Columns (6) and (7) are the same as column (4) and (5) except that coefficients are estimated relative to the baseline omitted category, age of ten years. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit).

higher. The shape of the learning process is consistent with the theory: age has a strong effect in early years, and matters less for more experienced firms (section C in the online appendix provides a graphical depiction of the result and a discussion of our functional form assumption). Note that most of our estimated coefficients are statistically different from each other up to year seven, which supports the existence of a learning process over this time horizon. After seven years, our results no longer provide clear evidence of learning (note however that the coefficient of the last category is less precisely estimated due to the small number of observations). However, even the most experienced firms in our sample still significantly adjust their quantities following demand shocks. Assuming that part of the demand signals received is persistent would explain this finding: in that case, experienced firms would continue to adjust their quantities to demand shocks even if they have fully discovered their idiosyncratic demand.

## 4.3 Learning and market uncertainty

Our second prediction is that a higher level of uncertainty in the market (a higher $\sigma_\varepsilon^2$ in the model) should slowdown the updating process. The underlying intuition is that a demand signal is less informative when uncertainty is higher. It follows that the speed at which firms update their beliefs should decrease with age, but less so when uncertainty is larger (see proof of prediction 1 in the appendix).

Table 3: Prediction 1: the role of uncertainty

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | $\Delta\varepsilon^q_{ijk,t+1}$ | | | |
| Sample | | | High Uncertainty | Low Uncertainty |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.102^a$ | $0.113^a$ | $0.054^a$ | $0.163^a$ |
| | (0.006) | (0.005) | (0.001) | (0.002) |
| $\times$ Age$_{ijkt}$ | $-0.003^a$ | $-0.006^a$ | $-0.002^a$ | $-0.007^a$ |
| | (0.000) | (0.001) | (0.000) | (0.000) |
| $\times$ Uncertainty | $-0.005^a$ | $-0.006^a$ | | |
| | (0.001) | (0.001) | | |
| $\times$ Age$_{ijkt}\times$ Uncertainty | | $0.001^a$ | | |
| | | (0.000) | | |
| Age$_{ijkt}$ | $-0.033^a$ | $-0.027^a$ | $-0.037^a$ | $-0.028^a$ |
| | (0.000) | (0.001) | (0.001) | (0.001) |
| Uncertainty | $-0.004^b$ | $0.004^c$ | | |
| | (0.001) | (0.002) | | |
| Age$_{ijkt}\times$ Uncertainty | | $-0.002^a$ | | |
| | | (0.000) | | |
| Observations | 1848126 | 1848126 | 928963 | 919146 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Uncertainty is the standard deviation of $a_{ijkt}$, computed by market $jk$. High and low uncertainty mean above and below sample median.

We use our theory-based measure of market uncertainty (the standard deviation of $a_{ijkt}$, computed by product and destination over the entire period). We then add to specification (19) an interaction term between our uncertainty measure and $(a_{ijkt} - \varepsilon^q_{ijkt})$, and a triple interaction between age, $(a_{ijkt} - \varepsilon^q_{ijkt})$ and uncertainty (as well as an interaction term between age and uncertainty). Table 3 contains the results. Column (1) shows that, as predicted, the extent of belief updating following a demand shock is smaller in markets characterized by a higher level of uncertainty. On the other hand, the coefficient on the interaction term between age and the demand shocks is virtually unaffected. Quantitatively, the role of uncertainty is non negligible. A standard deviation increase from the mean of the level of uncertainty decreases the response of beliefs to demand shocks from 0.090 to 0.082 in column (1).

Moreover, when uncertainty is large, gaining experience has a lower effect on belief updating, as shown by the coefficient of the triple interaction term in column (2). Another way to represent these results is to separate the sample into high and low uncertainty markets, defined according to the sample median of our uncertainty measure. We run our baseline specification (column (4) of Table 2) separately on each of the two sub-samples. The results are displayed in column (3) and (4) of Table 3. We clearly see that the average extent of belief updating is much larger in markets with low uncertainty levels, and that updating decreases more with age in the least uncertain markets. In the online appendix, section D we use bins of age categories and a more extreme sample split (first and last quartile of uncertainty). In these specifications, we find that the updating coefficient decreases from 0.171 in the second year to 0.128 after ten years in the least uncertain markets, while in the most uncertain markets the relationship is flatter and updating is almost nonexistent as the coefficients decrease from 0.035 to 0.021.

# 5 Robustness

In this section we first assess the implications of several key assumptions of our model for our identification strategy and the interpretation of our results. We then discuss how our results might be affected by endogenous exit, before considering a series of additional sensitivity tests.

## 5.1 Modelling assumptions

Our model makes three important assumptions. First, firms set their quantities before observing the demand realization, as in Jovanovic (1982). Second, firms face CES demand and monopolistic competition (hence markups are constant). Third, firm productivity is not market-specific. In this subsection we assess the sensitivity of our results to these hypotheses (we discuss the validity of our demand-side modelling of learning in section 6). In particular, we show how they affect (i) the identification of beliefs and demand signals and (ii) our test of prediction 1. Relaxing these assumptions implies in general that the residuals $\varepsilon^q_{ijkt}$ can no longer be interpreted as reflecting beliefs only. However – and provided that we control for market-specific firm size in some cases –, these extensions do not alter the qualitative interpretation of our results, in the sense that our baseline estimates of Table 2 can still be viewed as evidence of belief updating. For each extension, we summarize here the main intuitions and refer the reader to the online appendix E for details.

#### 5.1.1 Fixed quantities

We have assumed so far that quantities are set before firms observe their idiosyncratic demand in each market, while prices adjust to the demand shocks. We relax this assumption in two directions: we start by considering the possibility that prices are set first, with or without a constant price elasticity. Second, we assume that firms can adjust their quantity decision after observing part of the demand shock.

If we completely reverse our assumption and suppose that prices are set *ex-ante* while quantities fully adjust to demand shocks, due to CES demand, prices will only depend on supply side characteristics. They take the form of a constant mark-up over marginal costs and do not vary with the quantity produced, the firm's beliefs or the demand shock. Quantities on the other hand fully adjust and depend solely on the demand shocks. Regressing $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$ should therefore generate insignificant $\widehat{\beta}$ coefficients, and $\varepsilon^q_{ijkt}$ should not vary with age. Both these predictions are clearly at odds with our findings.

Now, assume that prices are set *ex-ante* but the market structure is oligopolistic, which implies variable markups. In this case, prices reflect the firm's beliefs, as markups depend on its expected market share. Quantities reflect both these beliefs and the demand shocks. We can still estimate demand signals, but our identification strategy should be reversed: $\varepsilon^q_{ijkt}$ should be regressed on $\varepsilon^p_{ijkt}$, and the updating process should be observed on $\Delta\varepsilon^p_{ijkt}$. The main prediction of such a model is that a positive demand shock should lead firms to update upwards their beliefs, which would increase their markup and their prices. In the online appendix (section E.2), we follow this alternative methodology and find that prices slightly *decrease* with demand shocks, which is inconsistent with this alternative model of Bertrand competition with a non-constant price elasticity.

Finally, we consider an intermediate case where firms can revise their quantity decision after observing part of the demand shock. In this case, our theoretical predictions still hold, but the identification of the demand shock is affected: $\varepsilon^q_{ijkt}$ now also captures part of the demand shock and becomes a noisy measure of the firm's belief. This may affect our estimates of the demand shocks, although the direction of this bias is unclear. Yet, unless this bias is correlated with age, our main results that young firms update more their beliefs should not be affected. One way to gauge the importance of this possible bias is to focus on sectors or destinations for which quantities are more likely to be rigid – i.e. those for which the demand shocks are more likely to be correctly estimated – and to compare the results with our baseline estimates of Table 2. We expect less quantity adjustment for complex goods (in which many different relationship-specific inputs are used in the production process) and in destinations characterized by longer time-to-ship. In section E.3 of the online appendix, we restrict our sample to sectors or destinations which are above the sample median in terms of time-to-ship or input complexity. The estimated magnitude of belief updating and the coefficient on the interaction terms between demand shocks and age are quantitatively similar to our baseline estimates.[22] Altogether, these results suggest that our assumption of fixed quantities is not unrealistic and does not lead our identification strategy to artificially generate our results.

### 5.1.2 Other extensions and control for size

Our next two extensions allow respectively for variable mark-ups and for productivity to be market-specific. We reach similar conclusions in both cases. $\varepsilon^q_{ijkt}$ can no longer be interpreted as beliefs about demand only – it is also affected by mark-ups or productivity. $\Delta\varepsilon^q_{ijkt+1}$ therefore reflects changes in beliefs as well as variations in mark-ups or productivity. The key point, however, is that we are still able to interpret the reaction of firms to demand shocks as evidence of belief updating, provided that we control for size. The complete derivations are provided in sections E.4 and E.5 of the online appendix.

**Variable mark-ups**. The first implication of variable mark-ups for our empirical strategy is that prices could now depend on local market conditions, i.e. the price equation (13) should include a set of $jkt$ fixed-effects. Columns (1) and (2) of Table 4 at the end of this section shows that this modification leaves our results largely unchanged.

Second and more importantly, the quantities residuals $\varepsilon^q_{ijkt}$ should now capture the firms' beliefs, but also their expected markups. Hence, changes in expected mark-ups should affect $\Delta\varepsilon^q_{ijkt+1}$. To take into account this possibility, we extend the model to an oligopolistic market structure. Formally, we simply assume that the number of competitors in each sector $K$, $\Omega_{kt}$, is small enough so that each competitor takes into account the impact of his own decisions on the sectoral price index. As shown in the online appendix, our methodology still produce unbiased estimates of the demand shock. Our main equation however becomes:

$$\Delta\varepsilon^q_{ijkt+1} = g_t \left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t \frac{\sigma^2_\varepsilon}{2\sigma_k} - \sigma_k g_t \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right) - \sigma_k \Delta \ln\left(\frac{\mathbb{E}_t\left[\varepsilon(s_{ijkt+1})\right]}{\mathbb{E}_t\left[\varepsilon(s_{ijkt+1})\right]-1}\right)$$

where $\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]$ is the expected elasticity of demand faced by firm $i$ in market $jk$ at the beginning of period $t$, which itself depends on the expected market share $\mathbb{E}_{t-1}\left[s_{ijkt}\right]$. With variable mark-ups, our main equation includes two new terms.

The first term is the level of the expected mark-ups. It comes from the fact that the expected mark-up also affects our measure of beliefs, $\varepsilon^q_{ijkt}$, and in turn $\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)$. We thus need to control for firm size/market share to avoid a standard omitted variable bias.

The second term captures the change in expected mark-ups, and it depends on the updating process through the change in the expected market share. Our measure of belief updating is now underestimated: when firms update positively, they tend to increase their

[22] The coefficient on the interaction term between demand shocks and age is slightly lower than our baseline in the case of complex goods (col. (5) of Table A.7). In column (6), however, we see that this result is only driven by the effect of the last age category, 10 years of experience, which is itself quite imprecisely estimated.

quantities but also their prices, which dampens their overall quantity reaction. It follows that in the case of variable mark-ups, $\varepsilon^q_{ijkt}$ becomes an increasing function of firm's beliefs[23] and we only capture the overall reaction of purged quantities to belief updating. Put differently, our results still provide evidence for the updating process, but in a qualitative sense.

Importantly, two firms of different sizes may not have the same mark-up reaction to a given belief update. This is another reason to control for market share: to be able to compare the extent of updating of firms of different age, but with the same market share.

**Product-destination productivity**. In the model, we have assumed that productivity was firm-product-specific. Here we relax this assumption and consider the case of product-destination-specific productivity. This again introduces a new source of dynamics in $\varepsilon^q_{ijkt}$. We assume that the unit cost of producing good $k$ for market $j$ at time $t$ is $\frac{w_{it}}{\varphi_{ikt}}\frac{1}{\varphi_{ijkt}}$. This could reflect differences in productivity for the same good across markets, but also differences in product quality. Again, our methodology still produces unbiased estimates of the shock, as shown in the online appendix. But the dynamics of quantities now also reflects the evolution of $\varphi_{ijkt}$. We get:

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t\frac{\sigma^2_\varepsilon}{2\sigma_k} + \sigma_k g_t \ln\left(\varphi_{ijkt}\right) + \sigma_k\Delta\ln\left(\varphi_{ijkt+1}\right).$$

As for the case of variable mark-ups, because $\varepsilon^q_{ijkt}$ contains a new element, our equation now has two additional terms: one in level because $\ln\left(\varphi_{ijkt}\right)$ alters our measure of beliefs, and one in difference $\Delta\ln\left(\varphi_{ijkt+1}\right)$, because $\Delta\varepsilon^q_{ijkt+1}$ also reflects the dynamics of productivity. Again, the first term implies that we need to control for firm size, to avoid a standard omitted variable bias. Second, the dynamics of $\ln\left(\varphi_{ijkt}\right)$ also affects $\Delta\varepsilon^q_{ijkt+1}$. If this dynamics is uncorrelated with the updating process, the interpretation of our results should be unaffected. If however $\Delta\ln\left(\varphi_{ijkt+1}\right)$ is positively affected by the updating process – if a positive updating leads firms to invest to improve $\varphi_{ijkt}$ – then our measure of updating becomes a measure of the overall impact of the updating process on $\Delta\varepsilon^q_{ijkt+1}$: it does not only capture the updating process itself but also how the quantity response is magnified by a change in productivity.[24] Again, $\varepsilon^q_{ijkt}$ would become an increasing function of firm's beliefs, and our evidence of the updating process would become qualitative as we would not identify firms' beliefs *per se*. This productivity response could be size dependent, which again requires to control for firm size. The decline of the overall response of $\Delta\varepsilon^q_{ijkt+1}$ to demand shocks over time, conditional on size, however still provides evidence

Table 4: Prediction 1: controlling for size

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. var. | | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | |
| Robustness | Controlling for FE$_{jkt}$ in prices | | Controlling for FE$_{jkt}$ in prices and size | | | |
| Size | | | Linear | | Bins | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.103$^a$ | | 0.103$^a$ | | 0.102$^a$ | |
| | (0.002) | | (0.002) | | (0.002) | |
| $\times$ Age$_{ijkt}$ | -0.003$^a$ | | -0.003$^a$ | | -0.003$^a$ | |
| | (0.000) | | (0.000) | | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | 0.096$^a$ | | 0.096$^a$ | | 0.096$^a$ |
| | | (0.002) | | (0.002) | | (0.002) |
| ... | | | | | | |
| $\times$ Age$_{ijkt} = 10$ | | 0.074$^a$ | | 0.074$^a$ | | 0.075$^a$ |
| | | (0.009) | | (0.009) | | (0.009) |
| Age$_{ijkt}$ | -0.034$^a$ | | -0.040$^a$ | | 0.019$^a$ | |
| | (0.001) | | (0.001) | | (0.002) | |
| Size$_{ijkt}$ | | | -1.053$^a$ | -1.015$^a$ | | |
| | | | (0.016) | (0.017) | | |
| $\times$ Age$_{ijkt}$ | | | 0.109$^a$ | 0.101$^a$ | | |
| | | | (0.003) | (0.003) | | |
| Observations | 1870377 | 1870377 | 1870377 | 1870377 | 1501840 | 1501840 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17). In this table, $jkt$ fixed effects are included in the estimation of the price residuals $\varepsilon^p_{ijkt}$ used to identify demand shocks. $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (17). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Size$_{ijkt}$ is proxied by the value sold by firm $i$ on market $jk$ during year $t$ divided by the total value exported by French firms in market $jk$ during year $t$. Columns (5) and (6) include size bins corresponding to the ten deciles of size variable, computed by market-year. Age dummies included alone in columns (2), (4) and (6) but coefficients not reported. See Table A.8 in the online appendix for the full set of coefficients on the interaction terms.

for an updating process.

**Controlling for size.** The two extensions of the models discussed above suggest that firm-market size should be included in our regressions, together with its interaction with firm-market age. We do so in Table 4. Columns (1) and (2) are similar to our baseline regressions (Table 2, columns (2) and (4)), except that $jkt$ fixed effects are introduced in the estimation of the price residuals $\varepsilon^p_{ijkt}$, as predicted by models with variable markups. The average level of belief updating is slightly larger than in our baseline estimates, but

[23] Formally, we derive in the online appendix two alternatives sufficient conditions ensuring that the overall quantity response to a positive updating is still positive: either $\sigma_k \geq 2$ or $s_{ijkt} \leq 1/2$.

[24] Note that this possibility does not violate the orthogonality conditions that we need to identify demand shocks. As discussed in section 3.3, we need the beliefs a firm in market $jk$ at time $t$ to be orthogonal to overall firm-product characteristics; yet, beliefs can be correlated with the characteristics of a firm-product in that particular market $j$.

the effect of age is similar. In columns (3) to (6) we additionally control for firm size, as measured by the value sold by firm $i$ on market $jk$ during year $t-1$ divided by the total value exported by French firms in market $jk$ during year $t-1$. Size is introduced either linearly in columns (3) and (4) or through bins computed using market-specific deciles in columns (5) and (6). Our coefficients of interest are extremely stable across specifications.[25] In the online appendix E.6, we consider a number of alternative measures of firm size and include interaction terms between size and $a_{ijkt} - \varepsilon^q_{ijkt}$ to account for the fact that age and size are correlated. In all instances the results are similar to our benchmark estimates.

## 5.2 Survival and selection bias

Our main prediction is tested on the sample of firms which survive in period $t$. Endogenous sample selection could be a concern in equation (19). The error term $u_{ijkt}$ might be correlated in particular with demand shocks: the observed sample includes firms with relatively positive demand shocks (as those with negative shocks are more likely to exit), and firms which do not update downward their beliefs too much following a negative signal (otherwise they would exit). In other words, endogenous exit might create a correlation between the error term of (19) and demand shocks.

The predictions of the learning model for survival are discussed in details in section F of the online appendix. We show that exit probability depends (negatively) on demand signals, age, and beliefs, as well as on the dynamics of firm productivity and market conditions. Predicted exit probabilities can therefore be estimated as a function of $a_{ijkt}$, $\varepsilon^q_{ijkt}$, $\text{Age}_{ijkt}$ and fixed effects in the $ikt$ and $jkt$ dimensions. We use a linear probability model which allows the inclusion of our two high-dimensional fixed effects. Once these survival probabilities have been estimated, we perform two different types of exercises to check that our results are not affected by endogenous selection.

First, we gauge the importance of this selection bias by estimating (19) on sub-samples defined according to the survival probability. This is an application of the "identification-at-infinity" method (Chamberlain, 1986; Mulligan and Rubinstein, 2008). The general idea is to restrict the estimation sample to firms that are most likely to survive, the selection bias being lower for firms with high survival probability. We allocate firms in 5 bins of survival probability and estimate (19) on sub-samples that include only firms above the 20th, 40th, 60th and 80th percentiles of survival probability. The results are presented in Table A.12 (section H) in the online appendix. Starting from the full sample in column (1), we progressively drop the quintiles of observations with the highest exit probabilities from the sample. Accordingly, column (5) only includes the quintiles of observations with the lowest exit probabilities (i.e. the highest survival probability). If endogenous exits were driving our results, we would expect the patterns of belief updating to substantially differ across samples. On the contrary, we find that the coefficients on $(a_{ijkt} - \varepsilon^q_{ijkt})$ and

[25]The positive coefficient on age in column (5) may appear surprising at first, but this coefficient cannot be directly interpreted as this estimation also includes a full set of interaction terms between age and size.

its interaction with age are extremely stable across different bins of survival probability.[26]

Table 5: Demand shocks and beliefs updating: controlling for endogenous exit

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijkt+1}$ | | | | $\Delta\varepsilon^q_{ijkt+1}$ | | |
| Selection correction | | Linear | | | | Semi-parametric | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.065$^a$ | 0.075$^a$ | 0.075$^a$ | | 0.065$^a$ | 0.075$^a$ | 0.075$^a$ | |
| | (0.001) | (0.002) | (0.001) | | (0.001) | (0.002) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | -0.003$^a$ | -0.003$^a$ | | | -0.003$^a$ | -0.003$^a$ | |
| | | (0.000) | (0.000) | | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | 0.069$^a$ | | | | 0.069$^a$ |
| | | | | (0.001) | | | | (0.001) |
| ... | | | | | | | | |
| $\times$ Age$_{ijkt} = 9$ | | | | 0.054$^a$ | | | | 0.054$^a$ |
| | | | | (0.007) | | | | (0.007) |
| $\widehat{\Pr(\text{exit}_{ijkt})}$ | -0.409$^a$ | -0.409$^a$ | -0.409$^a$ | -0.417$^a$ | | | | |
| | (0.005) | (0.005) | (0.003) | (0.005) | | | | |
| Age$_{ijkt}$ | -0.054$^a$ | -0.054$^a$ | -0.054$^a$ | | -0.057$^a$ | -0.057$^a$ | -0.057$^a$ | |
| | (0.001) | (0.001) | (0.000) | | (0.001) | (0.001) | (0.001) | |
| Observations | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 |

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3) and (7)). $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (4) and (8) but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). In columns (1)-(4), predicted exit probabilities are obtained from the estimation of Table A.11, column (4) and introduced directly in equation (19). In columns (5) to (8), they are introduced semi-parametrically in the second step, i.e. we included 100 bins corresponding to each percentile of the variable. Online appendix Table A.14 reports the full set of coefficients.

These results suggest that endogenous exit does not bias our results. We can go further and try to account for a potential selection bias by including a correction term in our estimations. Given the structure of our selection equation (which includes two high dimensional sets of fixed effects), we cannot use probit or other maximum likelihood estimators to implement a standard Heckman procedure. Instead, we follow Olsen (1980) and include a correction term constructed from a linear estimation of the selection equation. Crucially, Olsen's correction term is linear, which implies that the selection equation needs to include variables which do not appear in the second step.[27] This is not a problem in our case as $ikt$ and $jkt$ fixed effects can be used as exclusion variables. Results appear in Table 5, columns (1) to (4) and are again close to our baseline estimates. Alternatively, we can relax the linearity assumption of the correction term and use a partially linear approach in the second step. More precisely, as suggested by Cosslett (1991), we replace the linear correction term by a hundred indicator variables constructed from predicted exit

[26]In the online appendix (Table A.13), we perform a similar analysis but define quintiles based on both exit probability and firm-market size. The results are similar.

[27]See Vella (1998) for a summary of Olsen (1980) and alternative procedures to correct for endogenous sample selection. More details about the procedure appear in the online appendix, section H.

probabilities. Results are provided in Table 5, columns (5) to (8). Again our coefficients of interest are largely unaffected.[28]

### 5.3 Measurement issues

In section I of the online appendix, we perform some additional robustness checks. In particular: (i) we restrict the sample to extra EU destinations to account for the different treatment of EU trade flows by the customs (section I.1); (ii) we use alternative definitions of firm age (section I.2); (iii) we reconstruct the years, beginning the month of the first entry at the firm-product-destination level, to account for the fact that the first year of export measured over a calendar year is potentially incomplete, as pointed out by Berthou and Vicard (2015) and Bernard et al. (2017), which can affect growth rates in the first period (section I.3); (iv) we replicate the results with equation (17) being estimated at the 4-digit (HS4) instead of 6-digit level, as some 6-digit categories might include few observations, leading to imprecise estimates (section I.4); and (v) we re-estimate $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ including $ijt$ fixed effects in equations (14) and (15) to control for the potential informational spillovers from selling other products in the same destination (section I.5). Each set of results is discussed in details in the online appendix. In all cases, they are extremely close to our baseline estimates of Table 2.

Overall, the results presented in this section show that both the magnitude of belief updating and its age dependence are extremely stable across various samples and specifications, which strongly suggests that our findings are not driven by specific sectors, firms or modelling assumptions.

## 6 Discussion: alternative mechanisms on the demand side?

Several alternative demand side mechanisms have been proposed in the literature to explain firm dynamics. They mainly give rise to demand accumulation, either endogenously or exogenously. A first category of models considers firms engaging in market-specific investment to increase their profitability, or in a costly search for new buyers (see for instance Ericson and Pakes, 1995, Luttmer, 2011, Eaton et al., 2014, Fitzgerald et al., 2016). A second possibility is that firms price low in their first years to build a consumer base (Foster et al., 2016, Gourio and Rudanko, 2014). Finally, demand could simply evolve exogenously over time as in Ruhl and Willis (2017). All these mechanisms would generate the increase in average sales observed over time for surviving firms that we documented in

[28] Online appendix Table A.15 shows the results of an alternative semi-parametric procedure consisting in using a polynomial expansion of the first-step prediction as a correction term. We also try a standard Heckman procedure, estimating the first step by probit without fixed effects and relying on the nonlinearity of the inverse mills ratio in the second step to identify the selection term.

section 2.2 (Figure 2) – and this is precisely the stylized fact that motivated many of these papers. As already underlined, models of demand accumulation, if they do not include some learning about demand, cannot deliver our main prediction, i.e. that firms adjust less and less their quantities to past demand shocks as they grow older in a market. Yet, we cannot exclude *a priori* the possibility that some demand accumulation is at play on top of the updating process. Put differently, our estimates of $\varepsilon^q_{ijkt}$, which we interpret as beliefs, could in theory reflect other types of dynamics of market-specific demand. In this section we first show that our assumption of firms learning about a constant demand parameter is consistent with our data, i.e. that variations in $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ can indeed be interpreted as being driven, at least to a first order, by the updating process.[29] We then show that the variance of estimated beliefs explains a large part of the observed variance of sales growth within cohort.

## 6.1 Dynamics of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$

To further check the validity of the model, we study how the quantities and prices residuals $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ vary with age within cohorts, as the predictions of the learning model differ from those of demand accumulation theories. In the passive learning model, the dynamics of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ are affected by both within firm-markets dynamics and selection effects. Indeed, conditional on age and fixed effects, the decision to stay or exit the market depends on the firm's beliefs: there is a threshold value below which firms exit the market.[30] Exit decisions thus depend on the beliefs at the beginning of the period and on the demand shocks received. First, for a given demand shock, the smallest firms – firms with the lowest $\varepsilon^q_{ijkt}$ – are more likely to exit. Second, for a given level of beliefs, firms that decrease in size – those facing negative demand shocks – exit more. Therefore, survivors are firms that received positive demand shocks on average.

**Dynamics of $\varepsilon^q_{ijkt}$.** Both effects imply that, conditional on survival, $\varepsilon^q_{ijkt}$ should grow on average over time within cohorts. This is due to composition effects: as prior beliefs are unbiased on average, firms have equal probabilities to update upward or downward. Hence, when focusing on within firm-market variations (i.e. controlling for firm-product-destination fixed effects), quantities should become much flatter.[31] This is indeed what we find in Figure 2 (the complete set of coefficients and standard errors is provided in Tables A.23 and A.24 in the online appendix K). Figure 2.a plots the coefficients obtained when we simply regress $\varepsilon^q_{ijkt}$ on firm-market age: $\varepsilon^q_{ijkt}$ sharply increases with age. When instead we focus on variation within firms-markets (Figure 2.b), $\varepsilon^q_{ijkt}$ becomes almost flat: it only

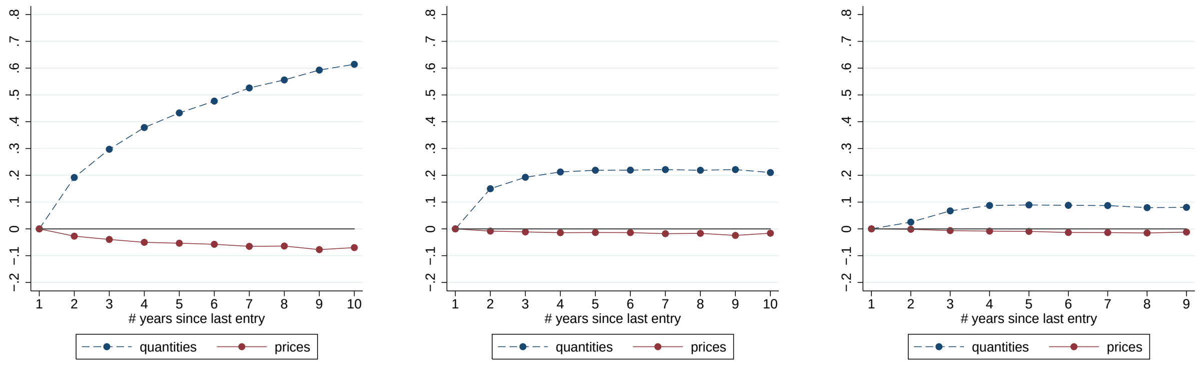
[29] In online appendix section J, we also implement a test proposed by Pakes and Ericson (1998) to discriminate between models of "active"" and "passive" learning. This test mainly shows that firms' beliefs do not follow a Markov process, a feature that should be natural in models of demand accumulation where the decision to accumulate demand depends on firm size.

[30] See online appendix section F for details.

[31] The passive learning model actually generates positive *unconditional* growth rates of quantities, i.e. even in the absence of composition effects triggered by selection. As shown in the online appendix G, this is however a weaker prediction, as it is driven by functional form assumptions.

exhibits a slight positive growth in the first years, especially at age 2. This is mostly due to the incompleteness of the first year of export measured over the calendar year; as shown in Figure 2.c, when years are reconstructed to start the month of the first entry,[32] the increase observed in the second year almost vanishes. After 3 years, $\varepsilon^{q}_{ijkt}$ is only 9 percent higher than at the time of entry, and remains constant afterwards.

Figure 2: Dynamics of $\varepsilon^{q}_{ijkt}$ and $\varepsilon^{p}_{ijkt}$




(a) All firms

(b) All firms with firm-market FE

(c) All firms with firm-market FE (reconstructed years)

Note: This figure plots the coefficients obtained when regressing the prices and quantities residuals $\varepsilon^{q}_{ijkt}$ and $\varepsilon^{p}_{ijkt}$ on a set of age dummies. Age is defined at the firm-market (firm-product-destination) level. Panel (b) controls for firm-product-destination fixed effects. Panel (c) considers the same specification as panel (b) but on the dataset of reconstructed years (see section I.3 in the online appendix). The complete set of coefficients and standard errors are shown in the online appendix Table A.23 (columns (2) and (6) for panel (a) and (4) and (8) for panel (b)) and Table A.24 (columns (3) and (6) for panel (c)).

These results contrast with the prediction of demand accumulation theories. In these models, we would expect quantities to increase more gradually and more strongly over time. Moreover, such an increase should not only be observed in the pooled regressions of Figure 2.a, but also in the within firms-markets estimations of Figures 2.b and 2.c.[33] We do find some growth at early age even after accounting for composition effects, which is consistent both with demand accumulation theories and with the passive learning model. Yet, this growth is extremely limited in magnitude and in duration, which suggests that the role of demand accumulation processes, if any, seems modest in our data at the firm-market level.

**Dynamics of $\varepsilon^{p}_{ijkt}$.** When interpreted through the lens of the learning model, $\varepsilon^{p}_{ijkt}$ represents the difference between demand shocks and the firms' expected demand. Composition effects imply that $\varepsilon^{p}_{ijkt}$ should decrease over time. Because they receive positive

[32] See section I.3 in the online appendix.

[33] Note that the pattern shown in our stylized fact #2 and in Figure A.5 of the online appendix should not be seen as evidence of some sort of demand accumulation: even in the learning model, the subsample of firms-markets surviving the entire period grow over time, as they received positive demand shocks on average.

demand shocks on average, survivors initially set their price above their optimal pricing rule, to "jump" on realized demand. They next update their beliefs, which progressively become more accurate over time. $\varepsilon^p_{ijkt}$ should thus decrease on average and converge toward its steady state value. But again, controlling for firm-market fixed effects, $\varepsilon^p_{ijkt}$ should remain constant. These predictions are confirmed in Figure 2: without firm-market fixed effects, $\varepsilon^p_{ijkt}$ is decreasing in age, although the effect is quantitatively limited (Figure 2.a). This is what the passive learning model predicts as changes in the firm beliefs are supposed to affect more $\Delta\varepsilon^q_{ijkt}$ than $\Delta\varepsilon^p_{ijkt}$.[34] Note that all coefficients statistically differ from zero at conventional levels (Table A.23). More importantly, when composition effects are accounted for, prices become flat (Figures 2.b and 2.c).

While consistent with the learning model, these findings are difficult to reconcile with theories of demand accumulation. In models where such accumulation is driven by firm pricing policy (i.e. pricing low in the first years to attract consumers), prices of young firms should be lower than those of experienced exporters: $\varepsilon^p_{ijkt}$ should increase over time in Figures 2.b and 2.c. If demand accumulation is not driven by firm pricing, prices should stay constant over time; they should not decline with age as in Figure 2.a.[35] Overall, the results shown in Figure 2 therefore support our interpretation of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ as being mostly driven by the updating process.[36]

## 6.2 The variance of firms' growth

We have seen in section 2.2 (Figure 1) that the variance of observed growth rates within market-specific cohorts of firms decline with the age of the cohort conditional on size, a fact that does not arise naturally in models where learning is absent.[37] On the other hand, with learning, younger firms update more than older firms and so have larger growth rates in absolute value. It follows that the variance of firms growth decreases with the cohort tenure on a specific market. As formally shown in the appendix, we get the following prediction which is a direct consequence of firm updating:

**Prediction # 3** (variance of growth rate): *The within cohort variances of growth rates of $Z^q_{ijkt}$ and $Z^p_{ijkt}$ decrease with cohort age.*

[34] The magnitude of the difference in growth rates should be a factor $\sigma_k$ (equations (14) and (15)), which is indeed close to what we find in Table A.23 when comparing the price and quantity equations.

[35] The price decrease we find in Figure 2.a also suggests that at least part of the updating process we uncover is directly about demand. Indeed, if firms were fully informed about the demand function (and would learn about something else, for instance productivity), they would choose a quantity - prices couple on the demand function and prices should not deviate from the optimal pricing rule.

[36] This does not imply that demand accumulation processes are not relevant to explain other dimensions of firm dynamics. For instance, firms may accumulate demand due to investment or marketing expenses that affect simultaneously their sales in many markets, or because of product-specific trends in consumer tastes: firms with the "right" product would experience positive growth in demand. Since these elements are purged from our quantities and prices residuals, we cannot infer their importance.

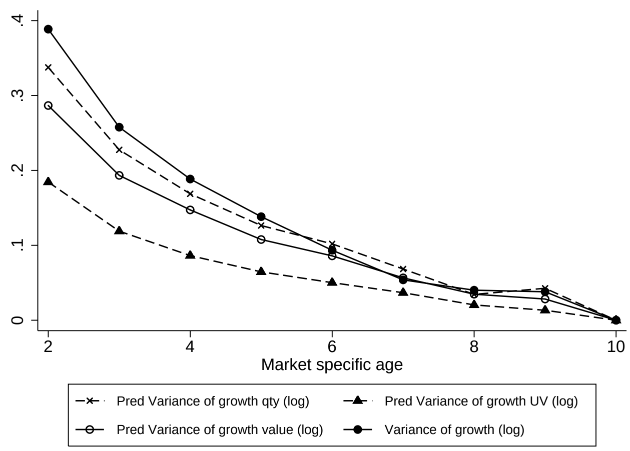
[37] The literature has however proposed mechanisms allowing to explain the decline in variance of growth rate with size, *conditional* on age (see for instance Luttmer, 2011).

We test this prediction by estimating the following equation:

$$\mathbb{V}\left(\Delta\varepsilon_{ijkt}^{X}\right) = \delta^{X} \times \text{AGE}_{cjkt} + \mathbf{FE}_{cjk} + u_{ijkt} \qquad \forall X = \{q, p\} \tag{20}$$

where $\mathbf{FE}_{cjk}$ represent cohort fixed effects. As in section 2.2, a cohort of new exporters on a product-destination market is defined as all firms entering market $jk$ in year $t$. We again expect our coefficient of interest $\delta^X$ to be negative: because firms update less their beliefs when they gain experience in a market, their quantities and prices become less volatile. Using the estimated coefficients from (20), we can also check whether the variance of the growth in beliefs match the observed variance of sales growth.

Figure 3: Impact of firm-market specific age conditional on size: predicted patterns




Note: this figure plots the variance of quantities, prices and values residuals within cohort over age on each product-destination market. The coefficients are shown in columns (2) and (6) in table A.25 and column (2) in table A.26 respectively. The coefficients on the growth and variance of growth of sales from figure 1.a.

Figure 3 shows the results. We plot the variance of the growth of quantities (beliefs) and prices residuals, as well as of the predicted value of sales and compare it with the observed variance of sales growth. The full set of estimates appear in online appendix L. Within cohort, the variance of the growth rate of both beliefs and prices residuals sharply decreases with age in all columns. Note that this is still true when controlling for the number of observations in the cohort, for average size, or for attrition by concentrating on the firms-markets which survive over the entire period (see Tables A.25 and A.26). The variance of $\varepsilon_{ijkt}^{q}$ follows quite closely the variance of observed sales. Again, given that this decline in variance with age conditional on size cannot be explained by models without learning, this provides further support for the learning model.

# 7 Conclusion

In this paper we have provided direct evidence that passive learning about demand is an important determinant of firm dynamics. We derived a core prediction from a standard model of market-specific firm dynamics incorporating Bayesian learning about local demand that theories without learning cannot generate: a demand signal leads firms to update their beliefs, especially when they are young. Combining the structure of the model with detailed exporter-level data, we developed a methodology to identify demand shocks and firms beliefs about demand.

The learning process generates the decline in the growth rates and their variance within cohort with firms' age found in the data. Our framework is also consistent with heterogeneous patterns of growth of surviving firms since over-optimistic firms upon entry may experience negative growth. We have focused on a specific dimension of firm dynamics – the post-entry firm behavior at the product-destination level –, yet this dimension explains more than half of the variance in overall firm growth.

Our results open several paths for future research. An implication of the model is that the learning process creates a form of hysteresis: the most experienced firms are less sensitive to demand shocks in terms of sales and exit decisions. This suggests that aggregate uncertainty shocks, thought as an increase in the dispersion of micro-level shocks (Bloom et al., 2014), should have heterogeneous effects across industries depending on their age structure. Another natural extension of our paper would be to go beyond post-entry dynamics and extend our framework to include explicitly informational spillovers across products, destinations or firms. Such spillovers could affect firms' entry decisions and size upon entry. Quantifying the respective contributions of each of these sources of information to firm dynamics would bear direct policy relevance.

# References


Abbring, Jaap H. and Jeffrey R. Campbell (2005), "A Firm's First Year." Tinbergen Institute Discussion Papers 05-046/3, Tinbergen Institute.

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

Arkolakis, Costas (2016), "A unified theory of firm selection and growth." *Quarterly Journal of Economics*, 131(1), 89–155.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott (2009), "The Margins of US Trade." *American Economic Review*, 99, 487–93.

Bernard, Andrew B., Renzo Massari, Jose-Daniel Reyes, and Daria Taglioni (2017), "Exporter Dynamics and Partial Year Effects." *American Economic Review*, 107, 3211–3228.

Berthou, Antoine and Vincent Vicard (2015), "Firms' export dynamics: experience vs. size." *World Economy*, 38, 1130–1158.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry (2014), "Really Uncertain Business Cycles." WP 14-18, Center for Economic Studies.

Buono, Ines and Harald Fadinger (2012), "The micro-dynamics of exporting: evidence form French firms." Working Paper 2012-880, Banca d'Italia.

Cabral, Luis and Jose Mata (2003), "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review*, 93, 1075–1090.

Chamberlain, Gary (1986), "Asymptotic Efficiency in Semiparametric Models with Censoring." *Journal of Econometrics*, 32, 182–218.

Clementi, Gina Luca and Hugo A Hopenhayn (2006), "A Theory of Financing Constraints and Firm Dynamics." *The Quarterly Journal of Economics*, 121, 229–265.

Cosslett, Stephen (1991), "Semiparametric estimation of a regression model with sample selectivity." In *Nonparametric and Semiparametric Methods in Econometrics and Statistics* (Barnett. James Powell William A and George Tauchen, eds.), Cambridge University Press.

Defever, Fabrice, Benedikt Heid, and Mario Larch (2015), "Spatial Exporters." *Journal of International Economics*, 95, 145–156.

Dunne, Timothy, Mark J Roberts, and Larry Samuelson (1989), "The Growth and Failure of U.S. Manufacturing Plants." *The Quarterly Journal of Economics*, 104, 671–98.

Eaton, Jonathan, Marcela Eslava, David Jinkins, C. J. Krizan, and James Tybout (2014), "A search and learning model of export dynamics." Unpublished manuscript.

Eaton, Jonathan, Marcela Eslava, Maurice Kugler, and James Tybout (2008), "Export dynamics in colombia: Firm-level evidence." In *The Organization of Firms in a Global Econom* (D. Marin E. Helpman and T. Verdier, eds.), Cambridge, MA: Harvard University Press.

Eaton, Jonathan, Samuel Kortum, and Francis Kramarz (2011), "An anatomy of international trade: Evidence from french firms." *Econometrica*, 79, 1453–1498.

Ericson, Richard and Ariel Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, 62, 53–82.

Evans, David (1987), "Test of alternative theories of firm growth." *Journal of Policial Economy*, 95, 657–74.

Fernandes, Ana and Heiwai Tang (2014), "Learning from Neighbors' Export Activities: Evidence from Exporters' Survival." *Journal of International Economics*, 94, 87–94.

Fitzgerald, Doireann, Stephanie Haller, and Yaniv Yedid-Levi (2016), "How Exporters Grow." Staff Report 524, Federal Reserve Bank of Minneapolis.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98, 394–425.

Foster, Lucia, John Haltiwanger, and Chad Syverson (2016), "The slow growth of new plants: learning about demand?" *Economica*, 83, 91–129.

Gourio, Francois and Leena Rudanko (2014), "Customer capital." *Review of Economic Studies*, 81, 1102–1136.

Guimaraes, Paulo and Pedro Portugal (2010), "A simple feasible alternative procedure to estimate models with high-dimensional fixed effects." *Stata Journal*, 10, 628–649.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda (2013), "Who Creates Jobs? Small versus Large versus Young." *The Review of Economics and Statistics*, 95, 347–361.

Hopenhayn, Hugo A (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica*, 60, 1127–50.

Imbs, Jean and Isabelle Mejean (2015), "Elasticity optimism." *American Economic Journal: Macroeconomics*, 7, 43–83.

Impullitti, Giammario, Alfonso A. Irarrazabal, and Luca David Opromolla (2013), "A theory of entry into and exit from export markets." *Journal of International Economics*, 90, 75–90.

Jovanovic, Boyan (1982), "Selection and the Evolution of Industry." *Econometrica*, 50, 649–70.

Klette, Tor Jakob and Samuel Kortum (2004), "Innovating Firms and Aggregate Innovation." *Journal of Political Economy*, 112, 986–1018.

Lawless, Martina (2009), "Firm export dynamics and the geography of trade." *Journal of International Economics*, 77, 245–254.

Luttmer, Erzo G. J. (2007), "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics*, 122, 1103–1144.

Luttmer, Erzo G. J. (2011), "On the Mechanics of Firm Growth." *Review of Economic Studies*, 78, 1042–1068.

Morales, Eduardo, Gloria Sheu, and Andres Zahler (2014), “Extended Gravity.” unpublished manuscript, Princeton University.

Mulligan, Casey B. and Yona Rubinstein (2008), “Selection, Investment, and Women’s Relative Wages Over Time.” *The Quarterly Journal of Economics*, 123, 1061–1110.

Munch, Jakob R and Daniel Nguyen (2014), “Decomposing Firm-Level Sales Variation.” *Journal of Economic Behavior and Organization*, 106, 317–334.

Olsen, Randall J (1980), “A least squares correction for selectivity bias.” *Econometrica*, 48, 1815–1820.

Pakes, Ariel and Richard Ericson (1998), “Empirical Implications of Alternative Models of Firm Dynamics.” *Journal of Economic Theory*, 79, 1–45.

Piveteau, Paul (2016), “An empirical model of export entry with customer accumulation.” Unpublished manuscript.

Rauch, James E. (1999), “Networks versus markets in international trade.” *Journal of International Economics*, 48, 7–35.

Rossi-Hansberg, Esteban and Mark L. J. Wright (2007), “Establishment Size Dynamics in the Aggregate Economy.” *American Economic Review*, 97, 1639–1666.

Ruhl, Kim J. and Jonathan L. Willis (2017), “New Exporter Dynamics.” *International Economic Review*, 58, 703–726. Forthcoming.

Timoshenko, Olga A. (2015), “Product Switching in a Model of Learning.” *Journal of International Economics*, 95, 233–249.

Vella, Francis (1998), “Estimating Models with Sample Selection Bias: A Survey.” *Journal of Human Resources*, 33, 127–169.

# A Appendix - Detailed derivations and proofs

**Optimal quantities and prices.** Firms choose quantities by maximizing expected profits subject to demand. Using (1), we get:

$$\max_q \int \pi_{ijkt} dG_{t-1}(a_{ijkt}) = \max_q q_{ijkt}^{1-\frac{1}{\sigma_k}} \left(\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}\right)^{\frac{1}{\sigma_k}} \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] - \frac{w_{it}}{\varphi_{ikt}} q_{ijkt} - F_{ijk}$$

The FOC writes:

$$\left(1-\frac{1}{\sigma_k}\right) q_{ijkt}^{-\frac{1}{\sigma_k}} \left(\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}\right)^{\frac{1}{\sigma_k}} \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] = \frac{w_{it}}{\varphi_{ikt}} \Leftrightarrow q_{ijkt}^* = \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k} \left(\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}\right) \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]^{\sigma_k}$$

And from the constraint, we get $p^*_{ijkt} = \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right)\left(\frac{e^{\frac{a_{ijkt}}{\sigma_k}}}{E_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]}\right)$

**Updating of firm's beliefs about expected demand.** First note that firm $i$ has a prior about the demand shock given by $a_{ijkt} \sim \mathcal{N}(\widetilde{\theta}_{ijkt-1}, \widetilde{\sigma}^2_{ijkt-1} + \sigma^2_\varepsilon)$ and thus $e^{\frac{a_{ijkt}}{\sigma_k}} \sim LN(\frac{\widetilde{\theta}_{ijkt-1}}{\sigma_k}, \frac{\widetilde{\sigma}^2_{ijkt-1} + \sigma^2_\varepsilon}{\sigma^2_k})$. It follows that $\mathbb{E}_{t-1}[e^{\frac{a_{ijkt}}{\sigma_k}}] = \int \left(e^{\frac{a_{ijkt}}{\sigma_k}}\right) dG_{t-1}(a_{ijkt}) = e^{\frac{1}{\sigma_k}\left(\widetilde{\theta}_{ijkt-1} + \frac{\widetilde{\sigma}^2_{ijkt-1} + \sigma^2_\varepsilon}{2\sigma_k}\right)}$. Hence:

$$\Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] = \frac{1}{\sigma_k}\left(\Delta\widetilde{\theta}_{ijkt} + \frac{\widetilde{\sigma}^2_{ijkt} - \widetilde{\sigma}^2_{ijkt-1}}{2\sigma_k}\right)$$

Using the the definition of $\Delta\widetilde{\theta}_{ijkt}$, $g_t$, $\widetilde{\sigma}^2_{ijkt-1}$ and $\widetilde{\sigma}^2_{ijkt}$ (see (3) and (4)), it is easy to show that $\frac{\widetilde{\sigma}^2_{ijkt-1} - \widetilde{\sigma}^2_{ijkt}}{g_t} = \widetilde{\sigma}^2_{ijkt-1}$. It follows the expression in the text (11):

$$\Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] = \frac{g_t}{\sigma_k}\left(a_{ijkt} - \widetilde{\theta}_{ijkt-1}\right) - \frac{g_t}{\sigma_k}\frac{\widetilde{\sigma}^2_{ijkt-1}}{2\sigma_k}$$

But we only observe $\Delta\varepsilon^q_{ijkt+1} = \sigma_k \Delta \ln \mathbb{E}_t[e^{\frac{a_{ijkt+1}}{\sigma_k}}]$. It follows that $\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \widetilde{\theta}_{t-1}\right) - g_t\frac{\widetilde{\sigma}^2_{ijkt-1}}{2\sigma_k}$. As $\varepsilon^q_{ijkt} = \sigma_k \ln \mathbb{E}_{t-1}[e^{\frac{a_{ijkt}}{\sigma_k}}]$, we get $\varepsilon^q_{ijkt} = \widetilde{\theta}_{ijkt-1} + \frac{\widetilde{\sigma}^2_{ijkt-1} + \sigma^2_\varepsilon}{2\sigma_k}$, or $\widetilde{\theta}_{ijkt-1} = \varepsilon^q_{ijkt} - \frac{\widetilde{\sigma}^2_{t-1} + \sigma^2_\varepsilon}{2\sigma_k}$. This leads to the equation we test in the main text:

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t\frac{\sigma^2_\varepsilon}{2\sigma_k} \tag{21}$$

**Prediction 1.** Prediction 1 states that $a_{ijkt} - \varepsilon^q_{ijkt}$, has a larger impact on firms' updating, the younger the firms are. Using (21), we immediately get:

$$\frac{\partial \Delta\varepsilon^q_{ijkt+1}}{\partial\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)} = g_t > 0$$

Updating is larger for younger firms, as $g_t$ decreases with $t$.

**Prediction 2: Impact of market uncertainty.** Moreover, the updating process is also affected by the level of market uncertainty $\sigma^2_\epsilon$. Formally:

$$\frac{\partial^2\left(\Delta\varepsilon^q_{ijkt+1}\right)}{\partial\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)\partial\sigma^2_\epsilon} = -\frac{g^2_t}{\sigma^2_{jk0}} < 0$$

Updating decreases with uncertainty, as a signal is less informative when market uncertainty is larger. As a consequence, market uncertainty dampens the speed of learning. In other words, updating decreases less with age, the more uncertain the market. This can be seen noting that:

$$\frac{\partial^2\left(\Delta\varepsilon^q_{ijkt+1}\right)}{\partial\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)\partial t} = -\frac{1}{\left(\frac{\sigma^2_\epsilon}{\sigma^2_{jk0}} + t\right)^2}$$

which is larger (less negative) in more uncertain markets (with larger $\sigma^2_\epsilon$).

**Dynamics of prices and quantities**. The model predicts expected growth rates of opposite signs for quantities and prices. This result comes from (14) and (15). Taking the first difference of these equations in expected terms, we directly get the expected growth rates. We find:

$$\mathbb{E}\left[\Delta \ln Z^q_{ijkt+1}\right] = -\frac{1}{\sigma_k}\mathbb{E}\left[\Delta \ln Z^p_{ijkt+1}\right]$$

Given that firms that decrease in size will on average be more likely to exit, the expected growth rate of quantities must be positive for survivors. Hence, the expected growth rate of prices for these firms should be negative and smaller by a factor $-\frac{1}{\sigma_k}$. Quantitatively, this is very close to what we find in table A.23.

**Prediction 3.** Prediction 3 states that the variance of growth rates within cohort decrease with cohort age. The variance of these growth rates can be expressed as:

$$\mathbb{V}\left[\Delta \ln Z^q_{ijkt+1}\right] = \sigma_k^2 \mathbb{V}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) \tag{22}$$

$$\begin{aligned}\mathbb{V}\left[\Delta \ln Z^p_{ijkt+1}\right] &= \left(\frac{1}{\sigma_k}\right)^2 \mathbb{V}\left(\Delta a_{ijkt+1}\right) + \mathbb{V}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) \\ &\quad -\frac{2}{\sigma_k} Cov\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right)\end{aligned} \tag{23}$$

First, $a_{ijkt+1}$ and $a_{ijkt}$ being drawn from the same distribution, $\mathbb{V}\left[\Delta a_{ijkt+1}\right] = 2\sigma_\epsilon^2$. Second, using (11), we get:

$$\mathbb{V}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) = \left(\frac{\sigma_\epsilon}{\sigma_k\left(\frac{\sigma_\epsilon^2}{\sigma_{jk0}^2} + t\right)}\right)^2$$

As $\mathbb{E}\left[\Delta a_{ijkt+1}\right] = 0$, we get:

$$Cov\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right) = \mathbb{E}\left[\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\Delta a_{ijkt+1}\right]$$

Expanding this expression and using the fact that $a_{ijkt}$ and $a_{ijkt+1}$ are independent and that $\mathbb{E}\left[a_{ijkt}\right] = \mathbb{E}\left[a_{ijkt+1}\right] = \overline{a}_{ijkt-1}$, we get:

$$Cov\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right) = -\frac{\sigma_\epsilon^2}{\sigma_k\left(\frac{\sigma_\epsilon^2}{\sigma_{jk0}^2} + t\right)}$$

Finally, plugging this term into (22) and (23) and after rearranging, we get the following expressions which are both strictly decreasing with $t$:

$$\mathbb{V}\left[\Delta \ln Z^q_{ijkt+1}\right] = \left(\frac{\sigma_\epsilon}{\left(\frac{\sigma_\epsilon^2}{\sigma_{jk0}^2} + t\right)}\right)^2 \tag{24}$$

$$\mathbb{V}\left[\Delta \ln Z^{p}_{ijkt+1}\right] = \left(\frac{\sigma_{\epsilon}}{\sigma_{k}}\right)^{2} \left(\left(\frac{1}{\left(\frac{\sigma^{2}_{\epsilon}}{\sigma^{2}_{jk0}} + t\right)} + 1\right)^{2} + 1\right) \tag{25}$$

# Demand learning and firm dynamics: evidence from exporters

Online Appendix

December 5, 2017

Nicolas Berman[1] Vincent Rebeyrol[2] Vincent Vicard[3]

[1] Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS, Graduate Institute Geneva and CEPR. E-mail: nicolas.berman@univ-amu.fr.

[2] Toulouse School of Economics. E-mail: vincent.rebeyrol@ut-capitole.fr

[3] Banque de France. E-mail: vincent.vicard@banque-france.fr.

# Contents

# A Data and descriptive evidence on firm dynamics

## A.1 Dataset construction

We use data on values and quantities sold by French firms, by destination, HS6 product and year, over the period 1994-2005. We focus on the subset of HS6 product categories that remain stable in the HS classification over the period in order to be able to track firms over time on specific markets.[4] As we use the first two years to define entry, we concentrate on the years 1996-2005. Note that firms-products-destinations that already export at the beginning of the period (in 1994 or 1995) are not considered, as we are interested in post-entry dynamics.

Because intra-EU and extra-EU flows are treated differently by the French Customs, we harmonize the data in several ways. The declaration of extra-EU export flows is mandatory when a transaction exceeds 1,000 euros or 1,000 kg. For shipments to EU countries, firms have to report their detailed expeditions when their total exports to all EU countries exceed a threshold over the year of 38,100 euros before 2001, 99,100 euros in 2001 and 100,000 euros between 2002 and 2005. Firms below the reporting threshold are required to fill a simplified form without the details on the product exported and the destination market. In order to harmonize the data requirement over the different destinations, we drop all intra-EU export flows below 1,000 euros, as well as firms that report at least once under the simplified procedure (as for these firms, we do not observe their flows in all markets). We also check that all our results are unchanged when removing EU destinations from the sample.

## A.2 Additional descriptive evidence

This section provides further details on the computation of the stylized facts presented in section 2 of the main text.

**Contribution to aggregate sales growth.** The literature has documented the essential contribution of young firms to industry dynamics, either in terms of aggregate output, employment or trade. Haltiwanger et al. (2013) show for instance that US start-ups display substantially higher rates of job creation and destruction in their first ten years, and that these firms represent a large share of total employment after a decade of existence. These patterns are also found for other countries (see Criscuolo *et al.*, 2014 for evidence on 18 OECD countries; Lawless (2014) on Irish firms, Ayyagari *et al.*, 2011 for developing countries). Similar facts characterize trade dynamics: Eaton et al. (2008) and Bernard et al. (2009) show that exporters start small but that, conditional on survival, they account for large shares of total export growth after a few years.

[4] The frequent changes in the combined nomenclature (CN8) prevents us to use this further degree of disaggregation of the customs' product classification.

Our exporter-level data exhibit comparable features. We compute the contribution of the intensive margin (incumbent firm×product×destination) and the firm- and firm-market extensive margins to the growth of total French exports on a year-on-year basis or over the entire time frame of the sample (between 1996 and 2005). We use mid-point growth rates to account for entries and exits Bricongne *et al.* (2012). Initial size relates to firms' sales the first year of entry on a specific market on which they export up to 2005.

Over the 1996-2005 period, we find that, on average, new firm-destination-product triplets represent only 12.3% of total export value after a year, but their share reaches 53.5% after a decade (27.3% due to new markets served by incumbents and 26.2% by new firms exporting, see Table A.1). The contribution of the extensive margin to aggregate exports is determined by three components of firm dynamics: entry, survival and post entry growth on new markets. Since new exporters typically do not survive more than a few years in export markets,[5] firm selection and growth are important drivers of aggregate trade growth over longer horizons, besides the size at entry. Column (2) of Table A.1 shows that pure growth after entry accounts for around 40% of the end-of-period share of newly created firm-destination-product triplets. The objective of our paper is precisely to understand how learning about demand can explain this post-entry dynamics.

Table A.1: Shares in end-of-period French aggregate exports

| | (1)<br>Average<br>yoy 1996/2005 | (2)<br>Overall<br>1996/2005 |
|---|---|---|
| New exporters | 2.4% | 25.9% |
| *Initial size* | - | *16.4%* |
| *Growth since entry* | - | *9.6%* |
| New product-destination | 9.9% | 27.7% |
| *Initial size* | - | *16.7%* |
| *Growth since entry* | - | *11.0%* |
| Incumbent exporter-product-destination | 87.7% | 46.4% |
| Total | 100% | 100% |

Note: Source: French Customs. Column (1) presents the average contribution to year-on-year growth rates, i.e. the contribution of each subcomponent to the yearly growth rates, observed for each year of our sample, then averaged across years. Column (2) reports the contributions of each subcomponent to the total growth of French exports between 1996 and 2005. Initial entry measures firms' sales the year of first entry on a specific market on which they still export in 2005; and growth since entry measures the contribution of sales growth between the first entry and 2005.

**Firm-product-destination specific factors are a key component of sales' growth.** We decompose the variance of sales growth, in a way similar in spirit to Eaton et al.

[5] For French exporters, the average survival rate at the firm-product-destination level is 32% between the first and second year, and 9% over a five-year horizon.

(2011).[6] We first regress firm-market specific sales growth on a set of destination-product-time dummies. The $R^2$ of such a regression is 0.14: market-specific dynamics play a limited role. Adding firm-product-time fixed effects increases the $R^2$ to 0.46, suggesting that supply-side factors such as productivity do a good job at explaining variations of firms' sales over time. However, it appears clearly that sales growth remains largely driven by firm-market specific factors. Our paper concentrates on this part of firm dynamics, with the objective of understanding the extent to which it is consistent with firms learning about their demand.

Table A.2: Decomposition of the variance of sales

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent var. | Growth of exports | | Value of exports | |
| Product-destination-time FE | Yes | Yes | - | - |
| Firm-product-time FE | - | Yes | - | - |
| Product-destination FE | - | - | Yes | - |
| Firm-product FE | - | - | Yes | - |
| Firm-product-destination FE | - | - | - | Yes |
| $R^2$ | 0.14 | 0.46 | 0.57 | 0.80 |

Note: OLS estimations based on French customs data. Each column contains the $R^2$ of a separate regression of the dependent variable on a specific set of fixed effects.

**Firm-market growth and its variance decline with age, conditional on size.** Columns (1), (4) and (5) of Table A.3 below shows the coefficients used to plot Figure 1 of the main text. Column (2) shows that similar results for firm growth are obtained when including in the estimations firm×product×year fixed effects. Specifications reported in columns (1)-(3) include dummies by decile of firms size computed by HS4-product×destination, and HS2 sector and year fixed effects. Firms size is defined as average firm×product×destination sales over $t$ and $t-1$ (Haltiwanger et al., 2013). Standard errors are clustered at the firm level. The exit probability in column (3) is estimated using a linear probability model. Column (4) includes year fixed effects and controls for average size, computed as the mean of average sales over $t$ and $t-1$ across firms by cohort. Standard errors are clustered at the market level.

**Post-entry growth dynamics are heterogenous across survivors.** Finally, Figures

[6] Eaton et al. (2011) show, using firm-destination data, that firm-specific effects explain well the probability of serving a market (57%), but less so sales variations conditional on selling in a market (39%). Munch and Nguyen (2014) find that the mean contribution of the firm component to unconditional sales variations is 49%. They also show that the firm-specific effects are more important for firms already established in a product-destination market. Lawless and Whelan (2014) find an adjusted pseudo-$R^2$ of 45% on a sample of Irish exporters.

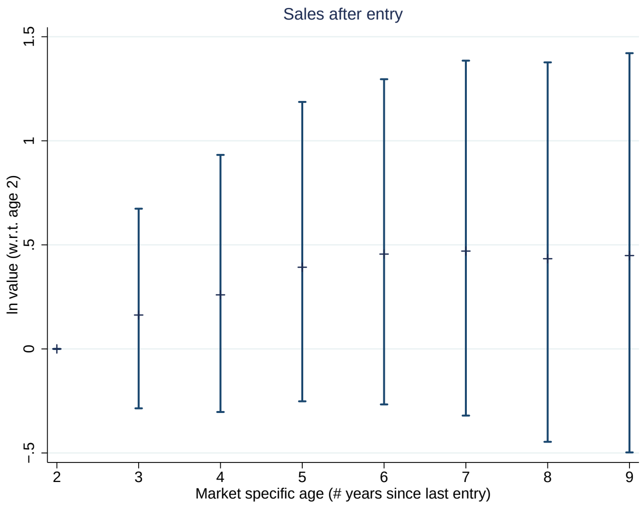
Table A.3: Age, growth and volatility of sales and exit rates

| Dep. var. | (1) Growth value (log) | (2) | (3) Exit probability | (4) Variance of growth value (log) |
|---|---|---|---|---|
| $\text{Age}_{ijkt} = 1$ | | | 0.296[a] | |
| | | | (0.009) | |
| $\text{Age}_{ijkt} = 2$ | 0.437[a] | 0.398[a] | 0.169[a] | 0.389[a] |
| | (0.015) | (0.015) | (0.009) | (0.015) |
| $\text{Age}_{ijkt} = 3$ | 0.132[a] | 0.174[a] | 0.093[a] | 0.258[a] |
| | (0.014) | (0.015) | (0.008) | (0.015) |
| $\text{Age}_{ijkt} = 4$ | 0.079[a] | 0.105[a] | 0.052[a] | 0.189[a] |
| | (0.014) | (0.015) | (0.008) | (0.015) |
| $\text{Age}_{ijkt} = 5$ | 0.055[a] | 0.069[a] | 0.025[a] | 0.138[a] |
| | (0.014) | (0.015) | (0.008) | (0.015) |
| $\text{Age}_{ijkt} = 6$ | 0.047[a] | 0.049[a] | 0.008 | 0.093[a] |
| | (0.014) | (0.015) | (0.008) | (0.015) |
| $\text{Age}_{ijkt} = 7$ | 0.032[b] | 0.031[b] | -0.002 | 0.054[a] |
| | (0.014) | (0.015) | (0.008) | (0.015) |
| $\text{Age}_{ijkt} = 8$ | 0.033[b] | 0.031[b] | -0.007 | 0.040[b] |
| | (0.014) | (0.015) | (0.007) | (0.016) |
| $\text{Age}_{ijkt} = 9$ | 0.018 | 0.011 | | 0.038[b] |
| | (0.016) | (0.017) | | (0.016) |
| $\text{Size}_{ijkt/t-1}$ - decile 1 | -0.251[a] | -0.156[a] | 0.326[a] | |
| | (0.012) | (0.021) | (0.004) | |
| $\text{Size}_{ijkt/t-1}$ - decile 2 | -0.219[a] | -0.142[a] | 0.290[a] | |
| | (0.006) | (0.008) | (0.004) | |
| $\text{Size}_{ijkt/t-1}$ - decile 3 | -0.210[a] | -0.193[a] | 0.258[a] | |
| | (0.005) | (0.006) | (0.004) | |
| $\text{Size}_{ijkt/t-1}$ - decile 4 | -0.189[a] | -0.175[a] | 0.228[a] | |
| | (0.005) | (0.005) | (0.003) | |
| $\text{Size}_{ijkt/t-1}$ - decile 5 | -0.169[a] | -0.156[a] | 0.197[a] | |
| | (0.005) | (0.005) | (0.003) | |
| $\text{Size}_{ijkt/t-1}$ - decile 6 | -0.143[a] | -0.130[a] | 0.163[a] | |
| | (0.005) | (0.004) | (0.003) | |
| $\text{Size}_{ijkt/t-1}$ - decile 7 | -0.120[a] | -0.105[a] | 0.133[a] | |
| | (0.004) | (0.004) | (0.003) | |
| $\text{Size}_{ijkt/t-1}$ - decile 8 | -0.090[a] | -0.077[a] | 0.097[a] | |
| | (0.004) | (0.004) | (0.002) | |
| $\text{Size}_{ijkt/t-1}$ - decile 9 | -0.051[a] | -0.039[a] | 0.055[a] | |
| | (0.004) | (0.004) | (0.002) | |
| Average $\text{Size}_{cjkt/t-1}$ | | | | 0.022[a] |
| | | | | (0.001) |
| Observations | 1,666,317 | 1,456,113 | 3,061,865 | 348,536 |
| Year FE | Yes | Yes | Yes | Yes |
| Sector (HS2) FE | Yes | Yes | Yes | - |
| Firm-product-year FE | - | Yes | - | - |

Robust standard errors clustered by firm (respectively destination-product in columns (4)) in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Size computed as average size in $t$ and $t+1$; size bins by decile are computed at the destination-product(HS4) level. The omitted age category is 10 years.

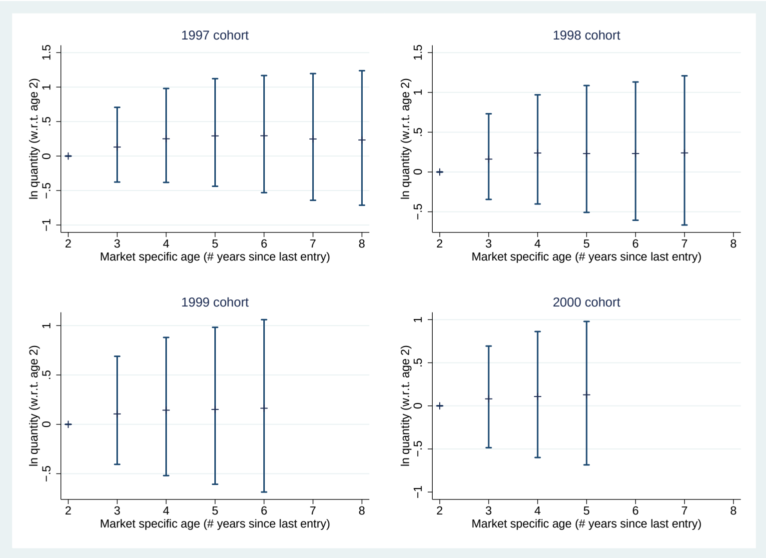
A.1 and A.2 show that the heterogenous growth dynamics of quantities that we discuss in the main text also hold for the value of sales, and for different cohorts of firm-markets.

Figure A.1: Sales dynamics over time for surviving firms




Note: This figure plots statistics about market-specific firm sales values with respect to age. Values are normalized to 1 in age 2. The upper and lower limits of the boxes represent the first and last quartiles of the variable, with the median in between.

Figure A.2: Quantity dynamics over time for surviving firms: robustness




Note: This figure plots statistics about market-specific firm quantities with respect to age. Quantities are normalized to 1 in age 2. The upper and lower limits of the boxes represent the first and last quartiles of the variable, with the median in between.

# B Elasticity of substitution estimates

Table A.4 reports the descriptive statistics on the elasticities of substitution estimated from equation 18, for the full estimation sample (upper panel) and across HS6 products (lower panel). We first report the overall estimates, then the statistics obtained when products with insignificant $\beta$ coefficients (at the 5% level) are removed from the sample, and finally the statistics obtained when products with $\sigma_k < 1$ are excluded. As can be seen from the upper panel, insignificant coefficients only represent 1.6% of the observations in the final sample, while dropping theory inconsistent elasticities (lower than 1) further eliminates only 0.1% of the observations. This clearly show that, whenever we can precisely estimate these elasticities, we get plausible coefficients.

Across products, our estimates yield a mean (resp. median) $\sigma_k$ of 7.17 (resp. 5.51) after dropping insignificant or theory-inconsistent ones (lower panel). The median is largely unaffected by our cleaning rules. Except in the case in which insignificant estimates are kept, the distribution of $\sigma_k$ does not contain extreme value: the 99% is equal to 28.7. The last three rows report $\sigma_k$ for different categories of products according to the Rauch (1999)'s liberal classification. As expected, differentiated goods exhibit a mean (resp. median) of 6.2 (resp. 5.2), lower than referenced priced goods (9.1 and 7.2 respectively) and homogenous goods (11.1 and 9.0 resp.). Those means are statistically different at the 1% level, with t-stat of -12.4 (differentiated vs. referenced), -13.3 (differentiated vs. homogenous), and -3.1 (referenced vs. homogenous).

Table A.4: Statistics on elasticities of substitution

| | Obs. | Mean | S.D. | 1% | 25% | Median | 75% | 99% |
|---|---|---|---|---|---|---|---|---|
| | | | | Full sample | | | | |
| $\sigma_k$, all estimates | 1883748 | 7.30 | 60.42 | 2.21 | 3.58 | 5.11 | 6.70 | 33.06 |
| $\sigma_k$, if $\beta$ significant | 1854359 | 6.20 | 4.79 | 2.24 | 3.57 | 5.09 | 6.59 | 26.16 |
| $\sigma_k$, if $\beta$ significant and $\sigma_k > 1$ | 1854141 | 6.20 | 4.78 | 2.24 | 3.57 | 5.09 | 6.59 | 26.16 |
| | | | | Across HS6 products | | | | |
| $\sigma_k$, all estimates | 3542 | 13.91 | 221.62 | -69.55 | 3.82 | 5.83 | 10.27 | 116.10 |
| $\sigma_k$, if $\beta$ significant | 2780 | 7.10 | 5.55 | 1.68 | 3.86 | 5.49 | 8.41 | 28.73 |
| $\sigma_k$, if $\beta$ significant and $\sigma_k > 1$ | 2767 | 7.17 | 5.47 | 1.89 | 3.88 | 5.51 | 8.42 | 28.73 |
| $\sigma_k$, if $> 1$, differentiated goods | 1778 | 6.24 | 4.21 | 2.00 | 3.77 | 5.17 | 7.20 | 21.82 |
| $\sigma_k$, if $> 1$, referenced priced goods | 670 | 9.09 | 6.78 | 1.80 | 4.46 | 7.24 | 11.61 | 32.98 |
| $\sigma_k$, if $> 1$, homogenous goods | 159 | 11.14 | 8.71 | 1.63 | 5.23 | 8.97 | 15.07 | 50.82 |

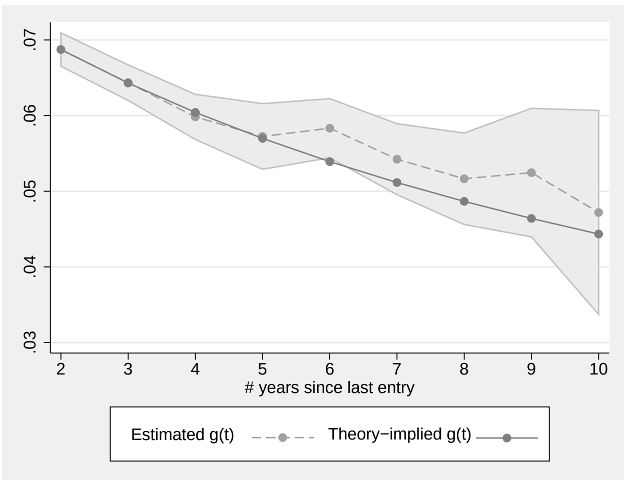
Source: Authors computations from French Customs data. Elasticities of substitution estimated from equation (17). $\sigma_k$, if significant $\beta$ means that we keep only the estimate when $\beta$ estimated from equation (17) is statistically different from zero at the 5% level. $\sigma_k$, if $> 1$ means that we further drop the observation if $\sigma_k > 1$. The classification of goods into differentiated, referenced priced and homogeneous comes from Rauch (1999).

# C Main results – Graphical representation

The figures below depicts the coefficients of 2, column (4). Equation (18) predicts that these coefficient should follow the following shape: $g_t = \frac{1}{\sigma_\varepsilon^2/\sigma_{jk0}^2+t}$ with $g_t$ measuring the speed of learning. To determine whether our set of estimated coefficients significantly differ from this shape, in Figure we have taken year 2 (the first coefficient) as a benchmark; from this coefficient we can infer the value of $\sigma_\varepsilon^2/\sigma_{jk0}^2 \approx 12$. Assuming the coefficient of year 2 is indeed correct, the shape of our coefficients is quite similar to the one implied by our functional form assumption.

We can go further and test this restriction: does our model perform significantly better than a model in which we would constrain the coefficients (from year 3 onwards) to follow the shape of $g_t$? When we test these restrictions, they are rejected at the 5% level: the shape implied by our coefficient is different from the one implied by the normality assumption (the p-value associated with the hypothesis that the models are the same is 0.02). However, as is apparent in Figure R3.1 below, this is mostly due to a difference in the coefficients after age 5. In fact, when we consider all coefficients but the one of year 6, the restrictions are no longer rejected (the p-value of 0.17). When we concentrate on the first four years (for which we have more observations and therefore more precise estimates), the p-value of the F-test of the restrictions is as high as 0.66.

Figure A.3: Functional form assumption

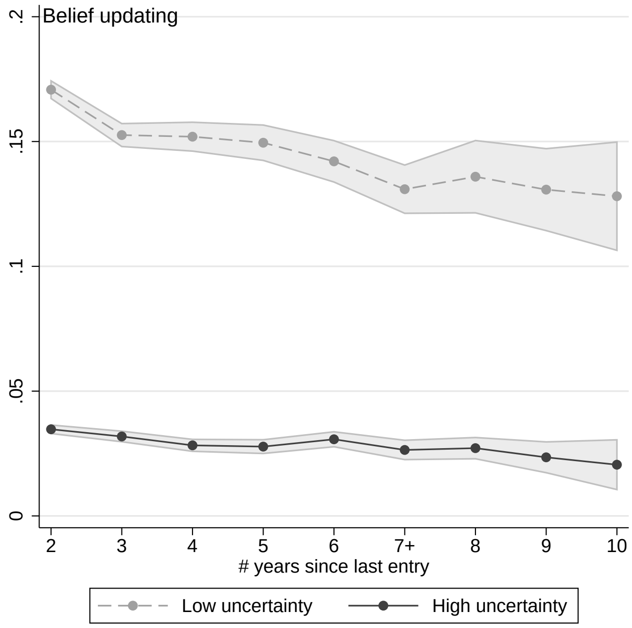



Note: This figure plot the coefficient and 95% confidence intervals from Table 2, column (4), and the coefficient implied by our functional form assumption $g_t = 1/(\sigma_\varepsilon^2/\sigma_{jk0}^2 + t)$. At age 2 we assume that the coefficient is correctly estimated and infer the rest of the theory-based coefficient from the expression of $g_t$.

# D Belief updating and uncertainty

To illustrate the impact of uncertainty on the updating process and how it evolves with age, we split our sample into markets with low (below the first quartile of uncertainty) and high (above the third quartile) uncertainty in Table A.5 and Figure A.4. We still find evidence for the updating process on both sub-samples but the average level of belief updating following a demand shock is larger on less uncertain markets (0.171 versus 0.035 for firms of age two). As expected from prediction 2, the profile of learning is also flatter on more uncertain markets.

Figure A.4: Uncertainty and belief updating




The figure plots the coefficients of regressions similar to Table 2, column (4), ran on two different sub-samples defined according to the market level of uncertainty (below the first quartile and above the third quartile). Uncertainty is computed as the standard deviation of the demand shocks $a_{ijkt}$, computed by market. Grey areas represent 90% confidence bands.

Table A.5: Prediction 1: the role of uncertainty (subsamples)

| | (1) | (2) |
|---|---|---|
| Dep. var. | $\Delta\varepsilon^q_{ijk,t+1}$ | |
| Uncertainty | High | Low |
| $(a_{ijkt} - \varepsilon^q_{ijkt}) \times \text{Age}_{ijkt} = 2$ | $0.035^a$ | $0.171^a$ |
| | (0.001) | (0.002) |
| $\times\ \text{Age}_{ijkt} = 3$ | $0.032^a$ | $0.153^a$ |
| | (0.001) | (0.003) |
| $\times\ \text{Age}_{ijkt} = 4$ | $0.028^a$ | $0.152^a$ |
| | (0.001) | (0.004) |
| $\times\ \text{Age}_{ijkt} = 5$ | $0.028^a$ | $0.150^a$ |
| | (0.002) | (0.004) |
| $\times\ \text{Age}_{ijkt} = 6$ | $0.031^a$ | $0.142^a$ |
| | (0.002) | (0.005) |
| $\times\ \text{Age}_{ijkt} = 7$ | $0.026^a$ | $0.131^a$ |
| | (0.002) | (0.006) |
| $\times\ \text{Age}_{ijkt} = 8$ | $0.027^a$ | $0.136^a$ |
| | (0.003) | (0.009) |
| $\times\ \text{Age}_{ijkt} = 9$ | $0.023^a$ | $0.131^a$ |
| | (0.004) | (0.010) |
| $\times\ \text{Age}_{ijkt} = 10$ | $0.021^a$ | $0.128^a$ |
| | (0.006) | (0.013) |
| Observations | 454040 | 438324 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). $\text{Age}_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Uncertainty is the standard deviation of $a_{ijkt}$, computed by market. High and low mean above the third quartile and below the first quartile of the uncertainty variable.

# E Extensions of the model

We consider in this section alternative versions of the model and discuss their implications for our identification strategy.

## E.1 Firms set price first, monopolistic competition

Let us first consider the opposite of our baseline assumption: prices are set first, before demand shocks are realized. Once the demand shock is observed, firms then choose quantities. The maximization problem becomes:

$$\max_p \int \pi_{ijkt} dG_{t-1}(a_{ijkt}) \qquad \text{s.t.} \qquad q_{ijkt} = e^{a_{ijkt}} p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}$$

$$\max_p p_{ijkt}^{1-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} \mathbb{E}_{t-1}\left[e^{a_{ijkt}}\right] - \frac{w_{it}}{\varphi_{ikt}} \mathbb{E}_{t-1}\left[e^{a_{ijkt}}\right] p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} - F_{ijk}$$

From the FOC and the constraint we get:

$$\begin{aligned} p^*_{ijkt} &= \frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}} \\ q^*_{ijkt} &= e^{a_{ijkt}} \left(\frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} \end{aligned}$$

With constant price elasticity, firms choose prices as constant mark-ups over marginal costs: prices do not depend on sales, but solely on supply side characteristics. Quantities then adjust to the demand level. Therefore, if prices are determined before observing the demand shocks, while quantities can fully adjust to it, neither prices nor quantities depend on firm beliefs. We would get:

$$\begin{aligned} \varepsilon^q_{ijkt} &= \ln Z^q_{ijkt} = a_{ijkt} \\ \varepsilon^p_{ijkt} &= \ln Z^p_{ijkt} = 0 \end{aligned}$$

Regressing $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$ should generate insignificant $\widehat{\beta}$ coefficients and the absolute value of $\varepsilon^q_{ijkt}$ should not decrease with age.

## E.2 Firms set price first, oligopolistic competition

Second, we consider the case of an oligopolistic market structure with Bertrand competition (so still price first), to allow for variable markups. Formally, we assume that consumers in country $j$ maximize utility derived from the consumption of goods from $K$ sectors. Each sector is composed of a small enough set of differentiated varieties of product $k$:

$$\begin{aligned} U_j &= \mathbb{E}\sum_{t=0}^{+\infty}\beta^t \ln\left(C_{jt}\right), \text{ with } C_{jt} = \prod_{k=0}^{K} C_{jkt}^{\mu_k} \\ \text{and } C_{jkt} &= \left(\sum_{\Omega_{kt}} \left(e^{a_{ijkt}}\right)^{\frac{1}{\sigma_k}} q_{ijkt}(\omega)^{\frac{\sigma_k-1}{\sigma_k}} d\omega\right)^{\frac{\sigma_k}{(\sigma_k-1)}} \end{aligned}$$

with $\Omega_{kt}$ the (small enough) set of varieties of product $k$ available at time $t$. We assume that firms take income $Y_{jt}$ as constant, i.e. we assume that $K$ is large enough.

The upper tier utility maximization implies $C_{jkt} = \frac{\mu_k Y_{jt}}{P_{jkt}}$. It follows the demand in market $j$ at time $t$ for a variety of product $k$:

$$q_{ijkt} = e^{a_{ijkt}} p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} = C_{jkt} e^{a_{ijkt}} \frac{p_{ijkt}^{-\sigma_k}}{P_{jkt}^{-\sigma_k}}$$

with the price index of sector $k$ in country $j$ defined as:

$$P_{jkt} = \left(\sum_{\Omega_{kt}} e^{a_{ijkt}} p_{ijkt}^{1-\sigma_k} di\right)^{\frac{1}{1-\sigma_k}}$$

The firm maximization program writes:

$$\max_p \int \pi_{ijkt} dG_{t-1}(a_{ijkt}) \qquad \text{s.t.} \qquad q_{ijkt} = e^{a_{ijkt}} p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}$$

It follows:

$$\begin{aligned} p_{ijkt}^* &= \frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\frac{w_{it}}{\varphi_{ikt}} \\ q_{ijkt}^* &= e^{a_{ijkt}}\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k}\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} \end{aligned}$$

with

$$\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right] = \sigma_k - (\sigma_k - 1)\,\mathbb{E}_{t-1}\left[s_{ijkt}\right]$$

where $\mathbb{E}_{t-1}\left[s_{ijkt}\right]$ is the expected market share at the beginning of period $t$. The residuals from the estimation in logs with fixed effects are:

$$\begin{aligned} \varepsilon_{ijkt}^q &= a_{ijkt} - \sigma_k \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{E_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right) \\ \varepsilon_{ijkt}^p &= \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right) \end{aligned}$$

As the demand shock now appears in the residual quantities, we regress $\varepsilon_{ijkt}^q$ on $\varepsilon_{ijkt}^p$:

$$a_{ijkt} - \sigma_k \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right] - 1}\right) = \beta\left(\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right] - 1}\right)\right) + \lambda_{ijk} + v_{ijkt}$$

We obtain:

$$\widehat{\beta} = -\sigma_k \text{ and } \widehat{v}_{ijkt} = \varepsilon_{ijkt}$$

and

$$\widehat{\lambda_{ijk}} + \widehat{v}_{ijkt} = a_{ijkt}$$

To test this alternative specification, we look at the dynamics of prices, that reflect the evolution of firms beliefs:

$$\Delta\varepsilon^p_{ijkt+1} = \Delta \ln\left(\frac{E_t\left[\varepsilon(s_{ijkt})\right]}{E_t\left[\varepsilon(s_{ijkt})\right] - 1}\right)$$

As $\frac{\partial\left(\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]\right)}{\partial\left(\mathbb{E}_{t-1}\left[s_{ijkt}\right]\right)} < 0$, a positive shock (i.e. generating a positive updating) implies a decrease in the expected price elasticity and an increase in markup. In Table A.6, we assess the empirical relevance of this alternative model. We do not find evidence of a positive relationship between prices and demand shocks. Overall, our data are therefore not consistent with the assumption of firms choosing their price first.

Table A.6: Prediction 1: demand shocks and beliefs updating (assuming Bertrand)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^p_{ijkt+1}$ | | |
| $a_{ijkt}$ | -0.004$^a$ | -0.005$^a$ | -0.005$^a$ | |
| | (0.001) | (0.001) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | 0.000 | 0.000 | |
| | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | -0.003$^a$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | | -0.001$^a$ |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 4$ | | | | -0.001$^a$ |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 5$ | | | | -0.000 |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 6$ | | | | -0.001$^a$ |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 7$ | | | | -0.001$^a$ |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 8$ | | | | -0.000 |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 9$ | | | | -0.000 |
| | | | | (0.000) |
| $\times$ Age$_{ijkt} = 10$ | | | | 0.001 |
| | | | | (0.000) |
| Age$_{ijkt}$ | 0.001$^a$ | 0.001$^a$ | 0.001$^a$ | |
| | (0.000) | (0.000) | (0.000) | |
| Observations | 1883748 | 1883748 | 1883748 | 1883748 |

Robust standard errors clustered by firm in parentheses (bootstrapped in column (3)). $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (4) but coefficients not reported. Shocks $a_{ijkt}$ are computed assuming Bertrand competition, i.e. by regressing $\varepsilon^q_{ijkt}$ on $\varepsilon^p_{ijkt}$ instead of the opposite. See text for more details. Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit).

## E.3 Partial quantity adjustment

Here, we maintain our assumption that quantities are set first, but allow firms to observe part of the demand shock before taking their quantity decision. Prices then fully adjust once the other part of the demand shock is observed.

Suppose that the demand shock $a_{ijkt}$ can be decomposed into 2 components: $a_{ijkt} = a^1_{ijkt} + a^2_{ijkt}$, with $a^1_{ijkt} \sim \mathcal{N}\left(\overline{a}^1_{ijk}, \varsigma\sigma^2_\varepsilon\right)$, $a^2_{ijkt} \sim \mathcal{N}\left(\overline{a}^2_{ijk}, (1-\varsigma)\,\sigma^2_\varepsilon\right)$ and $\overline{a}^1_{ijk} + \overline{a}^2_{ijk} = \overline{a}_{ijk}$. Firms can observe $a^1_{ijkt}$ before taking their quantity decision. $a^2_{ijkt}$ is then realized and firms fully adjust their prices. For simplicity, we assume that $a^1_{ijkt}$ does not bring additional information, i.e. $Cov(a^1_{ijkt}, a^2_{ijkt}) = 0$.

$\overline{a}^1_{ijk}$ and $\varsigma$ capture the relative importance of the first (observed) shock and therefore the importance of the learning process for firms: if $a^1_{ijkt}$ captures the entire demand shock ($\overline{a}_{ijk} = \overline{a}^1_{ijk}$ and $\varsigma = 1$), there is nothing to learn about. Beliefs are only related to $a^2_{ijkt}$, the part of the demand shock which is not observed at the time of the quantity decision. The distribution of beliefs is now described by $G_{t-1}(a^2_{ijkt})$.

After having observed $a^1_{ijkt}$, firms choose quantities by maximizing expected profits subject to demand. We get:

$$\max_q \int \pi_{ijkt} dG_{t-1}(a^2_{ijkt}) = \max_q q_{ijkt}^{1-\frac{1}{\sigma_k}} \left(\frac{\mu_k Y_{jt}}{P^{1-\sigma_k}_{jkt}}\right)^{\frac{1}{\sigma_k}} e^{\frac{a^1_{ijkt}}{\sigma_k}}\, \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right] - \frac{w_{it}}{\varphi_{ikt}} q_{ijkt} - F_{ijk}.$$

The constraint can now be written $p_{ijkt} = \left(\frac{\mu_k Y_{jt} e^{a^1_{ijkt}} e^{a^2_{ijkt}}}{q_{ijkt} P^{1-\sigma_k}_{jkt}}\right)^{\frac{1}{\sigma_k}}$. From the FOC and the constraint we get:

$$\begin{aligned} p^*_{ijkt} &= \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right)\left(\frac{e^{\frac{a^2_{ijkt}}{\sigma_k}}}{\mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]}\right) \\ q^*_{ijkt} &= \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k}\left(\frac{\mu_k Y_{jt}}{P^{1-\sigma_k}_{jkt}}\right) e^{a^1_{ijkt}}\mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]^{\sigma_k}. \end{aligned}$$

As before, quantities depend on firms' beliefs while prices are still a constant markup over marginal cost in expected terms. We get:

$$\begin{aligned} \varepsilon^q_{ijkt} &= \ln Z^q_{ijkt} = a^1_{ijkt} + \sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right] \\ \varepsilon^p_{ijkt} &= \ln Z^p_{ijkt} = \frac{1}{\sigma_k} a^2_{ijkt} - \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]. \end{aligned}$$

Note that if $\overline{a}^1_{ijk} = \overline{a}_{ijk}$ and $\varsigma = 1$, all the demand shock is observed and $\varepsilon^q_{ijkt}$ captures the demand shock only while $\varepsilon^p_{ijkt}$ does not depend neither on the demand shock, nor on firm beliefs (which are irrelevant in that case). This case is equivalent to the one where prices are set first. If on the other hand $\overline{a}^1_{ijk} = \varsigma = 0$, we are back to our baseline assumption of fixed quantities. Importantly, all our theoretical predictions still hold in the intermediate case. In particular, equation (11) still describes the evolution of beliefs, which are now related to the distribution of $a^2_{ijkt}$.

**Identification**. If quantities can partly adjust, $\varepsilon^q_{ijkt}$ captures both the firm beliefs and part of the demand shock, i.e. our measure of beliefs becomes noisy. This is innocuous when looking at the dynamics of $\varepsilon^q_{ijkt}$ (see 6.1) or when looking at the relationship between the variance of growth rates and age cohorts (see 6.2), but it has implications for the identification of the demand shocks $v_{ijkt}$. Regressing $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$ gives:

$$\left(\frac{1}{\sigma_k}a^2_{ijkt} - \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]\right) = \beta\left(a^1_{ijkt} + \sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]\right) + \lambda_{ijk} + v_{ijkt}.$$

It follows:

$$\widehat{\beta} = -\frac{1}{\sigma_k}\Lambda_P \text{ with } \Lambda_P = \frac{\mathbb{V}\left(\sigma_k \ln \widetilde{\mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]}\right)}{\mathbb{V}\left(\sigma_k \ln \widetilde{\mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]}\right) + \mathbb{V}\left(\widetilde{a^1_{ijkt}}\right)}$$

where variables with the sign$\,\widetilde{}\,$ are demeaned in the $ijk$ dimension. We get $0 < \Lambda_P < 1$: $\widehat{\beta}$ is underestimated due to the attenuation bias introduced by the noisy measure of firms' beliefs.

Hence, the estimated shock $\widehat{\lambda_{ijk}} + \widehat{v}_{ijkt}$ may be biased, but the direction of this bias is unclear as we would like now to isolate $a^2_{ijkt}$ and not $a_{ijkt} = a^1_{ijkt} + a^2_{ijkt}$. Indeed, firms now form beliefs about the part of the demand shock which is not observed at the time of the quantity decision. $\widehat{v}_{ijkt}$ may thus be larger or smaller than $a^2_{ijkt}$.

Suppose for instance that $\Lambda_P = 1/2$. This implies that $\mathbb{V}\left(\sigma_k \ln \widetilde{\mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]}\right) = \mathbb{V}\left(\widetilde{a^1_{ijkt}}\right)$. In this case our estimated demand shock would be:

$$\widehat{\lambda_{ijk}} + \widehat{v}_{ijkt} = \frac{1}{\sigma_k}a^2_{ijkt} + 2\sigma_k\left(a^1_{ijkt} - \sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]\right)$$

The direction of the bias depends on, among $\sigma_k \ln \mathbb{E}_{t-1}\left[e^{\frac{a^2_{ijkt}}{\sigma_k}}\right]$ and $a^1_{ijkt}$, which one is the most important component of $\varepsilon^q_{ijkt}$.

**Equation under test**. We obtain:

$$\Delta\varepsilon^q_{ijkt+1} = \sigma_k \Delta \ln \mathrm{E}_t \left[ e^{\frac{a^2_{ijkt+1}}{\sigma_k}} \right] + \Delta a^1_{ijkt+1}$$

It is worth noting that $\Delta\varepsilon^q_{ijkt}$ still fully captures the updating process, as $\Delta a^1_{ijkt+1} = 0$ in expected terms. It is now about the true value of $\overline{a}^2_{ijk}$. We get: $\Delta\widetilde{\theta}_t = g_t \left( a^2_{ijkt} - \widetilde{\theta}_{t-1} \right)$, and thus $\Delta \ln \mathrm{E}_t \left[ e^{\frac{a^2_{ijkt+1}}{\sigma_k}} \right] = \frac{1}{\sigma_k} \left( \Delta\widetilde{\theta}_t + \frac{\widetilde{\sigma}^2_t - \widetilde{\sigma}^2_{t-1}}{2\sigma_k} \right)$.

It follows:

$$\Delta\varepsilon^q_{ijkt+1} = g_t \left( a^2_{ijkt} - \widetilde{\theta}_{t-1} - \frac{\widetilde{\sigma}^2_{t-1} - \widetilde{\sigma}^2_t}{g_t 2\sigma_k} \right) + \Delta a^1_{ijkt+1}$$

Further,

$$\widetilde{\theta}_{ijkt-1} = \varepsilon^q_{ijkt} - \frac{\widetilde{\sigma}^2_{t-1} + \sigma^2_\varepsilon}{2\sigma_k} - a^1_{ijkt}$$

Hence:

$$\Delta\varepsilon^q_{ijkt+1} = g_t \left( a^2_{ijkt} - \varepsilon^q_{ijkt} \right) + g_t \left( \frac{\sigma^2_\varepsilon}{2\sigma_k} + a^1_{ijkt} \right) + \Delta a^1_{ijkt+1}$$

The possibility of some partial quantity adjustment may just generate some extra unconditional growth at the firm-market level (see the second term), as $\Delta a^1_{ijkt+1} = 0$ in expected terms. So, beyond the fact that $a^2_{ijkt}$ may be biased upwards or downwards, our strategy is left unaffected.

How this potential bias may affect our results on beliefs updating (prediction 1)? Consider our baseline specification, equation (19). There are two distinct issues here. First, $\widehat{\alpha}_1$ – the average extent of belief updating – might be upward or downward biased, depending on the direction of the bias of our estimated demand shocks. Second, if this bias depends on firm-market age, this may affect how $\widehat{\alpha}_1$ evolves with age, which is key for our findings.

As discussed in the main text, a simple way to gauge the importance of this issue is to focus on sectors or destinations for which quantities are more likely to be rigid (those for which $\widehat{\lambda_{ijk}} + \widehat{v}_{ijkt}$ is more likely to be correctly estimated) and to compare the results with our baseline estimates of Table 2. We expect less quantity adjustment for complex goods (in which many different relationship-specific inputs are used in the production process) and in destinations characterized by longer time-to-ship. In Table A.7, we restrict our sample to sectors or destinations which are above the sample median in terms of time-to-ship or input complexity. Data on sector-specific complexity comes from Nunn (2007), and data on time-to-ship between France's main port (Le Havre) and each of the destinations' main port from Berman *et al.* (2013).

Results in Table A.7 show that the updating of the firms' beliefs following a demand shock is quantitatively close to our baseline estimates (columns (1) and (4)), which suggests that the bias of our estimated demand shocks, if any, is limited. Further, the coefficient on

Table A.7: Prediction 1: robustness (fixed quantities)

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | $\Delta\varepsilon^q_{ijk,t+1}$ | |
| Sample | | Long time-to-ship | | | Complex goods | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.072$^a$ | 0.081$^a$ | | 0.077$^a$ | 0.084$^a$ | |
| | (0.002) | (0.003) | | (0.002) | (0.003) | |
| $\times$ Age$_{ijkt}$ | | -0.003$^b$ | | | -0.002$^a$ | |
| | | (0.001) | | | (0.001) | |
| $\times$ Age$_{ijkt} = 2$ | | | 0.075$^a$ | | | 0.080$^a$ |
| | | | (0.002) | | | (0.002) |
| $\times$ Age$_{ijkt} = 3$ | | | 0.071$^a$ | | | 0.078$^a$ |
| | | | (0.003) | | | (0.002) |
| $\times$ Age$_{ijkt} = 4$ | | | 0.068$^a$ | | | 0.075$^a$ |
| | | | (0.004) | | | (0.003) |
| $\times$ Age$_{ijkt} = 5$ | | | 0.062$^a$ | | | 0.070$^a$ |
| | | | (0.005) | | | (0.004) |
| $\times$ Age$_{ijkt} = 6$ | | | 0.069$^a$ | | | 0.074$^a$ |
| | | | (0.005) | | | (0.004) |
| $\times$ Age$_{ijkt} = 7$ | | | 0.068$^a$ | | | 0.071$^a$ |
| | | | (0.008) | | | (0.004) |
| $\times$ Age$_{ijkt} = 8$ | | | 0.062$^a$ | | | 0.072$^a$ |
| | | | (0.010) | | | (0.005) |
| $\times$ Age$_{ijkt} = 9$ | | | 0.063$^a$ | | | 0.065$^a$ |
| | | | (0.013) | | | (0.008) |
| $\times$ Age$_{ijkt} = 10$ | | | 0.040$^b$ | | | 0.079$^a$ |
| | | | (0.018) | | | (0.012) |
| Age$_{ijkt}$ | -0.029$^a$ | -0.029$^a$ | | -0.031$^a$ | -0.031$^a$ | |
| | (0.001) | (0.001) | | (0.001) | (0.001) | |
| Observations | 358418 | 358418 | 358418 | 800015 | 800015 | 800015 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Age dummies included alone in columns (3) and (6) but coefficients not reported. Complex goods means in the above the sample median of the variable, and large time-to-ship above the median for extra-EU observations. Data on sector-specific complexity comes from Nunn (2007), and data on time-to-ship between France's main port (Le Havre) and each of the destinations' main port from Berman *et al.* (2013).

the interaction term between demand shocks and age (columns (2)-(3) and (5)-(6)), is also similar our baseline estimates. The coefficient on the interaction term between demand shocks and age is slightly lower than our baseline in the case of complex goods (col. (5) of Table A.7). In column (6), however, we see that this result is only driven by effect of the last age category, 10 years of experience, which is itself quite imprecisely estimated.

Altogether, these results suggest that our assumption of fixed quantities should not be rejected in light of our data.

## E.4 Oligopolistic Competition - Cournot

We investigate here the possibility that firms markups are variable. To do so, we consider the same model as before, but assume that competition is oligopolistic within sectors and not monopolistic. Formally, we assume that consumers in country $j$ maximize utility derived from the consumption of goods from $K$ sectors. Each sector is composed of a small enough set of differentiated varieties of product $k$:

$$\begin{aligned} U_j &= E\sum_{t=0}^{+\infty} \beta^t \ln\left(C_{jt}\right), \text{ with } C_{jt} = \prod_{k=0}^{K} C_{jkt}^{\mu_k} \\ \text{and } C_{jkt} &= \left(\sum_{\Omega_{kt}} \left(e^{a_{ijkt}}\right)^{\frac{1}{\sigma_k}} q_{ijkt}(\omega)^{\frac{\sigma_k - 1}{\sigma_k}} d\omega\right)^{\frac{\sigma_k}{(\sigma_k-1)}} \end{aligned}$$

with $\rho$ the discount factor. $\Omega_{kt}$ the (small enough) set of varieties of product $k$ available at time $t$, and $\sum^K \mu_k = 1$. We assume that firms take income $Y_{jt}$ as constant, i.e. we assume that $K$ is large enough.

The upper tier utility maximization implies $C_{jkt} = \frac{\mu_k Y_{jt}}{P_{jkt}}$. It follows the demand in market $j$ at time $t$ for a variety of product $k$:

$$q_{ijkt} = e^{a_{ijkt}} p_{ijkt}^{-\sigma_k} \frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}} = C_{jkt} e^{a_{ijkt}} \frac{p_{ijkt}^{-\sigma_k}}{P_{jkt}^{-\sigma_k}}$$

with the price index of sector $k$ in country $j$ defined as:

$$P_{jkt} = \left(\sum_{\Omega_{kt}} e^{a_{ijkt}} p_{ijkt}^{1-\sigma_k} di\right)^{\frac{1}{1-\sigma_k}}$$

**Equilibrium**. Firms maximize profits, given the demand they face. They maximize:

$$\max_q \int \pi_{ijkt} dG_{t-1}(a_{ijkt}) \qquad \text{s.t.} \qquad p_{ijkt} = \left(\frac{C_{jkt} e^{a_{ijkt}}}{q_{ijkt}}\right)^{\frac{1}{\sigma_k}} \frac{\mu_k Y_{jt}}{C_{jkt}}$$

We get:

$$\frac{\partial \int \pi_{ijkt} dG_{t-1}(a_{ijkt})}{\partial q_{ijkt}} = \left(\left(C_{jkt}\right)^{\frac{1}{\sigma_k}} q_{ijkt}^{-\frac{1}{\sigma}} \frac{\mu_k Y_{jt}}{C_{jkt}} \mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma}}\right]\right)\left(1 - \frac{1}{\sigma_k}\right)\left(1 - \mathbb{E}_{t-1}\left[s_{ijkt}\right]\right) - \frac{w_{it}}{\varphi_{ikt}}$$

where $\mathbb{E}_{t-1}\left[s_{ijkt}\right]$ is the expected market share at the beginning of period $t$:

$$\mathbb{E}_{t-1}\left[s_{ijkt}\right] = \frac{\mathbb{E}_{t-1}\left[q_{ijkt} p_{ijkt}\right]}{C_{jkt} P_{jkt}} = \frac{\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma}}\right] \mathbb{E}_{t-1}\left[p_{ijkt}^{1-\sigma_k}\right]}{P_{jkt}^{1-\sigma_k}}$$

It follows:

$$
\begin{aligned}
q^*_{ijkt} &= \left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k}\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}\left(\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right)^{\sigma_k}\\
p^*_{ijkt} &= \frac{e^{\frac{a_{ijkt}}{\sigma_k}}}{\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]}\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\frac{w_{it}}{\varphi_{ikt}}\right)
\end{aligned}
$$

with

$$\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]=\frac{1}{\frac{1}{\sigma_k}+\left(1-\frac{1}{\sigma_k}\right)\mathbb{E}_{t-1}\left[s_{ijkt}\right]}$$

**Identification**. Purged quantities and prices:

$$
\begin{aligned}
\varepsilon^q_{ijkt} &= \sigma_k\left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]-\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right)\right)\\
\varepsilon^p_{ijkt} &= \frac{a_{ijkt}}{\sigma_k}-\left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]-\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right)\right)
\end{aligned}
$$

We regress $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$:

$$
\begin{aligned}
&\frac{a_{ijkt}}{\sigma_k}-\left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]-\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right)\right)\\
&=\beta\left(\sigma_k\left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]-\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right)\right)\right)+\lambda_{ijk}+v_{ijkt}
\end{aligned}
$$

We get:

$$\widehat{\beta}=-\frac{1}{\sigma_k}\quad\text{and}\quad\widehat{v}_{ijkt}=\frac{1}{\sigma_k}\varepsilon_{ijkt}$$

And

$$\widehat{\lambda_{ijk}}+\widehat{v}_{ijkt}=\frac{a_{ijkt}}{\sigma_k}$$

Put differently, our strategy to identify demand signals is still valid if firms markups are variable. This is because firms' markups affect purged prices and quantities in the same way as beliefs do.

**Equation under test**. We obtain:

$$\Delta\varepsilon^q_{ijkt+1}=\sigma_k\left[\Delta\ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]-\Delta\ln\left(\frac{E_t\left[\varepsilon(s_{ijkt})\right]}{E_t\left[\varepsilon(s_{ijkt})\right]-1}\right)\right]$$

$\Delta\varepsilon^q_{ijkt}$ does not only capture the updating process, but is also impacted by changes in expected mark-up.

The updating process itself does not change however, we still get $\Delta\widetilde{\theta}_t = g_t\left(a_{ijkt} - \widetilde{\theta}_{t-1}\right)$, and $\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] = \frac{1}{\sigma_k}\left(\Delta\widetilde{\theta}_t + \frac{\widetilde{\sigma}_t^2 - \widetilde{\sigma}_{t-1}^2}{2\sigma_k}\right)$.

It follows

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(\left(a_{ijkt} - \widetilde{\theta}_{t-1}\right) - \frac{\widetilde{\sigma}_{t-1}^2 - \widetilde{\sigma}_t^2}{g_t 2\sigma_k}\right) - \sigma_k \Delta \ln\left(\frac{E_t\left[\varepsilon(s_{ijkt})\right]}{E_t\left[\varepsilon(s_{ijkt})\right] - 1}\right)$$

Further,

$$\widetilde{\theta}_{ijkt-1} = \varepsilon^q_{ijkt} - \frac{\widetilde{\sigma}_{t-1}^2 + \sigma_\varepsilon^2}{2\sigma_k} + \sigma_k \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right] - 1}\right)$$

And we obtain:

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t\frac{\sigma_\varepsilon^2}{2\sigma_k} - \sigma_k\left(g_t \ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right] - 1}\right) + \Delta \ln\left(\frac{E_t\left[\varepsilon(s_{ijkt})\right]}{E_t\left[\varepsilon(s_{ijkt})\right] - 1}\right)\right)$$

With variable mark-ups, our main equation includes two new terms.

The first term is the level of the expected mark-ups, $\ln\left(\frac{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]}{\mathbb{E}_{t-1}\left[\varepsilon(s_{ijkt})\right]-1}\right)$. This term comes from the fact that the expected mark-up also affects our measure of beliefs, $\varepsilon^q_{ijkt}$. As it is a component of $\varepsilon^q_{ijkt}$ and therefore of $\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)$, we need to control for it to avoid a standard omitted variable bias. Hence, we need to control for $\varepsilon^q_{ijkt}$ in the estimation.

The second term captures the change in expected mark-ups $\Delta \ln\left(\frac{E_t\left[\varepsilon(s_{ijkt})\right]}{E_t\left[\varepsilon(s_{ijkt})\right]-1}\right)$, and it depends on the updating process through the change in the expected market share. It follows that our measure of beliefs updating is now underestimated as the quantity reaction to a demand shock is dampened by the mark-up reaction: when firms update positively, they tend to increase their quantities but also their prices, which dampens their overall quantity reaction. Under very weak conditions however, the quantity reaction to beliefs updating is still positive (see the proof below). It means that in the case of variable mark-ups, what we interpret quantitatively as beliefs updating becomes the overall reaction of purged quantities $\varepsilon^q_{ijkt}$ to belief updating. $\varepsilon^q_{ijkt}$ becomes an increasing function of firm's beliefs, but cannot be seen as identical to firm's beliefs. Thus, our results still provide evidence for the updating process, but in a qualitative sense.

Importantly, the relation that goes from beliefs to expected markups (through the expected market share) is not log linear. Put differently, two firms of different sizes, but updating in the exact same way, will not have the same mark-up reaction to this updating. This means that we need again to control for firm size, to be able to compare the beliefs updating of firms with the same initial market share.

**Proof**. Quantity increase after a positive updating: $\frac{\partial\left(\Delta\varepsilon^q_{ijkt}\right)}{\partial\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right)} > 0$

First, we express $\varepsilon^q_{ijkt}$ in terms of beliefs and market share. We have:

$$\varepsilon^q_{ijkt} = \sigma_k \left( \ln \mathrm{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] - \ln \left( \frac{\mathbb{E}_{t-1} \left[ \varepsilon(s_{ijkt}) \right]}{\mathbb{E}_{t-1} \left[ \varepsilon(s_{ijkt}) \right] - 1} \right) \right)$$

Note that:

$$\ln \left( \frac{\mathbb{E}_{t-1} \left[ \varepsilon(s_{ijkt}) \right]}{\mathbb{E}_{t-1} \left[ \varepsilon(s_{ijkt}) \right] - 1} \right) = - \ln \left( 1 - \frac{1}{\sigma_k} \right) - \ln \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right)$$

As we have purged $\varepsilon^q_{ijkt}$ of its $ikt$ components, we get:

$$\varepsilon^q_{ijkt} = \sigma_k \left( \ln \mathrm{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] + \ln \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right) \right)$$

Second, let's find the relation between beliefs and market share. Market share is given by:

$$\mathbb{E}_{t-1} \left[ s_{ijkt} \right] = \frac{\mathbb{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] \mathbb{E}_{t-1} \left[ p^{1-\sigma_k}_{ijkt} \right]}{P^{1-\sigma_k}_{jkt}}$$

And expected price:

$$\mathbb{E}_{t-1} \left[ p^{1-\sigma_k}_{ijkt} \right] = \mathbb{E}_{t-1} \left[ \left( \frac{w_{it}}{\varphi_{ikt}} \frac{\sigma_k}{\sigma_k - 1} \right)^{1-\sigma_k} \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right)^{\sigma_k - 1} \right]$$

Given that we work with purged prices and quantities, we obtain:

$$\begin{aligned} \mathbb{E}_{t-1} \left[ s_{ijkt} \right] &= \mathbb{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] \mathbb{E}_{t-1} \left[ p^{1-\sigma_k}_{ijkt} \right] \\ \mathbb{E}_{t-1} \left[ p^{1-\sigma_k}_{ijkt} \right] &= \mathbb{E}_{t-1} \left[ \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right)^{\sigma_k - 1} \right] \end{aligned}$$

It follows that $\mathbb{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma}} \right] = \frac{\mathbb{E}_{t-1}\left[ s_{ijkt} \right]}{\left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right)^{\sigma - 1}}$ and we obtain:

$$\ln \mathrm{E}_{t-1} \left[ e^{\frac{a_{ijkt}}{\sigma_k}} \right] = \ln \mathbb{E}_{t-1} \left[ s_{ijkt} \right] - (\sigma_k - 1) \ln \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right)$$

As the expected market share is an increasing function of the beliefs, we only need to show that $\Delta \varepsilon^q_{ijkt+1}$ is increasing in firm's expected market share.

Third, we can now express $\varepsilon^q_{ijkt}$ as a function of the expected market share only:

$$\varepsilon^q_{ijkt} = \sigma_k \left( \ln \mathbb{E}_{t-1} \left[ s_{ijkt} \right] - (\sigma_k - 2) \ln \left( 1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right] \right) \right)$$

We get:

$$\frac{\partial \varepsilon^q_{ijkt}}{\partial \mathbb{E}_{t-1} \left[ s_{ijkt} \right]} = \sigma_k \left( \frac{1}{\mathbb{E}_{t-1} \left[ s_{ijkt} \right]} + (\sigma_k - 2) \frac{1}{1 - \mathbb{E}_{t-1} \left[ s_{ijkt} \right]} \right)$$

It follows that $\varepsilon^q_{ijkt}$ is an increasing function of the expected market share if:

$$\frac{\partial \varepsilon^q_{ijkt}}{\partial \mathbb{E}_{t-1}\left[s_{ijkt}\right]} > 0 \Leftrightarrow 1 + \left(\sigma_k - 3\right) \mathbb{E}_{t-1}\left[s_{ijkt}\right] > 0$$

This condition is necessarily fulfilled if $\sigma_k > 2$. If $\sigma_k < 2$, we can concentrate on the limiting case $\sigma_k = 1$. $\mathbb{E}_{t-1}\left[s_{ijkt}\right] < 1/2$ provides another sufficient condition for the above condition to hold.

## E.5 Product-destination specific productivity

Here, we introduce a product-destination component to productivity. Specifically, we assume that the unit cost of producing good $k$ for market $j$ at time $t$ is:

$$\frac{w_{it}}{\varphi_{ikt}} \frac{1}{\varphi_{ijkt}}$$

with $\varphi_{ijkt} > 0$ and where $\frac{1}{\varphi_{ijkt}}$ can be understood as a cost wedge for market $jk$ with respect to the average cost of this good. Further, it could also capture differences in product quality for the same good across markets. Finally, it could capture differences in transportation costs between French competitors in market $jk$ at time $t$.

**Equilibrium**. The optimal price and quantities are given by:

$$\begin{aligned} q^*_{ijkt} &= \left(\frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}\varphi_{ijkt}}\right)^{-\sigma_k} \left(\frac{\mu_k Y_{jt}}{P^{1-\sigma_k}_{jkt}}\right) \left(\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right)^{\sigma_k} \\ p^*_{ijkt} &= \left(\frac{\sigma_k}{\sigma_k - 1} \frac{w_{it}}{\varphi_{ikt}\varphi_{ijkt}}\right) \left(\frac{e^{\frac{a_{ijkt}}{\sigma_k}}}{\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]}\right) \end{aligned}$$

**Identification**. Purged quantities and prices are:

$$\begin{aligned} \varepsilon^q_{ijkt} &= \sigma_k \left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] + \ln\left(\varphi_{ijkt}\right)\right) \\ \varepsilon^p_{ijkt} &= \frac{1}{\sigma_k} a_{ijkt} - \left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] + \ln \varphi_{ijkt}\right) \end{aligned}$$

We regress $\varepsilon^p_{ijkt}$ on $\varepsilon^q_{ijkt}$:

$$\frac{1}{\sigma_k} a_{ijkt} - \left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] + \ln \varphi_{ijkt}\right) = \beta\left(\sigma_k \left(\ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] + \ln\left(\varphi_{ijkt}\right)\right)\right) + \lambda_{ijk} + v_{ijkt}$$

We obtain:

$$\widehat{\beta} = -\frac{1}{\sigma_k} \quad \text{and} \quad \widehat{v}_{ijkt} = \frac{1}{\sigma_k}\varepsilon_{ijkt}$$

And

$$\widehat{\lambda_{ijk}} + \widehat{v}_{ijkt} = \frac{1}{\sigma_k} a_{ijkt}$$

Our identification strategy is still valid if productivity incorporates a $ijk$ component.

**Equation under test**. We now get:

$$\Delta\varepsilon^q_{ijkt+1} = \sigma_k \left[\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] + \Delta \ln\left(\varphi_{ijkt+1}\right)\right]$$

And $\Delta\varepsilon^q_{ijkt+1}$ cannot be seen as updating only. The updating process itself does not change, we still get $\Delta\widetilde{\theta}_t = g_t\left(a_{ijkt} - \widetilde{\theta}_{t-1}\right)$, and thus $\Delta\,\ln\mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] = \frac{1}{\sigma_k}\left(\Delta\widetilde{\theta}_t + \frac{\widetilde{\sigma}^2_t - \widetilde{\sigma}^2_{t-1}}{2\sigma_k}\right)$.

It follows

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \widetilde{\theta}_{t-1} - \frac{\widetilde{\sigma}^2_{t-1} - \widetilde{\sigma}^2_t}{g_t 2\sigma_k}\right) + \sigma_k \Delta\ln\left(\varphi_{ijkt+1}\right)$$

Further,

$$\widetilde{\theta}_{ijkt-1} = \varepsilon^q_{ijkt} - \frac{\widetilde{\sigma}^2_{t-1} + \sigma^2_\varepsilon}{2\sigma_k} - \sigma_k \ln\left(\varphi_{ijkt}\right)$$

Hence:

$$\Delta\varepsilon^q_{ijkt+1} = g_t\left(a_{ijkt} - \varepsilon^q_{ijkt}\right) + g_t\frac{\sigma^2_\varepsilon}{2\sigma_k} + \sigma_k\left[g_t\ln\left(\varphi_{ijkt}\right) + \Delta\ln\left(\varphi_{ijkt+1}\right)\right]$$

As for the case of variable mark-ups, our equation includes two new terms.

This term comes from the fact that $\ln\left(\varphi_{ijkt}\right)$ affects our measure of beliefs, $\varepsilon^q_{ijkt}$. As it is a component of $\varepsilon^q_{ijkt}$ and thus of $\left(a_{ijkt} - \varepsilon^q_{ijkt}\right)$, we need to control for firm size to avoid a standard omitted variable bias.

The presence of the second term, $\Delta\ln\left(\varphi_{ijkt+1}\right)$ comes from the fact that $\Delta\varepsilon^q_{ijkt+1}$ also reflects the dynamics of productivity.

If $\Delta\ln\left(\varphi_{ijkt+1}\right)$ is uncorrelated with the updating process, the interpretation of our results should be unaffected. If however $\Delta\ln\left(\varphi_{ijkt+1}\right)$ is positively affected by the updating process, because a positive updating would lead firms to invest to improve $\varphi_{ijkt}$, then our measure of updating becomes a measure of the overall impact of the updating process on $\Delta\varepsilon^q_{ijkt+1}$: it would not only capture the updating process itself but also how the quantity response is magnified by a change in productivity. $\varepsilon^q_{ijkt}$ would become an increasing function of firm's beliefs, and our evidence of the updating process would become qualitative as we would not identify firms' beliefs, but only a function of it. As this productivity response could be size dependent, we need again to control for firm size. The decline of the overall reaction of $\Delta\varepsilon^q_{ijkt+1}$ to demand shocks over time, conditional on size, however still provides evidence for an updating process.

Finally, note that $\Delta\varepsilon^q_{ijkt+1}$ will also capture the dynamics of $\ln\left(\varphi_{ijkt}\right)$ that is uncorrelated with the updating process. In turn, it could introduce some noise into our measure of the updating process. But if this dynamics was important, it should be observed in the dynamics of $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$. Results in section 6.1 show however that, when we concentrate on the within firm-market dynamics of these elements (Figure 2.c), no important pattern emerge: $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ are roughly constant over time, which suggests that there should not be any important dynamics in $\ln\left(\varphi_{ijkt}\right)$, beyond the one possibly driven by the updating process.

## E.6 Controlling for size – robustness

This section presents additional results controlling for size when testing prediction 1.

Table A.8 reports the full set of coefficients corresponding to Table 4 in the main text. Columns (1)-(2) include market-year fixed effects in the estimation of price residuals $\varepsilon^{p}_{ijkt}$ in equation (13) as models with variable mark-ups would involve. Columns (3)-(6) additionally include controls for firms' market specific size and its interaction with age, introduced either linearly or non linearly through size bins by deciles. Size is measured by the market share of firm $i$ during year $t-1$ in total French exports in value to the market $jk$.

Table A.9 shows that our results are not sensitive to the measurement of firm size. In columns (1)-(4), we measure firm size as market shares in quantity and introduce it either linearly or through bins as in Table A.8. Alternatively, in columns (5)-(8) firm size is measured as the log of the value exported by firm $i$ to market $jk$ in year $t-1$.

Finally, Table A.10 includes an interaction term between size and $a_{ijkt} - \varepsilon^{q}_{ijk}$ to account for the fact that age and size are correlated. We report results using as a measure of firm size either the market share in value (columns (1)-(4)) and quantity (columns (5)-(8)) introduced linearly (odd columns) or through bins (even columns).

In all cases, the coefficients on $a_{ijkt} - \varepsilon^{q}_{ijk}$ and its interaction with age remain close to our benchmark results in Table A.8.

Table A.8: Prediction 1: controlling for size

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. var. | $\Delta\varepsilon^q_{ijk,t+1}$ | | | | | |
| Robustness | Controlling for FE$_{jkt}$ in prices | | Controlling for FE$_{jkt}$ in prices and size | | | |
| Size | | | Linear | | Bins | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.103$^a$ | | 0.103$^a$ | | 0.102$^a$ | |
| | (0.002) | | (0.002) | | (0.002) | |
| × Age$_{ijkt}$ | -0.003$^a$ | | -0.003$^a$ | | -0.003$^a$ | |
| | (0.000) | | (0.000) | | (0.000) | |
| × Age$_{ijkt} = 2$ | | 0.096$^a$ | | 0.096$^a$ | | 0.096$^a$ |
| | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 3$ | | 0.093$^a$ | | 0.093$^a$ | | 0.093$^a$ |
| | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 4$ | | 0.087$^a$ | | 0.087$^a$ | | 0.087$^a$ |
| | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 5$ | | 0.086$^a$ | | 0.086$^a$ | | 0.087$^a$ |
| | | (0.003) | | (0.003) | | (0.002) |
| × Age$_{ijkt} = 6$ | | 0.082$^a$ | | 0.082$^a$ | | 0.081$^a$ |
| | | (0.003) | | (0.002) | | (0.003) |
| × Age$_{ijkt} = 7$ | | 0.079$^a$ | | 0.079$^a$ | | 0.078$^a$ |
| | | (0.003) | | (0.003) | | (0.003) |
| × Age$_{ijkt} = 8$ | | 0.076$^a$ | | 0.076$^a$ | | 0.076$^a$ |
| | | (0.004) | | (0.004) | | (0.004) |
| × Age$_{ijkt} = 9$ | | 0.077$^a$ | | 0.076$^a$ | | 0.077$^a$ |
| | | (0.005) | | (0.005) | | (0.005) |
| × Age$_{ijkt} = 10$ | | 0.074$^a$ | | 0.074$^a$ | | 0.075$^a$ |
| | | (0.009) | | (0.009) | | (0.009) |
| Age$_{ijkt}$ | -0.034$^a$ | | -0.040$^a$ | | 0.019$^a$ | |
| | (0.001) | | (0.001) | | (0.002) | |
| Size$_{ijkt}$ | | | -1.053$^a$ | -1.015$^a$ | | |
| | | | (0.016) | (0.017) | | |
| × Age$_{ijkt}$ | | | 0.109$^a$ | 0.101$^a$ | | |
| | | | (0.003) | (0.003) | | |
| Observations | 1870377 | 1870377 | 1870377 | 1870377 | 1501840 | 1501840 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17). Compared to our baseline methodology, in this table we include $jkt$ fixed effects in the estimation of the price residuals $\varepsilon^p_{ijkt}$ used to identify demand shocks. $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Size$_{ijkt}$ is proxied by the value sold by firm $i$ on market $jk$ during year $t$ divided by the total value exported by French firms in market $jk$ during year $t$. Columns (5) and (6) include size bins corresponding to the ten deciles of size variable, computed by market-year. Age dummies included alone in columns (2), (4) and (6) but coefficients not reported.

Table A.9: Prediction 1: controlling for size (robustness 1/2)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\Delta\varepsilon^q_{ijk,t+1}$ | | | | | | | |
| Size | Market share (quantity) | | | | log values | | | |
| Functional form | Linear | | Bins | | Linear | | Bins | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.098$^a$ | | 0.082$^a$ | | 0.104$^a$ | | 0.103$^a$ | |
| | (0.002) | | (0.002) | | (0.002) | | (0.002) | |
| × Age$_{ijkt}$ | -0.003$^a$ | | -0.002$^a$ | | -0.004$^a$ | | -0.003$^a$ | |
| | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| × Age$_{ijkt} = 2$ | | 0.092$^a$ | | 0.079$^a$ | | 0.098$^a$ | | 0.097$^a$ |
| | | (0.002) | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 3$ | | 0.089$^a$ | | 0.077$^a$ | | 0.094$^a$ | | 0.093$^a$ |
| | | (0.002) | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 4$ | | 0.083$^a$ | | 0.071$^a$ | | 0.088$^a$ | | 0.087$^a$ |
| | | (0.002) | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 5$ | | 0.083$^a$ | | 0.072$^a$ | | 0.087$^a$ | | 0.087$^a$ |
| | | (0.002) | | (0.002) | | (0.002) | | (0.002) |
| × Age$_{ijkt} = 6$ | | 0.079$^a$ | | 0.068$^a$ | | 0.083$^a$ | | 0.082$^a$ |
| | | (0.002) | | (0.002) | | (0.002) | | (0.003) |
| × Age$_{ijkt} = 7$ | | 0.077$^a$ | | 0.066$^a$ | | 0.080$^a$ | | 0.078$^a$ |
| | | (0.003) | | (0.003) | | (0.003) | | (0.003) |
| × Age$_{ijkt} = 8$ | | 0.075$^a$ | | 0.065$^a$ | | 0.077$^a$ | | 0.076$^a$ |
| | | (0.004) | | (0.004) | | (0.004) | | (0.004) |
| × Age$_{ijkt} = 9$ | | 0.076$^a$ | | 0.068$^a$ | | 0.077$^a$ | | 0.078$^a$ |
| | | (0.005) | | (0.005) | | (0.005) | | (0.005) |
| × Age$_{ijkt} = 10$ | | 0.073$^a$ | | 0.065$^a$ | | 0.074$^a$ | | 0.074$^a$ |
| | | (0.009) | | (0.009) | | (0.009) | | (0.009) |
| Age$_{ijkt}$ | -0.039$^a$ | | 0.016$^a$ | | -0.148$^a$ | | 0.024$^a$ | |
| | (0.001) | | (0.002) | | (0.003) | | (0.002) | |
| Size$_{ijkt}$ | -0.891$^a$ | -0.857$^a$ | | | -0.184$^a$ | -0.180$^a$ | | |
| | (0.015) | (0.015) | | | (0.002) | (0.003) | | |
| × Age$_{ijkt}$ | 0.090$^a$ | 0.083$^a$ | | | 0.015$^a$ | 0.014$^a$ | | |
| | (0.002) | (0.002) | | | (0.000) | (0.000) | | |
| Observations | 1870377 | 1870377 | 1501840 | 1501840 | 1870377 | 1870377 | 1501840 | 1501840 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17). Compared to our baseline methodology, in this table we include $jkt$ fixed effects in the estimation of the price residuals $\varepsilon^p_{ijkt}$ used to identify demand shocks. $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). In columns (1)-(4), Size$_{ijkt}$ is proxied by the quantity sold by firm $i$ on market $jk$ during year $t$ divided by the total quantity exported by French firms in market $jk$ during year $t$. In columns (5)-(8), Size$_{ijkt}$ is proxied by the log of the value sold by firm $i$ on market $jk$ during year $t$. Columns (3), (4), (7) and (8) include size bins corresponding to the ten deciles of size variable, computed by market-year. Age dummies included alone in columns (2), (4), (6) and (8) but coefficients not reported.

# F Firm survival

This section develops the predictions of the learning model regarding firms' survival and provides evidence that the exit behavior of firms on specific markets is in line with the demand learning process.

A firm decides to stop exporting a particular product to a given destination whenever the expected value of the profits stream associated with this activity becomes negative. At the beginning of period $t$ (after having received $t-1$ signals), expected profits for period $t$ are given by:

$$E_{t-1}\left[\pi_{ijkt}\right] = \frac{C_{ikt}^{S} C_{jkt}^{S}}{\sigma_k} e^{\left(\widetilde{\theta}_{ijkt-1} + \frac{\widetilde{\sigma}_{ijkt-1}^{2} + \sigma_{\varepsilon}^{2}}{2\sigma_k}\right)} - F_{ijk}.$$

Of course, the exit decision also depends on the expected future stream of profits, which depends on the evolution of $C_{ikt}^{S}$, $C_{jkt}^{S}$, $\widetilde{\theta}_{ijkt-1}$ and $\widetilde{\sigma}_{ijkt-1}^{2}$ over time. Our assumption of normal prior beliefs provides the conditional distribution of $\widetilde{\theta}_{ijkt}$ given $\widetilde{\theta}_{ijkt-1}$ while the distribution of $\widetilde{\sigma}_{ijkt-1}^{2}$ is deterministic. So, the evolution of firms' beliefs can be summarized by $\widetilde{\theta}_{ijkt-1}$ and $t$. Up to now, we have made no assumption regarding the dynamics of the $C_{ikt}^{S}$ and $C_{jkt}^{S}$ terms. Here, to proceed further, we follow Hopenhayn (1992) and introduce some (mild) assumptions on their dynamics. We label $A_{ijkt} \equiv C_{ikt}^{S} C_{jkt}^{S}$ and we assume that: i) $A_{ijkt}$ follows a Markov process, ii) $A_{ijkt}$ is bounded and iii) the conditional distribution $F\left(A_{ijkt+1} \mid A_{ijkt}\right)$ is continuous in $A_{ijkt}$ and $A_{ijkt+1}$, and $F(.)$ is strictly decreasing in $A_{ijkt}$.[7]

The set of firm state variables at time $t$ can thus be summarized by $\Omega_{ijkt} = \left\{A_{ijkt}, \widetilde{\theta}_{ijkt-1}, t\right\}$. The value function of the firm $V_{ijk}\left(\Omega_{ijkt}\right)$ satisfies the following Bellman equation:

$$V_{ijk}\left(\Omega_{ijkt}\right) = \max\left\{\mathbb{E}\left[\pi_{ijkt}\left(\Omega_{ijkt}\right)\right] + \beta\mathbb{E}\left[V_{ijk}\left(\Omega_{ijkt+1} \mid \Omega_{ijkt}\right)\right], 0\right\} \tag{26}$$

where $\beta$ is the rate at which firms discount profits and where we have normalized the value of exiting to zero.[8] The value function $V_{ijk}$ is monotonically increasing in $A_{ijkt}$ and $\widetilde{\theta}_{ijkt-1}$.[9] Intuitively, the flow of future expected profits inherits the properties of expected profits at time $t$. It follows that there exists a threshold value $\underline{\widetilde{\theta}_{ijkt-1}}(A_{ijkt}, t)$ such that a firm exits market $jk$ at time $t$ if $\widetilde{\theta}_{ijkt-1} < \underline{\widetilde{\theta}_{ijkt-1}}(A_{ijkt}, t)$. This implies:

**Prediction # 4** (firm exit): *Given $A_{ijkt}$ and $t$ (firm age), (a) the probability to exit decreases with $\widetilde{\theta}_{ijkt-1}$ and (b) a given negative difference between realized and expected demand triggers less exit for older firms.*

The literature has usually associated learning with exit rates declining with age, and

[7] While not very demanding, these assumptions restrict the set of possible dynamics for firm productivity. In that sense, our results on firm exit decision are somewhat weaker than those about firm growth, which are robust to any dynamics of firm productivity.

[8] Here, we assume that an exiting firm loses all the information accumulated in the past. If the firm enters again market $jk$ in the future, new initial beliefs will be drawn. In consequence, we treat the exit decision as irreversible.

[9] See Hopenhayn (1992) and Jovanovic (1982).

we indeed find this to be the case in our estimations. However, this relation may not necessarily be monotonic (see Pakes and Ericson, 1998 for a discussion). The decision to exit not only depends on the extent of firm updating (which indeed declines with age) but also on how $\underline{\widetilde{\theta}_{ijkt-1}}(A_{ijkt}, t)$ evolves over time. If this threshold increases very rapidly for some $t$, the exit rate could actually increase temporarily. For old firms however, i.e. when beliefs become accurate, and conditional on $A_{ijkt}$ *and* $t$, the exit rate should tend to 0.

On the other hand, an important and general implication of our demand learning model is that negative demand shocks should trigger less exits for older firms (prediction 4.b). The reason is simply that firms' posterior beliefs $\widetilde{\theta}_{ijkt-1}$ depend less and less on demand shocks as firms age. Hence, the exit rate may not always be decreasing with age, but demand shocks should always have a lower impact on the exit decision in older cohorts, because they imply less updating. Note that this prediction can also be understood as another robustness check for our formulation of a passive learning model: in an active learning model, no matter the age of the firm, demand shocks may trigger new investments. Their impact on future expected profits stream should not be weakened for older firms (see Ericson and Pakes, 1995). This prediction is not directly tested in Pakes and Ericson (1998) because they use a much less parametric model than ours which prevents them to back out demand shocks and firms' beliefs. Their test is solely based on actual firm size.

To test prediction 4, note that from equation (5), $\widetilde{\theta}_{ijkt-1}$ depends positively on $\widetilde{\theta}_{ijkt-2}$ and $a_{ijkt-1}$. We want to test if, conditional on $A_{ijkt}$ and firm age, the probability to exit at the end of period $t-1$ (i.e. beginning of period $t$) decreases with $\widetilde{\theta}_{ijkt-2}$ and $a_{ijkt-1}$.

We estimate the following probabilistic model:

$$S_{ijkt} = \alpha \mathrm{AGE}_{ijkt-1} + \beta(a_{ijk,t-1} - \varepsilon^q_{ijk,t-1}) + \gamma \varepsilon^q_{ijkt-1} + \delta(a_{ijk,t-1} - \varepsilon^q_{ijk,t-1}) \times \mathrm{AGE}_{ijkt-1} + \mathbf{FE} + u_{ijkt} > 0$$

Where $S_{ijkt} = 0$ is a dummy that takes the value 1 if firm $i$ exits market $jk$ in year $t$. We expect $\beta$ and $\gamma$ to be negative, and $\delta$ to be positive. **FE** include the two sets of fixed effects $\mathbf{FE}_{ikt}$ and $\mathbf{FE}_{jkt}$, which capture $C^S_{ikt}$ and $C^S_{jkt}$. We estimate this equation using a linear probability model which does not suffer from incidental parameters problems, an issue that might be important here given the two large dimensions of fixed effects we need to include.

The results for prediction 4.a are shown in Table A.11, columns (1) to (3), and are largely in line with the model: conditional on age, the exit probability decreases with the value of demand shocks $\widehat{v}$ and firm's belief (columns (1) to (3)).

Columns (4) and (5) of Table A.11 test for prediction 4.b. We simply add to our baseline specification of column (3) an interaction term between age and demand shock in $t-1$.[10] We indeed find that the coefficient on this interaction term is positive: Young firms

[10] Given our need to control for all $jkt$-determinants here, we use the version of $\widehat{v}_{ijk,t-1}$ computed using $jkt$-specific fixed effects, as in Table 4. This has no importance in columns (1) to (3) as the vector of fixed effects includes $\mathbf{FE}_{jkt}$, but it does in columns (4) and (5) as the coefficient on the interaction between $\widehat{v}_{ijkt-1}$ and age might reflect differences in $\widehat{v}_{ijkt-1}$ along the $jkt$ dimension (as we focus on an interaction term in this case).

react more to a given demand shock than mature exporters on the market. In column (5), a negative demand shock of 10% increases exit probability by 3.3 percentage points for a young firm (2 years after entry), but by only 1.3 percentage points after 7 years.

Table A.10: Prediction 1: controlling for size (robustness 2/2)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | | | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | | |
| Size | Market share (value) | | | | Market share (quantity) | | | |
| Functional form | Linear | | Bins | | Linear | | Bins | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.102^a$ | | $0.095^a$ | | $0.097^a$ | | $0.111^a$ | |
| | (0.002) | | (0.003) | | (0.002) | | (0.003) | |
| $\times$ Age$_{ijkt}$ | $-0.004^a$ | | $-0.003^a$ | | $-0.003^a$ | | $-0.002^a$ | |
| | (0.000) | | (0.000) | | (0.000) | | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | $0.096^a$ | | $0.100^a$ | | $0.092^a$ | | $0.097^a$ |
| | | (0.002) | | (0.003) | | (0.002) | | (0.003) |
| $\times$ Age$_{ijkt} = 3$ | | $0.092^a$ | | $0.098^a$ | | $0.088^a$ | | $0.095^a$ |
| | | (0.002) | | (0.003) | | (0.002) | | (0.003) |
| $\times$ Age$_{ijkt} = 4$ | | $0.086^a$ | | $0.092^a$ | | $0.082^a$ | | $0.090^a$ |
| | | (0.002) | | (0.003) | | (0.002) | | (0.003) |
| $\times$ Age$_{ijkt} = 5$ | | $0.085^a$ | | $0.092^a$ | | $0.082^a$ | | $0.090^a$ |
| | | (0.003) | | (0.003) | | (0.002) | | (0.003) |
| $\times$ Age$_{ijkt} = 6$ | | $0.081^a$ | | $0.087^a$ | | $0.079^a$ | | $0.085^a$ |
| | | (0.003) | | (0.003) | | (0.003) | | (0.003) |
| $\times$ Age$_{ijkt} = 7$ | | $0.078^a$ | | $0.084^a$ | | $0.076^a$ | | $0.082^a$ |
| | | (0.003) | | (0.004) | | (0.003) | | (0.004) |
| $\times$ Age$_{ijkt} = 8$ | | $0.075^a$ | | $0.082^a$ | | $0.074^a$ | | $0.080^a$ |
| | | (0.004) | | (0.005) | | (0.004) | | (0.005) |
| $\times$ Age$_{ijkt} = 9$ | | $0.075^a$ | | $0.084^a$ | | $0.075^a$ | | $0.084^a$ |
| | | (0.005) | | (0.006) | | (0.005) | | (0.006) |
| $\times$ Age$_{ijkt} = 10$ | | $0.072^a$ | | $0.081^a$ | | $0.072^a$ | | $0.079^a$ |
| | | (0.009) | | (0.010) | | (0.009) | | (0.010) |
| $\times$ Size$_{ijk,t-1}$ | $0.011^c$ | $0.011^c$ | | | 0.008 | $0.009^c$ | | |
| | (0.006) | (0.006) | | | (0.005) | (0.005) | | |
| Age$_{ijkt}$ | $-0.040^a$ | | $0.019^a$ | | $-0.039^a$ | | $0.015^a$ | |
| | (0.001) | | (0.002) | | (0.001) | | (0.002) | |
| Size$_{ijkt}$ | $-1.053^a$ | $-1.014^a$ | | | $-0.884^a$ | $-0.849^a$ | | |
| | (0.016) | (0.017) | | | (0.015) | (0.015) | | |
| $\times$ Age$_{ijkt}$ | $0.109^a$ | $0.101^a$ | | | $0.090^a$ | $0.083^a$ | | |
| | (0.003) | (0.003) | | | (0.002) | (0.002) | | |
| Observations | 1870377 | 1870377 | 1501840 | 1501840 | 1870377 | 1870377 | 1501840 | 1501840 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17). Compared to our baseline methodology, in this table we include $jkt$ fixed effects in the estimation of the price residuals $\varepsilon^p_{ijkt}$ used to identify demand shocks. $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). In columns (1)-(4), Size$_{ijkt}$ is proxied by the value sold by firm $i$ on market $jk$ during year $t$ divided by the total value exported by French firms in market $jk$ during year $t$. In columns (1)-(4), Size$_{ijkt}$ is proxied by the quantity sold by firm $i$ on market $jk$ during year $t$ divided by the total quantity exported by French firms in market $jk$ during year $t$. Age dummies included alone in columns (2), (4), (6) and (8) but coefficients not reported. Compared to our baseline estimates, these regressions include additional interaction terms between $a_{ijkt} - \varepsilon^q_{ijk,t}$ and age.

Table A.11: Firm exit

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | $\Pr(S_{ijkt} = 0 \mid S_{ijkt-1} = 1)$ | | | |
| $\text{Age}_{ijk,t-1}$ | -0.024$^a$ | -0.028$^a$ | -0.022$^a$ | -0.022$^a$ |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| $\varepsilon^q_{ijkt-1}$ | -0.043$^a$ | | -0.080$^a$ | -0.097$^a$ |
| | (0.000) | | (0.000) | (0.001) |
| $\times$ $\text{Age}_{ijk,t-1}$ | | | | 0.004$^a$ |
| | | | | (0.000) |
| $(a_{ijk,t-1} - \varepsilon^q_{ijkt-1})$ | | 0.033$^a$ | -0.039$^a$ | -0.044$^a$ |
| | | (0.000) | (0.000) | (0.001) |
| $\times$ $\text{Age}_{ijk,t-1}$ | | | | 0.001$^a$ |
| | | | | (0.000) |
| Observations | 4885284 | 4885284 | 4885284 | 4885284 |

Robust standard errors clustered by firm-product-destination in parentheses. Estimator: LPM. All estimations include $jkt$ and $ikt$ fixed effects. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt-1}$ is the belief of the firm about future demand from equation (14). $\text{Age}_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit).

# G Firm growth

Our first stylized fact shows that the growth rates of quantities decline with age, conditional on size. This decline comes from two different mechanisms in the passive learning model: (i) selection (ii) larger growth rates for younger firms, *unconditional* on survival.

The impact of selection on growth rates is due to the fact that younger firms have greater variance in their growth rates, which comes from their larger updating. Firms that decrease in size are more likely to exit. Hence, the distribution of growth rates is truncated from below. As younger firms may experience more negative growth rates due to the larger variance, this truncation leads to larger growth rates for younger firms, conditional on survival. Note that this mechanism holds only if exit rates are not increasing with age, which is clearly the case in our data (see Figure 1.a in the main text and the results of the previous section on firm survival).

Second, the passive learning model is also consistent with larger growth rates for younger firms, even if we do not condition on firm survival. This unconditional growth is quite limited in our data (see figure 2.c), but is not in contradiction with the model. It should be noted however that this result is driven by the assumption that $exp(\frac{a_{ijkt}}{\sigma_k})$ is log-normally distributed and is thus sensitive to the functional form assumption. In the rest of this section we detail the proof of this result.

**Expected growth rate, conditional on size, *unconditional* on survival.** The expected (quantity) growth rate of firm $i$ at time $t$, conditional on its size, and non conditional on survival, is given by:

$$\frac{\mathbb{E}_{t-1}\left[q^*_{ijkt+1}\right]}{q^*_{ijkt}}$$

where $\mathbb{E}_{t-1}\left[q^*_{ijkt+1}\right]$ is the expected quantity at time $t+1$, conditional on the information available at time $t-1$, i.e. conditional on the information received from $t-1$ signals: $\overline{a}_{ijkt-1}$. In words, this is the expected value of $q^*_{ijkt+1}$, given that the shock in period $t$, $a_{ijkt}$, is not observed yet, and will lead to an updating of firm beliefs between $t$ and $t+1$.

Given the optimal quantity choice (see equation (7)), we get:

$$\begin{aligned}
\frac{\mathbb{E}_{t-1}\left[q^*_{ijkt+1}\right]}{q^*_{ijkt}} &= \frac{\mathbb{E}_{t-1}\left[\left(\frac{\sigma_k}{\sigma_k-1}\frac{w_{it+1}}{\varphi_{ikt+1}}\right)^{-\sigma_k}\left(\frac{\mu_k Y_{jt+1}}{P^{1-\sigma_k}_{jkt+1}}\right)\mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]^{\sigma_k}\right]}{\left(\frac{\sigma_k}{\sigma_k-1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k}\left(\frac{\mu_k Y_{jt}}{P^{1-\sigma_k}_{jkt}}\right)\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]^{\sigma_k}} \\
&= \frac{\left(\frac{\sigma_k}{\sigma_k-1}\frac{w_{it+1}}{\varphi_{ikt+1}}\right)^{-\sigma_k}\left(\frac{\mu_k Y_{jt+1}}{P^{1-\sigma_k}_{jkt+1}}\right)\mathbb{E}_{t-1}\left[\mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]^{\sigma_k}\right]}{\left(\frac{\sigma_k}{\sigma_k-1}\frac{w_{it}}{\varphi_{ikt}}\right)^{-\sigma_k}\left(\frac{\mu_k Y_{jt}}{P^{1-\sigma_k}_{jkt}}\right)\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]^{\sigma_k}}
\end{aligned}$$

As we work with purged quantities, we label $\mathbb{E}_{t-1}\left[g_q\right]$ the expected growth rate of purged quantities:

$$\mathbb{E}_{t-1}\left[g_q\right]=\frac{\mathbb{E}_{t-1}\left[\mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]^{\sigma_k}\right]}{\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]^{\sigma_k}}$$

As $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]=e^{\frac{1}{\sigma_k}\left(\widetilde{\theta}_{ijkt-1}+\frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}$ (see appendix main text), we get:

$$\mathbb{E}_{t-1}\left[g_q\right]=\frac{\mathbb{E}_{t-1}\left[e^{\left(\widetilde{\theta}_{ijkt}+\frac{\widetilde{\sigma}_t^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}\right]}{e^{\left(\widetilde{\theta}_{ijkt-1}+\frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}}$$

Note that, from $t-1$ perspective, $\widetilde{\theta}_{ijkt}$ is a random variable as $a_{ijkt}$ is not observed. We may rewrite $\mathbb{E}_{t-1}\left[g_q\right]$ as:

$$\mathbb{E}_{t-1}\left[g_q\right]=\frac{\mathbb{E}_{t-1}\left[e^{\widetilde{\theta}_{ijkt}}\right]e^{\left(\frac{\widetilde{\sigma}_t^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}}{e^{\left(\widetilde{\theta}_{ijkt-1}+\frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}}$$

We next rewrite $\widetilde{\theta}_{ijkt}$ to explicit $a_{ijkt}$:

$$\begin{aligned}\widetilde{\theta}_{ijkt} &= \theta_0\frac{\frac{1}{\sigma_0^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}+\frac{1}{t}\left(\left(t-1\right)\overline{a}_{ijkt-1}+a_{ijkt}\right)\frac{\frac{t}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\\ &= \theta_0\frac{\frac{1}{\sigma_0^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}+\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\left(t-1\right)\overline{a}_{ijkt-1}+\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}a_{ijkt}\end{aligned}$$

$\widetilde{\theta}_{ijkt}$ being linear in $a_{ijkt}$, it is also normally distributed. From $t-1$ perspective we get $\mathbb{E}\left[\widetilde{\theta}_{ijkt}\mid\widetilde{\theta}_{ijkt-1}\right]=\widetilde{\theta}_{ijkt-1}$.

Second, remind that $\mathbb{V}(a_{ijkt})=\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2$, so $\mathbb{V}\left(\widetilde{\theta}_{ijkt}\right)=\left(\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right)^2\left(\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2\right)$. Since $\widetilde{\theta}_{ijkt}$ is normally distributed, $e^{\widetilde{\theta}_{ijkt}}$ is lognormally distributed. We thus obtain:

$$\begin{aligned}\mathbb{E}_{t-1}\left[g_q\right] &= \frac{e^{\left(\widetilde{\theta}_{ijkt-1}+\frac{1}{2}\left(\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right)^2\left(\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2\right)\right)}e^{\left(\frac{\widetilde{\sigma}_t^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}}{e^{\left(\widetilde{\theta}_{ijkt-1}+\frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}}\\ &= e^{\frac{1}{2}\left(\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right)^2\left(\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2\right)+\frac{\widetilde{\sigma}_t^2-\widetilde{\sigma}_{t-1}^2}{2\sigma_k}}\end{aligned}$$

As we work with log (purged) quantities, let's take the log:

$$\ln \mathbb{E}_{t-1}[g_q] = \frac{1}{2}\left(\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right)^2 \left(\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2\right) - \frac{\widetilde{\sigma}_{t-1}^2-\widetilde{\sigma}_t^2}{2\sigma_k}$$

Given the definitions of $\widetilde{\sigma}_t^2$ and $\widetilde{\sigma}_{t-1}^2$ (see equation (4)), we get:

$$\begin{aligned}
\ln \mathbb{E}_{t-1}[g_q] &= \frac{1}{2}\left(\frac{\frac{1}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right)^2 \left(\frac{1}{\frac{1}{\sigma_0^2}+\frac{t-1}{\sigma_\epsilon^2}}+\sigma_\varepsilon^2\right) - \frac{1}{2\sigma_k}\left(\frac{1}{\frac{1}{\sigma_0^2}+\frac{t-1}{\sigma_\epsilon^2}} - \frac{1}{\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}}\right) \\
\ln \mathbb{E}_{t-1}[g_q] &= \left(\frac{1}{2}-\frac{1}{2\sigma_k}\right)\frac{\frac{1}{\sigma_\epsilon^2}}{\left(\frac{1}{\sigma_0^2}+\frac{t}{\sigma_\epsilon^2}\right)\left(\frac{1}{\sigma_0^2}+\frac{t-1}{\sigma_\epsilon^2}\right)}
\end{aligned}$$

Note that $\frac{1}{2}-\frac{1}{2\sigma_k} > 0$, so $\ln \mathbb{E}_{t-1}[g_q]$ is always positive. Moreover, $t$ appears in the denominator only, so this expression is strictly decreasing with $t$: Expected growth rates decline with firm age in market $jk$.

The source of this result comes from the functional form assumption: the profit function depends on the *exponential* of the demand shock $a_{ijkt}$. Given that $a_{ijkt}$ is normally distributed, $exp(\frac{a_{ijkt}}{\sigma_k})$ is log-normally distributed, its expectation thus depends on its variance. Without this effect, expected growth rate (non conditional on survival) of purged quantities should always be 0, no matter firm age. Second, note that expectation is taken over $exp(\frac{a_{ijkt}}{\sigma_k})$ and its variance is reduced by $\sigma_k$. But expectation is also taken over $exp(\widetilde{\theta}_{ijkt})$, which does not depend on $\sigma_k$. This is generating the result.

# H Belief updating and age: endogenous selection

This section presents the detailed results discussed in section 5.2 of the main text on survival and selection bias and provides additional evidence based on alternative nonlinear estimators. All these specifications draw on the predictions of our model regarding firms' exit decision detailed in section F and Table A.11. In particular, exit probabilities depend on $a_{ijkt}$, $\varepsilon^q_{ijkt}$, $\text{Age}_{ijkt}$ and fixed effects in the $ikt$ and $jkt$ dimensions and can be estimated using a linear probability model.

We start by documenting whether the firms' updating process identified in Table 2 varies depending on their survival probability. This is application of the "identification-at-infinity" method (Chamberlain, 1986; Mulligan and Rubinstein, 2008). We expect the potential selection bias related to endogenous exit decisions to be lower on sub-samples of firms, selected on observable characteristics, most likely to survive. We first estimate equation (F) and compute the predicted probability of exit by firm×market×year. Equation (19) is then estimated on four sub-samples including respectively firms above the 20th, 40th, 60th and 80th percentiles of survival probability (i.e. below the 80th, 60th, 40th and 20th of exit probability). Table A.12 presents the results when firms are allocated in quintiles depending on their raw probability of exit. Alternatively, in Table A.13 we allocate firms in quintile of exit probability by firm-market size. In both specifications, both the coefficient on $(a_{ijkt} - \varepsilon^q_{ijkt-1})$ and its interaction with age are stable across sub-samples of firms and the results on the sample of firms most likely to survive (column (5)) is very close to the full sample (column(1)).

Table A.12: Demand shocks and beliefs updating, by exit probability

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. var. | | | $\Delta\varepsilon^q_{ijk,t+1}$ | | |
| Exit prob. | All | Bottom 80% | Bottom 60% | Bottom 40% | Bottom 20% |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.075^a$ | $0.075^a$ | $0.074^a$ | $0.071^a$ | $0.070^a$ |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.003) |
| $\times$ $\text{Age}_{ijkt}$ | $-0.003^a$ | $-0.003^a$ | $-0.003^a$ | $-0.003^a$ | $-0.003^a$ |
| | (0.000) | (0.000) | (0.001) | (0.001) | (0.001) |
| $\text{Age}_{ijkt}$ | $-0.038^a$ | $-0.047^a$ | $-0.054^a$ | $-0.057^a$ | $-0.069^a$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Observations | 1501766 | 1154290 | 839245 | 531182 | 248194 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). $\text{Age}_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Predicted exit probabilities are obtained by from the estimation of Table A.11, column (4).

Tables A.14 and A.15 directly account for the potential selection bias by including a correction term in our estimation of equation (19). The high dimensionally of the fixed effects implied by prediction 4 in Section F (see equation (F)) for the selection equation

Table A.13: Demand shocks and beliefs updating, by exit probability (robustness)

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Dep. var. | | | $\Delta\varepsilon^q_{ijk,t+1}$ | | |
| Exit prob. | All | Bottom 80% | Bottom 60% | Bottom 40% | Bottom 20% |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.090$^a$ | 0.093$^a$ | 0.092$^a$ | 0.093$^a$ | 0.091$^a$ |
| | (0.002) | (0.003) | (0.003) | (0.003) | (0.004) |
| $\times$ Age$_{ijkt}$ | -0.004$^a$ | -0.005$^a$ | -0.005$^a$ | -0.005$^a$ | -0.005$^a$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Age$_{ijkt}$ | -0.041$^a$ | -0.047$^a$ | -0.049$^a$ | -0.053$^a$ | -0.060$^a$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Observations | 753646 | 552923 | 392223 | 248980 | 120723 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Predicted exit probabilities are obtained by from the estimation of Table A.11, column (4). Samples of exit probabilities are constructed by quintiles of firm size.

prevents us from using a probit or other maximum likelihood estimator and implementing the standard Heckman procedure. In his review of the literature on endogenous sample selection Vella (1998) however proposes a number of alternative procedures based on linear (Olsen, 1980), semi-parametric (Cosslett, 1991), or polynominal estimations of correction terms. We report results of these three alternative procedures as well as a standard Heckman estimator ignoring the $ikt$ and $jkt$ fixed effects in the selection equation in Tables A.14 and A.15. Vella (1998) shows that the assumption of normality in the Heckman procedure can be relaxed to allow for consistent two step estimation using methods based on alternative distributional assumptions than probit in the selection equation. In particular, Vella (1998) argues that Olsen's procedure generally produces results similar to a Heckman two-step procedure. Instead of assuming Normality of the selection equation's error term, Olsen assumes that it follows a uniform distribution. Exclusion of at least one variable from the first step is required in Olsen, not in Heckman, as the Heckman estimator includes as a correction term the Inverse Mills ratio which maps the prediction of the selection equation into a correction term in a nonlinear fashion (hence the correction term is never perfectly collinear with the second-step regressors). The $ikt$ and $jkt$ fixed effects included in equation (F) can serve as exclusion variables in a linear procedure. The complete set of results is reported in columns (1)-(4) of table A.14. Alternatively Cosslett (1991) proposes a semi-parametric estimator in which the selection correction is approximated through indicator variables. In columns (5)-(8) of Table A.14, we use 100 bins corresponding to each centile of the predicted exit probabilities as correction terms. Finally, in columns (1)-(4) of Table A.15 the predicted probability of exit is introduced directly when estimating equation (19) in the form of a 10 degree polynomial. The last three columns of Table A.15 report the results of a standard two-step Heckman procedure excluding the $ikt$ and $jkt$ fixed effects in the probit estimation of the selection equation and using the nonlinearity of the Inverse Mills Ratio to identify its coefficient. Overall, all these alternative treatments of the sample selection bias leave our coefficients of interest largely unaffected, suggesting that endogenous selection is not driving our results.

Table A.14: Demand shocks and beliefs updating: controlling for endogenous exit

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijkt+1}$ | | | | $\Delta\varepsilon^q_{ijkt+1}$ | | |
| Selection correction | | Linear | | | | Semi-parametric | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.065$^a$ | 0.075$^a$ | 0.075$^a$ | | 0.065$^a$ | 0.075$^a$ | 0.075$^a$ | |
| | (0.001) | (0.002) | (0.001) | | (0.001) | (0.002) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | -0.003$^a$ | -0.003$^a$ | | | -0.003$^a$ | -0.003$^a$ | |
| | | (0.000) | (0.000) | | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | 0.069$^a$ | | | | 0.069$^a$ |
| | | | | (0.001) | | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | | 0.064$^a$ | | | | 0.064$^a$ |
| | | | | (0.001) | | | | (0.001) |
| $\times$ Age$_{ijkt} = 4$ | | | | 0.060$^a$ | | | | 0.060$^a$ |
| | | | | (0.002) | | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | | 0.056$^a$ | | | | 0.056$^a$ |
| | | | | (0.002) | | | | (0.002) |
| $\times$ Age$_{ijkt} = 6$ | | | | 0.059$^a$ | | | | 0.059$^a$ |
| | | | | (0.002) | | | | (0.002) |
| $\times$ Age$_{ijkt} = 7$ | | | | 0.055$^a$ | | | | 0.055$^a$ |
| | | | | (0.003) | | | | (0.003) |
| $\times$ Age$_{ijkt} = 8$ | | | | 0.051$^a$ | | | | 0.051$^a$ |
| | | | | (0.004) | | | | (0.004) |
| $\times$ Age$_{ijkt} = 9$ | | | | 0.054$^a$ | | | | 0.054$^a$ |
| | | | | (0.007) | | | | (0.007) |
| $\widehat{\Pr(\text{exit}_{ijkt})}$ | -0.409$^a$ | -0.409$^a$ | -0.409$^a$ | -0.417$^a$ | | | | |
| | (0.005) | (0.005) | (0.003) | (0.005) | | | | |
| Age$_{ijkt}$ | -0.054$^a$ | -0.054$^a$ | -0.054$^a$ | | -0.057$^a$ | -0.057$^a$ | -0.057$^a$ | |
| | (0.001) | (0.001) | (0.000) | | (0.001) | (0.001) | (0.001) | |
| Observations | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 | 1501766 |

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3) and (7)). $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (4) and (8) but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). In columns (1)-(4), predicted exit probabilities are obtained from the estimation of Table A.11, column (4) and introduced directly in equation (19). In columns (5)-(8), they are introduced semi-parametrically in the second step, i.e. we included 100 bins corresponding to each percentile of the variable.

Table A.15: Demand shocks and beliefs updating: controlling for endogenous exit (robustness)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | | $\Delta\varepsilon^q_{ijk,t+1}$ | |
| Selection correction | | Polynomial | | | | Heckman | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.065^a$ | $0.075^a$ | $0.075^a$ | | $0.078^a$ | $0.086^a$ | |
| | (0.001) | (0.002) | (0.001) | | (0.002) | (0.004) | |
| $\times$ Age$_{ijkt}$ | | $-0.003^a$ | $-0.003^a$ | | | $-0.003^a$ | |
| | | (0.000) | (0.000) | | | (0.001) | |
| $\times$ Age$_{ijkt} = 2$ | | | | $0.069^a$ | | | $0.080^a$ |
| | | | | (0.001) | | | (0.002) |
| $\times$ Age$_{ijkt} = 3$ | | | | $0.064^a$ | | | $0.078^a$ |
| | | | | (0.001) | | | (0.003) |
| $\times$ Age$_{ijkt} = 4$ | | | | $0.060^a$ | | | $0.076^a$ |
| | | | | (0.002) | | | (0.004) |
| $\times$ Age$_{ijkt} = 5$ | | | | $0.056^a$ | | | $0.062^a$ |
| | | | | (0.002) | | | (0.005) |
| $\times$ Age$_{ijkt} = 6$ | | | | $0.059^a$ | | | $0.072^a$ |
| | | | | (0.002) | | | (0.006) |
| $\times$ Age$_{ijkt} = 7$ | | | | $0.055^a$ | | | $0.070^a$ |
| | | | | (0.003) | | | (0.008) |
| $\times$ Age$_{ijkt} = 8$ | | | | $0.051^a$ | | | $0.065^a$ |
| | | | | (0.004) | | | (0.011) |
| $\times$ Age$_{ijkt} = 9$ | | | | $0.055^a$ | | | $0.058^a$ |
| | | | | (0.007) | | | (0.016) |
| Age$_{ijkt}$ | $-0.057^a$ | $-0.057^a$ | $-0.057^a$ | | $-0.594^a$ | $-0.595^a$ | |
| | (0.001) | (0.001) | (0.001) | | (0.006) | (0.006) | |
| $\lambda$ | | | | | $4.922^a$ | $4.922^a$ | $4.918^a$ |
| | | | | | (0.039) | (0.039) | (0.039) |
| Observations | 1501766 | 1501766 | 1501766 | 1501766 | 1550474 | 1550474 | 1550474 |

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3). $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (4) but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). In columns (1)-(4), predicted exit probabilities are obtained by from the estimation of Table A.11, column (4) and introduced directly in equation (19) in the form of a 10-degree polynomial. In columns (5)-(6), we use a Heckman estimator which estimates a probit in the first step (omitting the $ikt$ and $jkt$fixed effects) and introduces the inverse mills ratio ($\lambda$) in the second step.

# I Belief updating and age: additional robustness

## I.1 Extra-EU results

In Table A.16, we restrict our sample to extra-EU destination countries to check that the different declaration thresholds applying to intra-EU expeditions and extra-EU exports (as explained in footnote 11 of the main text) do not affect our results. Focusing on extra-EU countries reduces the number of observations by 40%, but does not alter our coefficients of interest compared to the baseline results in Table 2.

Table A.16: Prediction 1: demand shocks and beliefs updating (extra EU)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^{q}_{ijkt+1}$ | | |
| $a_{ijkt} - \varepsilon^{q}_{ijkt}$ | $0.061^{a}$ | $0.072^{a}$ | $0.072^{a}$ | |
| | (0.001) | (0.002) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | $-0.003^{a}$ | $-0.003^{a}$ | |
| | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | $0.067^{a}$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | | $0.062^{a}$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 4$ | | | | $0.056^{a}$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | | $0.055^{a}$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 6$ | | | | $0.055^{a}$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 7$ | | | | $0.050^{a}$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 8$ | | | | $0.048^{a}$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 9$ | | | | $0.048^{a}$ |
| | | | | (0.004) |
| $\times$ Age$_{ijkt} = 10$ | | | | $0.045^{a}$ |
| | | | | (0.007) |
| Age$_{ijkt}$ | $-0.033^{a}$ | $-0.033^{a}$ | $-0.033^{a}$ | |
| | (0.001) | (0.001) | (0.001) | |
| Observations | 1109761 | 1109761 | 1109761 | 1109761 |

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3)). [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Sample extra EU destinations only. Age dummies included alone in columns (4) but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^{q}_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit).

## I.2 Alternative age definitions

So far we have treated each entry into a market as a new one: age was reset to zero in case of exit. We now check the sensitivity of our results to alternative definitions of age. We define two alternative measures of age. We first assumes that information on local demand is not forgotten by the firm when it does not serve a product-destination only one year and accordingly reset age to zero only after two consecutive years of exit. In the second definition, we assume that firms keep entirely their knowledge about local demand when they exit, regardless of the number of exit years; this third age variable is simply the number of exporting years since the first entry of the firm.

Table A.17 shows that the results using these alternative definitions are qualitatively similar to our baseline estimates. However, the effects of age – its direct effect and its effect on firms' reactions to demand signals – are slightly lower than in our baseline table.

Table A.17: Prediction 1: alternative age definitions

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijk,t+1}$ | | | $\Delta\varepsilon^q_{ijk,t+1}$ | |
| Age definition | | # years since last entry (reset after 2 years exit) | | | # years exporting since first entry | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.064^a$ | $0.073^a$ | | $0.064^a$ | $0.072^a$ | |
| | (0.001) | (0.001) | | (0.001) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | $-0.002^a$ | | | $-0.002^a$ | |
| | | (0.000) | | | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | $0.068^a$ | | | $0.068^a$ |
| | | | (0.001) | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | $0.065^a$ | | | $0.065^a$ |
| | | | (0.001) | | | (0.001) |
| $\times$ Age$_{ijkt} = 4$ | | | $0.061^a$ | | | $0.063^a$ |
| | | | (0.002) | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | $0.059^a$ | | | $0.061^a$ |
| | | | (0.002) | | | (0.002) |
| $\times$ Age$_{ijkt} = 6$ | | | $0.060^a$ | | | $0.061^a$ |
| | | | (0.002) | | | (0.002) |
| $\times$ Age$_{ijkt} = 7$ | | | $0.057^a$ | | | $0.058^a$ |
| | | | (0.002) | | | (0.003) |
| $\times$ Age$_{ijkt} = 8$ | | | $0.056^a$ | | | $0.057^a$ |
| | | | (0.003) | | | (0.003) |
| $\times$ Age$_{ijkt} = 9$ | | | $0.054^a$ | | | $0.054^a$ |
| | | | (0.004) | | | (0.004) |
| $\times$ Age$_{ijkt} = 10$ | | | $0.047^a$ | | | $0.047^a$ |
| | | | (0.007) | | | (0.007) |
| Age$_{ijkt}$ | $-0.030^a$ | $-0.030^a$ | | $-0.029^a$ | $-0.029^a$ | |
| | (0.001) | (0.001) | | (0.001) | (0.001) | |
| Observations | 1854141 | 1854141 | 1854141 | 1854141 | 1854141 | 1854141 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. Age dummies included alone in columns (3) and (6) but coefficients not reported. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14).

## I.3 Reconstructed years

The usual aggregation of export sales by calendar year is likely to bias downward the average sales of new exporters because some enter a given market late in the year (Berthou and Vicard, 2015). The average growth rate of quantities would in turn be inflated between the first, potentially incomplete, and the second (full) year of export. When estimating equation (19), the dummy for age two picks the average bias related to the incompleteness of the first year of export. In Table A.18, we go one step further and address this issue directly by performing our estimation strategy on reconstructed years beginning the month of first entry at the firm-product-destination level. The results shows that both the average updating of the firms' beliefs and its interaction with age are quantitatively similar to our baseline in Table 2.

The drawback of using such reconstructed yearly data is the inability to control consistently for market-year fixed effects in equations (14) and (15): introducing market$\times$year fixed effects specific by firms' month of entry reduces dramatically the number of observations for which we can identify beliefs and demand shocks. We therefore stick to the usual calendar year dataset in the main text.

Table A.18: Prediction 1: reconstructed years

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijkt+1}$ | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.067^a$ | $0.078^a$ | $0.078^a$ | |
| | (0.001) | (0.001) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | $-0.003^a$ | $-0.003^a$ | |
| | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | $0.072^a$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | | $0.066^a$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 4$ | | | | $0.066^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | | $0.058^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 6$ | | | | $0.061^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 7$ | | | | $0.054^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 8$ | | | | $0.056^a$ |
| | | | | (0.004) |
| $\times$ Age$_{ijkt} = 9$ | | | | $0.056^a$ |
| | | | | (0.005) |
| Age$_{ijkt}$ | $-0.010^a$ | $-0.010^a$ | $-0.010^a$ | |
| | (0.001) | (0.001) | (0.001) | |
| Observations | 1495774 | 1495774 | 1495774 | 1495774 |

Robust standard errors clustered by firm in parentheses (bootstrapped in column (3)). $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Age dummies included alone in column (4) but coefficients not reported. In this Table years are reconstructed beginning the month of first entry at the firm-product-destination level

### I.4 $\sigma_k$ computed at 4-digit (HS4) level

In Table A.19, we use demand shocks obtained by estimating equation (17) by 4-digit product instead of 6-digit product of the Harmonized System classification in order to allow for a larger number of observations when estimating $\sigma_k$. As expected, our estimates of $\sigma_k$ are slightly lower in this case than in the baseline 6-digit case (a median of 4.98 and a mean of 5.83). The results shown in Table A.19 are close to our baseline results.

Table A.19: Prediction 1: $\sigma_k$ computed at 4-digit (HS4) level

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | $\Delta\varepsilon^q_{ijkt+1}$ | | | |
| Robustness | $\sigma_k$ at HS4 level | | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | 0.073$^a$ | 0.084$^a$ | 0.084$^a$ | |
| | (0.001) | (0.001) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | -0.004$^a$ | -0.004$^a$ | |
| | | (0.000) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | 0.078$^a$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 3$ | | | | 0.072$^a$ |
| | | | | (0.001) |
| $\times$ Age$_{ijkt} = 4$ | | | | 0.067$^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | | 0.065$^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 6$ | | | | 0.064$^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 7$ | | | | 0.059$^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 8$ | | | | 0.060$^a$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 9$ | | | | 0.059$^a$ |
| | | | | (0.004) |
| $\times$ Age$_{ijkt} = 10$ | | | | 0.057$^a$ |
| | | | | (0.006) |
| Age$_{ijkt}$ | -0.032$^a$ | -0.033$^a$ | -0.033$^a$ | |
| | (0.001) | (0.001) | (0.000) | |
| Observations | 1877732 | 1877732 | 1877732 | 1877732 |

Robust standard errors clustered by firm in parentheses (bootstrapped in column (3)). $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17), estimated by HS4 products instead of HS6; $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand from equation (14). Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Age dummies included alone in column (4) but coefficients not reported.

### I.5 Controlling for $ijt$ fixed effects

The theoretical framework developed in section 3 assumes no informational spillovers, considering $\theta_{ijk0}$ as exogenous. While our identification strategy controls *de facto* for several sources of informational spillovers – the firm×product×year fixed effects included in equations (14) and (15) account for past experience gathered from selling the same product on the domestic or other markets –, it does not take into those from selling other products in the same destination. To this end, we extend our identification strategy by including $ijt$ fixed effects in equations (12) and (13) and re-estimate $a_{ijkt}$ from these alternative $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ to test prediction 1. Table A.20 reports the results and show that our conclusion remain robust qualitatively as well as quantitatively. This lends support to our assumption that information is indeed mostly product-market specific. If shocks and beliefs were correlated across products within destinations, the firms' response to a demand shock would partly reflect its belief updating behavior on other products: including $ijt$ fixed effects should dampen the extent of estimated belief updating.

Table A.20: Prediction 1: controlling for $ijt$ fixed effects

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dep. var. | | $\Delta\varepsilon^q_{ijkt+1}$ | | |
| $a_{ijkt} - \varepsilon^q_{ijkt}$ | $0.091^a$ | $0.102^a$ | $0.102^a$ | |
| | (0.001) | (0.002) | (0.001) | |
| $\times$ Age$_{ijkt}$ | | $-0.003^a$ | $-0.003^a$ | |
| | | (0.001) | (0.000) | |
| $\times$ Age$_{ijkt} = 2$ | | | | $0.096^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 3$ | | | | $0.092^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 4$ | | | | $0.087^a$ |
| | | | | (0.002) |
| $\times$ Age$_{ijkt} = 5$ | | | | $0.085^a$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 6$ | | | | $0.082^a$ |
| | | | | (0.003) |
| $\times$ Age$_{ijkt} = 7$ | | | | $0.076^a$ |
| | | | | (0.004) |
| $\times$ Age$_{ijkt} = 8$ | | | | $0.078^a$ |
| | | | | (0.005) |
| $\times$ Age$_{ijkt} = 9$ | | | | $0.079^a$ |
| | | | | (0.005) |
| Age$_{ijkt}$ | $-0.013^a$ | $-0.013^a$ | $-0.013^a$ | |
| | (0.000) | (0.000) | (0.001) | |
| Observations | 1217810 | 1217810 | 1217810 | 1217810 |

Robust standard errors clustered by firm in parentheses (bootstrapped in column (3)). $^c$ significant at 10%; $^b$ significant at 5%. $^a$ significant at 1%. $a_{ijkt}$ is our estimate of the demand shock from equation (17); $\varepsilon^q_{ijkt}$ is the belief of the firm about future demand. $\varepsilon^q_{ijkt-1}$ and $\varepsilon^p_{ijkt-1}$ are respectively estimated from equation (14) and equation (15) including additionally fixed effects in the $ijt$ dimension. Age$_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). Age dummies included alone in column (4) but coefficients not reported.

# J Test of stationary demand

In the learning model, firms learn about an idiosyncratic demand parameter which is assumed to be constant over time. The initial size of a firm (i.e. its initial belief) should be a useful predictor of its beliefs and sales throughout its life, even controlling for past beliefs. In other words, the evolution of firms' beliefs should not be Markov. Such a prediction would not arise in models with "active learning" where firms invest to increase their profitability, possibly through demand accumulation. To discriminate between these two classes of models, Pakes and Ericson (1998) (see also Abbring and Campbell, 2005 for an application) propose to regress current firms beliefs on their immediate past beliefs and their initial prior beliefs. In Table A.21, we regress the beliefs of the firms after $x$ years, $x = 3, ..., 8$, on their belief at the time of entry, controlling for the immediate lag of the belief. We restrict our sample to firms present at least 8 years to avoid composition effects.[11] Two results are worth mentioning. First, initial beliefs have a positive and significant effect on future beliefs, and this effect remains highly significant even 8 years after entry. Second, the immediate lag of the belief becomes a better predictor of the current belief as the firm gets older, suggesting that firms indeed converge to their demand parameter. Both results are consistent with our assumption on $\overline{a}_{ijk}$. Note that these results are not sensitive to the number of lags used: Table A.22 focuses on firms aged 6 to 8 years for which we can include up to four lags of the belief (we find a similar pattern for firms aged 5 to 8 years for which we can include up to 3 lags). We find that the initial belief remains a significant predictor of current belief after 6, 7 or 8 years when increasing the number of lags of beliefs included as explanatory variables.

[11] Similar results are obtained when restricting the sample to firms present $j$ years, $j = 5, ..., 9$.

Table A.21: Passive versus active learning

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Dep. var. | $\varepsilon^q_{ijkt}$ | | | | | |
| Age definition | # years since last entry (reset after 1 year of exit) | | | | | |
| Age | 3 | 4 | 5 | 6 | 7 | 8 |
| $\varepsilon^q_{ijkt-1}$ | $0.541^a$ | $0.587^a$ | $0.632^a$ | $0.645^a$ | $0.659^a$ | $0.668^a$ |
| | (0.008) | (0.007) | (0.007) | (0.006) | (0.007) | (0.007) |
| $\varepsilon^q_{ijk0}$ | $0.144^a$ | $0.135^a$ | $0.097^a$ | $0.091^a$ | $0.079^a$ | $0.075^a$ |
| | (0.006) | (0.006) | (0.005) | (0.005) | (0.005) | (0.005) |
| Observations | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $\varepsilon^q_{ijkt-1}$ and $\varepsilon^q_{ijk0}$ are respectively the beliefs of the firm in market $jk$ in period $t-1$ and in the first period. Beliefs given by equation (14). Sample of firms-markets present at least 8 years.

Table A.22: Passive versus active learning: robustness

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\varepsilon^q_{ijkt}$ | | | | | | | | | | | |
| Age definition | # years since last entry (reset after 1 year of exit) | | | | | | | | | | | |
| Age | 6 | | | | 7 | | | | 8 | | | |
| $\varepsilon^q_{ijk0}$ | $0.091^a$ | $0.050^a$ | $0.031^a$ | $0.020^a$ | $0.079^a$ | $0.045^a$ | $0.029^a$ | $0.019^a$ | $0.075^a$ | $0.046^a$ | $0.032^a$ | $0.023^a$ |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.004) | (0.005) | (0.005) |
| $\varepsilon^q_{ijkt-1}$ | $0.645^a$ | $0.521^a$ | $0.504^a$ | $0.503^a$ | $0.659^a$ | $0.520^a$ | $0.503^a$ | $0.498^a$ | $0.668^a$ | $0.530^a$ | $0.509^a$ | $0.505^a$ |
| | (0.006) | (0.007) | (0.008) | (0.008) | (0.007) | (0.007) | (0.008) | (0.008) | (0.007) | (0.008) | (0.009) | (0.009) |
| $\varepsilon^q_{ijkt-2}$ | | $0.210^a$ | $0.172^a$ | $0.168^a$ | | $0.223^a$ | $0.181^a$ | $0.174^a$ | | $0.219^a$ | $0.171^a$ | $0.163^a$ |
| | | (0.007) | (0.008) | (0.008) | | (0.007) | (0.008) | (0.008) | | (0.007) | (0.008) | (0.008) |
| $\varepsilon^q_{ijkt-3}$ | | | $0.089^a$ | $0.076^a$ | | | $0.090^a$ | $0.069^a$ | | | $0.099^a$ | $0.075^a$ |
| | | | (0.006) | (0.007) | | | (0.007) | (0.008) | | | (0.007) | (0.008) |
| $\varepsilon^q_{ijkt-4}$ | | | | $0.034^a$ | | | | $0.050^a$ | | | | $0.054^a$ |
| | | | | (0.007) | | | | (0.006) | | | | (0.007) |
| Observations | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 | 41034 |

Robust standard errors clustered by firm in parentheses. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $\varepsilon^q_{ijkt-1}$ and $\varepsilon^q_{ijk0}$ are respectively the beliefs of the firm in market $jk$ in period $t-1$ and in the first period. Beliefs given by equation (14).

# K Profiles of prices and quantities

Table A.23 provides the full set of results used in figure 2 of the main text. In columns (9) and (10) we additionally include firms' size as an explaining variable when regressing $\varepsilon^p_{ijkt}$ on age, to control for the fact that size would affect firms' pricing decisions in a non-CES framework. It confirms that $\varepsilon^p_{ijkt}$ decreases with age (column (9)) but not when we account for composition effects through the inclusion of firm×market fixed effects (column (10)). Table A.24 reproduces estimations in Table A.23 on our dataset with reconstructed years. As expected, reconstructing years from the month of first entry by firm×product×destination dampens the initial increase in quantities sold between the first and second year but leaves unchanged the quantity profile thereafter.

Table A.23: Dynamics of quantity and prices

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| Dep. var. | | $\varepsilon^q_{ijkt}$ | | | | $\varepsilon^p_{ijkt}$ | | | | |
| Sample | | | Surv. | | | | Surv. | | | |
| $\text{Age}_{ijkt}$ | 0.095$^a$ | | | | -0.012$^a$ | | | | | |
| | (0.001) | | | | (0.001) | | | | | |
| $\text{Age}_{ijkt} = 2$ | | 0.192$^a$ | 0.317$^a$ | 0.150$^a$ | | -0.027$^a$ | -0.038$^a$ | -0.008$^a$ | | |
| | | (0.002) | (0.015) | (0.003) | | (0.001) | (0.014) | (0.001) | | |
| $\text{Age}_{ijkt} = 3$ | | 0.297$^a$ | 0.447$^a$ | 0.193$^a$ | | -0.039$^a$ | -0.047$^a$ | -0.011$^a$ | -0.007$^a$ | -0.002 |
| | | (0.003) | (0.018) | (0.004) | | (0.002) | (0.014) | (0.002) | (0.001) | (0.002) |
| $\text{Age}_{ijkt} = 4$ | | 0.378$^a$ | 0.511$^a$ | 0.212$^a$ | | -0.050$^a$ | -0.042$^a$ | -0.014$^a$ | -0.015$^a$ | -0.003 |
| | | (0.004) | (0.019) | (0.005) | | (0.002) | (0.014) | (0.002) | (0.002) | (0.002) |
| $\text{Age}_{ijkt} = 5$ | | 0.433$^a$ | 0.563$^a$ | 0.219$^a$ | | -0.053$^a$ | -0.047$^a$ | -0.014$^a$ | -0.018$^a$ | -0.002 |
| | | (0.005) | (0.021) | (0.007) | | (0.003) | (0.015) | (0.003) | (0.002) | (0.003) |
| $\text{Age}_{ijkt} = 6$ | | 0.477$^a$ | 0.594$^a$ | 0.219$^a$ | | -0.058$^a$ | -0.054$^a$ | -0.014$^a$ | -0.021$^a$ | -0.003 |
| | | (0.006) | (0.023) | (0.009) | | (0.003) | (0.015) | (0.004) | (0.003) | (0.003) |
| $\text{Age}_{ijkt} = 7$ | | 0.526$^a$ | 0.595$^a$ | 0.221$^a$ | | -0.065$^a$ | -0.050$^a$ | -0.018$^a$ | -0.028$^a$ | -0.006 |
| | | (0.008) | (0.022) | (0.011) | | (0.004) | (0.015) | (0.005) | (0.004) | (0.004) |
| $\text{Age}_{ijkt} = 8$ | | 0.556$^a$ | 0.571$^a$ | 0.219$^a$ | | -0.064$^a$ | -0.049$^a$ | -0.017$^b$ | -0.027$^a$ | -0.006 |
| | | (0.010) | (0.022) | (0.013) | | (0.006) | (0.015) | (0.007) | (0.006) | (0.007) |
| $\text{Age}_{ijkt} = 9$ | | 0.593$^a$ | 0.563$^a$ | 0.221$^a$ | | -0.077$^a$ | -0.043$^a$ | -0.024 | -0.040$^a$ | -0.012 |
| | | (0.014) | (0.022) | (0.018) | | (0.013) | (0.015) | (0.016) | (0.013) | (0.015) |
| $\text{Age}_{ijkt} = 10$ | | 0.614$^a$ | 0.496$^a$ | 0.210$^a$ | | -0.070$^a$ | -0.047$^a$ | -0.016$^c$ | -0.031$^a$ | -0.004 |
| | | (0.016) | (0.021) | (0.019) | | (0.006) | (0.015) | (0.009) | (0.007) | (0.008) |
| $\text{Size}_{ijk,t-1}$ | | | | | | | | | -0.016$^a$ | 0.006$^a$ |
| | | | | | | | | | (0.001) | (0.001) |
| Observations | 4382989 | 4382989 | 121775 | 4382989 | 4382989 | 4382989 | 121775 | 4382989 | 1883888 | 1883888 |
| Firm×Destination×Product FE | No | No | No | Yes | No | No | No | Yes | No | Yes |

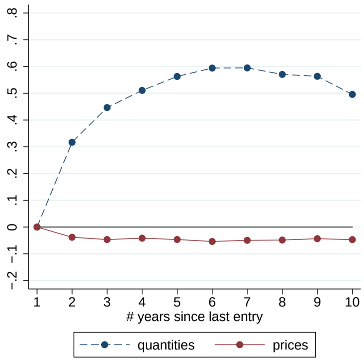
Standard errors clustered by firm in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ and the residuals from the quantity and prices equations (14) and (15). $\text{Size}_{ijk,t-1}$ is defined at the total quantity sold in market $jk$ by firm $i$ in year $t-1$. $\text{Age}_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). $\text{Size}_{ijkt}$ is proxied by the value sold by firm $i$ on market $jk$ during year $t$ divided by the total value exported by French firms in market $jk$ during year $t$. Similar price profiles are obtained when using alternative measures of firm size in columns (9)-(10). "Surv."

Table A.24: Dynamics of quantity and prices (reconstructed years)

| Dep. var. | (1) | (2) $\varepsilon^q_{ijkt}$ | (3) | (4) | (5) | (6) $\varepsilon^p_{ijkt}$ | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| $\text{Age}_{ijkt}$ | $0.085^a$ | | | $-0.012^a$ | | | | |
| | (0.001) | | | (0.001) | | | | |
| $\text{Age}_{ijkt} = 2$ | | $0.109^a$ | $0.025^a$ | | $-0.022^a$ | -0.002 | | |
| | | (0.002) | (0.003) | | (0.001) | (0.002) | | |
| $\text{Age}_{ijkt} = 3$ | | $0.229^a$ | $0.067^a$ | | $-0.037^a$ | $-0.006^b$ | $-0.013^a$ | -0.003 |
| | | (0.004) | (0.005) | | (0.002) | (0.003) | (0.002) | (0.002) |
| $\text{Age}_{ijkt} = 4$ | | $0.312^a$ | $0.087^a$ | | $-0.045^a$ | $-0.008^a$ | $-0.019^a$ | $-0.004^b$ |
| | | (0.004) | (0.006) | | (0.002) | (0.003) | (0.002) | (0.002) |
| $\text{Age}_{ijkt} = 5$ | | $0.364^a$ | $0.089^a$ | | $-0.049^a$ | $-0.009^a$ | $-0.023^a$ | $-0.005^c$ |
| | | (0.005) | (0.007) | | (0.003) | (0.003) | (0.002) | (0.003) |
| $\text{Age}_{ijkt} = 6$ | | $0.419^a$ | $0.088^a$ | | $-0.057^a$ | $-0.014^a$ | $-0.029^a$ | $-0.009^a$ |
| | | (0.007) | (0.009) | | (0.004) | (0.004) | (0.003) | (0.003) |
| $\text{Age}_{ijkt} = 7$ | | $0.461^a$ | $0.087^a$ | | $-0.059^a$ | $-0.014^b$ | $-0.030^a$ | $-0.009^b$ |
| | | (0.009) | (0.012) | | (0.004) | (0.005) | (0.004) | (0.005) |
| $\text{Age}_{ijkt} = 8$ | | $0.491^a$ | $0.079^a$ | | $-0.066^a$ | $-0.015^c$ | $-0.037^a$ | -0.010 |
| | | (0.011) | (0.014) | | (0.007) | (0.009) | (0.007) | (0.008) |
| $\text{Age}_{ijkt} = 9$ | | $0.515^a$ | $0.080^a$ | | $-0.063^a$ | -0.012 | $-0.032^a$ | -0.006 |
| | | (0.016) | (0.019) | | (0.006) | (0.008) | (0.007) | (0.007) |
| $\text{Size}_{ijk,t-1}$ | | | | | | | $-0.015^a$ | $0.011^a$ |
| | | | | | | | (0.001) | (0.001) |
| Observations | 3741140 | 3741140 | 3741140 | 3741140 | 3741140 | 3741140 | 1524261 | 1524261 |
| Firm×Destination×Product FE | No | No | Yes | No | No | Yes | No | Yes |

Standard errors clustered by firm in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. $\varepsilon^q_{ijkt}$ and $\varepsilon^p_{ijkt}$ and the residuals from the quantity and prices equations (14) and (15). $\text{Size}_{ijk,t-1}$ is defined at the total quantity sold in market $jk$ by firm $i$ in year $t-1$. $\text{Age}_{ijkt}$ is the number of years since the last entry of the firm on market $jk$ (reset to zero after one year of exit). $\text{Size}_{ijkt}$ is proxied by the value sold by firm $i$ on market $jk$ during year $t$ divided by the total value exported by French firms in market $jk$ during year $t$. "Surv." means that we restrict the sample to firms-markets surviving the entire period. This table contains estimations similar to table A.23, except that they are ran on a sample in which years have been reconstructed by firm-market, starting from the first month of export.

Figure A.5: Dynamics of prices and quantities residuals: surviving firms (1996-2005)




Note: This figure plots the coefficients obtained when regressing the prices and quantities residuals $\varepsilon^{p}_{ijkt}$ and $\varepsilon^{q}_{ijkt}$ on a set of age dummies and restricting the sample to firms-markets surviving the entire period. The complete set of coefficients and standard errors are shown in Table A.23 (columns (4) and (8)).

# L Variance of growth rates: robustness

Tables A.25 and A.26 report the full set of results used to draw figure 3 in the main text. Table A.25 shows that the variance of both $\varepsilon^{q}_{ijkt}$ and $\varepsilon^{p}_{ijkt}$ decreases with age. As expected, the decline in the variance is larger for the quantity residuals. These results are robust to controlling for the number of observations (columns (3) and (7)), focusing on permanent exporters that survive throughout our time span (columns (4) and (8)), controlling for the average firm size in the cohort (columns (3)-(6) of Table A.26) or using our alternative definitions of age (columns (7)-(14) of Table A.26). Finally, columns (1) and (2) of Table A.26 confirm that the variance of $\varepsilon^{value}_{ijkt}$ decreases sharply with age as well.

Table A.25: Prediction 2.b: age and variance of growth rates

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\mathbb{V}(\Delta\varepsilon^{q}_{ijkt})$ | | | | $\mathbb{V}(\Delta\varepsilon^{p}_{ijkt})$ | | | |
| Age definition | # years since last entry (reset after 1 year of exit) | | | | # years since last entry (reset after 1 year of exit) | | | |
| Sample | All | | | Permanent exporters[1] | All | | | Permanent exporters[1] |
| $\text{Age}_{ijkt}$ | -0.051$^a$ | | -0.045$^a$ | -0.018$^a$ | -0.029$^a$ | | -0.024$^a$ | -0.007$^a$ |
| | (0.001) | | (0.001) | (0.002) | (0.001) | | (0.001) | (0.001) |
| $\text{Age}_{ijkt} = 3$ | | -0.110$^a$ | | | | -0.066$^a$ | | |
| | | (0.004) | | | | (0.002) | | |
| $\text{Age}_{ijkt} = 4$ | | -0.169$^a$ | | | | -0.098$^a$ | | |
| | | (0.005) | | | | (0.003) | | |
| $\text{Age}_{ijkt} = 5$ | | -0.211$^a$ | | | | -0.120$^a$ | | |
| | | (0.005) | | | | (0.003) | | |
| $\text{Age}_{ijkt} = 6$ | | -0.236$^a$ | | | | -0.134$^a$ | | |
| | | (0.006) | | | | (0.004) | | |
| $\text{Age}_{ijkt} = 7$ | | -0.269$^a$ | | | | -0.148$^a$ | | |
| | | (0.008) | | | | (0.004) | | |
| $\text{Age}_{ijkt} = 8$ | | -0.303$^a$ | | | | -0.164$^a$ | | |
| | | (0.009) | | | | (0.005) | | |
| $\text{Age}_{ijkt} = 9$ | | -0.295$^a$ | | | | -0.171$^a$ | | |
| | | (0.012) | | | | (0.007) | | |
| $\text{Age}_{ijkt} = 10$ | | -0.338$^a$ | | | | -0.185$^a$ | | |
| | | (0.018) | | | | (0.011) | | |
| # observations | | | 0.008$^a$ | 0.006 | | | 0.005$^a$ | 0.005 |
| | | | (0.001) | (0.006) | | | (0.000) | (0.004) |
| Observations | 434593 | 434593 | 434593 | 44421 | 434593 | 434593 | 434593 | 44421 |
| Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Standard errors clustered by cohort in parentheses. A cohort of exporters in a product-destination market includes all firms starting to export to that market in a given year. Cohort fixed effects included in all estimations. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. [1] firms present all years on market $jk$. "# observations " is the number of observations of the cohort in the current year.

Table A.26: Prediction 2: variance of growth rates: robustness

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\mathbb{V}(\Delta\varepsilon^{value}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{q}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{p}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{q}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{p}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{q}_{ijkt})$ | | $\mathbb{V}(\Delta\varepsilon^{p}_{ijkt})$ | |
| Age definition | # years since last entry (reset after 1 year of exit) | | | | | | # years since last entry (reset after 2 years exit) | | | | # years exporting since first entry | | | |
| $\text{Age}_{ijkt}$ | -0.044$^a$ | | -0.049$^a$ | | -0.026$^a$ | | -0.053$^a$ | | -0.038$^a$ | | -0.050$^a$ | | -0.035$^a$ | |
| | (0.001) | | (0.001) | | (0.001) | | (0.001) | | (0.001) | | (0.001) | | (0.001) | |
| $\text{Age}_{ijkt} = 3$ | | -0.093$^a$ | | -0.107$^a$ | | -0.060$^a$ | | -0.082$^a$ | | -0.069$^a$ | | -0.063$^a$ | | -0.054$^a$ |
| | | (0.003) | | (0.004) | | (0.002) | | (0.004) | | (0.002) | | (0.004) | | (0.003) |
| $\text{Age}_{ijkt} = 4$ | | -0.139$^a$ | | -0.165$^a$ | | -0.089$^a$ | | -0.138$^a$ | | -0.111$^a$ | | -0.116$^a$ | | -0.090$^a$ |
| | | (0.004) | | (0.005) | | (0.003) | | (0.004) | | (0.003) | | (0.004) | | (0.003) |
| $\text{Age}_{ijkt} = 5$ | | -0.179$^a$ | | -0.206$^a$ | | -0.111$^a$ | | -0.188$^a$ | | -0.148$^a$ | | -0.165$^a$ | | -0.128$^a$ |
| | | (0.005) | | (0.006) | | (0.003) | | (0.005) | | (0.003) | | (0.005) | | (0.003) |
| $\text{Age}_{ijkt} = 6$ | | -0.201$^a$ | | -0.231$^a$ | | -0.124$^a$ | | -0.230$^a$ | | -0.173$^a$ | | -0.211$^a$ | | -0.154$^a$ |
| | | (0.005) | | (0.007) | | (0.004) | | (0.006) | | (0.004) | | (0.005) | | (0.004) |
| $\text{Age}_{ijkt} = 7$ | | -0.230$^a$ | | -0.264$^a$ | | -0.138$^a$ | | -0.279$^a$ | | -0.197$^a$ | | -0.258$^a$ | | -0.182$^a$ |
| | | (0.007) | | (0.008) | | (0.005) | | (0.006) | | (0.004) | | (0.006) | | (0.004) |
| $\text{Age}_{ijkt} = 8$ | | -0.252$^a$ | | -0.298$^a$ | | -0.153$^a$ | | -0.311$^a$ | | -0.223$^a$ | | -0.297$^a$ | | -0.209$^a$ |
| | | (0.008) | | (0.009) | | (0.005) | | (0.008) | | (0.005) | | (0.008) | | (0.005) |
| $\text{Age}_{ijkt} = 9$ | | -0.258$^a$ | | -0.290$^a$ | | -0.160$^a$ | | -0.346$^a$ | | -0.251$^a$ | | -0.336$^a$ | | -0.242$^a$ |
| | | (0.010) | | (0.012) | | (0.007) | | (0.010) | | (0.006) | | (0.010) | | (0.006) |
| $\text{Age}_{ijkt} = 10$ | | -0.287$^a$ | | -0.332$^a$ | | -0.170$^a$ | | -0.433$^a$ | | -0.293$^a$ | | -0.423$^a$ | | -0.283$^a$ |
| | | (0.015) | | (0.018) | | (0.011) | | (0.018) | | (0.010) | | (0.017) | | (0.010) |
| $\text{Size}_{t-1}$ | | | -0.015$^a$ | -0.008$^a$ | -0.013$^a$ | -0.009$^a$ | | | | | | | | |
| | | | (0.002) | (0.002) | (0.001) | (0.001) | | | | | | | | |
| Observations | 434593 | 434593 | 434593 | 434593 | 434593 | 434593 | 246772 | 246772 | 246772 | 246772 | 251236 | 251236 | 251236 | 251236 |
| Cohort FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Standard errors clustered by cohort in parentheses. A cohort of exporters in a product-destination market includes all firms starting to export to that market in a given year. Cohort fixed effects included in all estimations. $^c$ significant at 10%; $^b$ significant at 5%; $^a$ significant at 1%. $^1$ firms present all years on market $jk$. "# observations " is the number of observations of the cohort in the current year.

# M Additional references


Ayyagari, M., Demirguc-Kunt, A. and Maksimovic, V. (2011), "Small vs. young firms across the world : contribution to employment, job creation, and growth", Policy Research Working Paper Series 5631, The World Bank.

Berman, N., de Sousa, J., Martin P. and T. Mayer (2013), "Time to Ship during Financial Crises.", *NBER International Seminar on Macroeconomics*, 9, 225-260.

Bricongne, J.-C., Fontagne, L., Gaulier, G., Taglioni, D. and Vicard, V. (2012), "Firms and the global crisis: French exports in the turmoil", *Journal of International Economics*, vol. 87 no 1: pp. 134-146.

Criscuolo, C., Gal, P. N. and Menon, C. (2014), "The Dynamics of Employment Growth: New Evidence from 18 Countries", CEP Discussion Papers 1274, Centre for Economic Performance, LSE.

Lawless, M. (2014), "Age or size? Contributions to job creation", *Small Business Economics*, vol. 42 no 4: pp. 815-830.

Lawless, M. and Whelan, K. (2014), "Where do firms export, how much, and why?", *World Economy*, vol. 37.

Nunn, N. (2007), "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade.", *The Quarterly Journal of Economics*, 122, 569-600.