Demand learning and exporter dynamics*

Nicolas Berman[†] Vincent Rebeyrol[‡] Vincent Vicard[§]

First version: November 2014
This version: December 9, 2014

Abstract: This paper provides direct evidence that learning about local demand is an important driver of exporters' dynamics. We present a simple trade model with Bayesian learning in which firms are uncertain about their idiosyncratic demand parameter in each of the markets they serve, and update their beliefs as noisy information arrives at each period. We derive three main predictions: (i) a new demand shock leads firms to update more their beliefs about future demand, the younger they are; (ii) the absolute value of firms' idiosyncratic growth rates and (iii) their variance across firms decrease with age. We find strong support for these predictions on detailed French firm-level data. Our data contains both the values and the quantities sold by a given firm, for the same product, in different destination markets, which allows us to purge firm sales from productivity variations and to identify separately both the demand shocks faced by the firms and their belief about future demand. The last part of the paper shows that market-specific firm exit behavior is also consistent with a model of demand learning.

JEL classification: F12, F14, L11, L25
Keywords: firm dynamics, trade, learning, demand

*We thank Fabien Gensbittel, Sacha Kapoor, Marti Mestieri, Franck Portier for very useful discussions and comments. The authors gratefully acknowledge financial support from Fondation Banque de France. This paper features an online appendix containing additional results and available on the authors' webpages. The opinions expressed in this paper are those of the authors and do not reflect the views of the Banque de France.

[†]Department of International Economics, Graduate Institute of International and Development Studies, and CEPR. E-mail: nicolas.berman@graduateinstitute.ch.

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

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

1 Introduction

Why do some firms grow faster than others? While some firms 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.*, 2007). In the case of French firms, 53.5% of total foreign sales are made by firms that did not serve these markets a decade earlier.[1] Among these, 40% come from the post-entry growth of sales on each market. 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, for those that survive, the average growth of their sales declines with their age and size, suggesting a violation of Gibrat's law.[2] Similar behaviors have been recently reported for sales on foreign markets.[3] These facts can be rationalized by several theories, featuring very different underlying mechanisms, such as stochastic productivity growth, endogenous R&D investment or demand learning. Empirically however, disentangling the role of these specific channels has been proven difficult, as it requires identifying separately the contributions of idiosyncratic demand and productivity to the variations of firms sales. This paper focuses on demand learning and provides direct evidence that it is an important driver of post-entry firm dynamics on the export markets.

We first present a simple trade model with Bayesian demand learning, in the spirit of Jovanovic (1982).[4] 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 at each period. These signals determine the firms' posterior beliefs about demand, from which they make their quantity decision. We derive three standard predictions which are specific to the learning mechanism: (i) a new signal leads a firm to update more its belief, the younger the firm is; (ii) controlling for aggregate market conditions and firm productivity, the absolute values of firm-market specific growth rates of sales, quantities and prices decrease with age; (iii) within cohort, the variance of growth rates decreases with age.

In order to test these predictions, we derive from the model an empirical methodology which allows to separately identify the firms' beliefs and the demand shocks (the signal) they face each period, in each market they serve. We use detailed exporter-level data containing the values and the quantities sold by French firms, by product and destination, over the period 1994-2005. We proceed in two steps. First, we purge firms sales, quantities and prices from market-specific conditions and from firm-specific supply side dynamics (e.g, productivity). In doing so, we take advantage of a unique feature of international trade data, in which we can observe the same firm selling the same product in different markets. This is key as it enables to cleanly separate productivity from demand variations. In addition, observing different firms selling the

[1] These numbers are based on the 1996-2005 period – see Section 2.

[2] See Evans (1987), Dunne *et al.* (1989), Caves (1998), Cabral and Mata (2003) and Haltiwanger *et al.* (2013) among many others.

[3] Eaton *et al.* (2007), Berthou and Vicard (2014), Bernard *et al.* (2014), Albornoz *et al.* (2012) or Fernandes and Tang (2014) show that these dynamics are also observed for exporters, and quantitatively magnified.

[4] 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 (2012).

same product in the same destination allows to control for aggregate market-specific conditions. Second, we use the fact that, in our model, quantity decisions only depend on the firms' beliefs (while prices and the value of sales also depend on the realized demand shocks) to separate out the firms' beliefs from the demand signal. Therefore, while requiring few, standard assumptions, our methodology allows to test predictions which directly relate age to the firms' beliefs, rather than age to firm size as typically done by the literature.

We find strong support for all three predictions of the model. The learning process appears to be especially strong in the first years after entry, although even the most experienced firms in our sample still exhibit significant belief updating. Quantitatively, our results suggest that the growth of beliefs explains a larger part of the variations in firm-level export growth than supply side dynamics. We show that these results survive to controlling for firm size, and more generally to the relaxation of a number of modelling hypotheses, in particular regarding the demand function. We use a variety of definitions of age to account for the fact that exporters enter and exit markets frequently and that the accumulated learning of demand might be partially kept even during periods of exit. We show that the bulk of accumulated learning is lost during periods of exit exceeding one year.

The last part of the paper considers the predictions of the model in terms of firm survival. Under the additional assumption that firm productivity follows a Markov process, the model predicts that given age, the probability to exit decreases with the firm belief and with the idiosyncratic demand shocks it faces. Further, a demand shock should lead to more exit in younger cohorts than in older ones. Again, our empirical results are in line with these predictions.

Our paper therefore shows that demand learning is an important characteristic of the micro-dynamics of exporting. By specifically testing the mechanism of beliefs updating which lies at the core of models of firm dynamics with learning, we also more generally contribute to the literature on industry dynamics which tries to understand the determinants of firm growth and survival. Our results lend support to a class of models featuring learning (Jovanovic, 1982) which have recently been used to study exporters' dynamics (Eaton *et al.*, 2014; Albornoz *et al.*, 2012; Timoshenko, 2012; Fernandes and Tang, 2014).

An alternative class of models explains firm and exporter dynamics through stochastic shocks to productivity (Hopenhayn, 1992; Luttmer, 2007, 2011; Arkolakis, 2013)[5] and/or endogenous productivity variations (Klette and Kortum, 2004; Rossi-Hansberg and Wright, 2007). Both the theories based on demand learning and the ones based on productivity variations can replicate qualitatively most stylized facts that we observe in the data. Arkolakis (2013), for instance, shows that a model combining stochastic productivity growth (as in Luttmer, 2007) and market penetration costs (as in Arkolakis, 2010) can reproduce facts observed on the domestic and export markets on entry-exit rates, and on the relationship between average firm sales growth (or their variance) and firm age (or size). But the literature strikingly lacks direct empirical evidence documenting the relative relevance of these alternative mechanisms. A major contribution of our paper is to properly identify the idiosyncratic demand component of firms sales, i.e. to purge those sales from their firm-specific productivity component. This allows us to make a statement about the relevance of demand learning that is robust to *any possible* dynamics of

[5]See also Impullitti *et al.* (2013)

firm productivity.

Note that we concentrate on post-entry dynamics, i.e. exporters' growth and survival. Entry decisions in a given destination might be affected by the beliefs of the firm on other destinations (Albornoz *et al.*, 2012), or on other products for the same destination (Timoshenko, 2012). These effects might be stronger for similar destinations and products (Morales *et al.*, 2014; Defever *et al.*, 2011; Lawless, 2009). The behavior of other firms serving the same market might also play a role (Fernandes and Tang, 2014). These are interesting but quite vast and different questions, which we indeed plan to study in future work, but that are beyond the scope of this paper.

From a methodological point of view, our paper is related to Foster *et al.* (2008, 2013) and Li (2014). Foster *et al.* (2008) use data on the prices and quantities of US homogenous goods producers to separate idiosyncratic demand shocks from firms' productivity, and quantify the effect of both elements on firm selection. Using the same sample, Foster *et al.* (2013) find that demand accumulation explains a large part of the relationship between firm age and firm size. Contrary to these papers, our methodology does not require measuring firm productivity to identify demand shocks. We explicitly control for all time-varying, firm-specific determinants of sales (these include productivity but also for instance capital constraints). This ensures that market specific demand learning/accumulation is the only source of dynamics driving our results. Another difference is that we focus on "passive" demand learning while Foster *et al.* (2013) consider "active" demand accumulation (through pricing). Our paper also relates to Li (2014) who adds Bayesian demand learning to a structural model of export dynamics in the line of Roberts *et al.* (2012), and estimate it on a set firms belonging the Chinese ceramic industry. Beyond methodological differences, our focus is different: Li (2014) studies exporters' entry decisions, while we concentrate on post-entry dynamics.

In theory, firms can learn about demand passively (by observing demand shocks and consequently updating their beliefs), or actively (by engaging in specific investments).[6] We focus on the first type of process. While we do not rule out the possibility that both types of learning coexist, we show that our methodology makes very unlikely that our results reflect active demand learning, as it explicitly controls for all variations in firm-specific expenditures. We also provide results which directly support our interpretation using a test initially proposed by Pakes and Ericson (1998).

The empirical relevance of firm learning has implications for the modeling of firm and industry dynamics in general. The most direct one is that firm size is not only driven by supply side factors but also reflects the evolution of managers' beliefs about their profitability. Therefore, models which aim at explaining the dynamics of firm size distribution (within and across industries) based solely on productivity dynamics would gain at introducing demand learning mechanisms. Second, our results imply that firms at different stages of their learning process will respond differently to idiosyncratic demand shocks. They also suggest that firms of different ages do not face the same amount of uncertainty, which might have implications for the impact of uncertainty shocks on aggregate outcomes (Bloom *et al.*, 2012). Finally, we find that it takes time for firms to discover their profitability in a given market (we find evidence of learning even 7 years after entry), and that this "learning capital" is quickly forgotten during exit periods. The next step –

[6]See for instance Ericson and Pakes (1995); Pakes and Ericson (1998) or Abbring and Campbell (2005).

which bears important policy relevance – is to try to understand which factors affect the speed of learning.

The paper proceeds as follows. In the next section, we describe our data and provide descriptive evidence on the contributions of new firms / markets growth and survival to overall export growth. In section 3 we present our model and its predictions related to firm growth. Section 4 describes our identification strategy, and section 5 our main results as well as a number of robustness exercises. In section 6 we consider firm survival. The last section concludes.

2 Firm dynamics on foreign markets and export growth

In this section we describe our data and present descriptive statistics about the dynamics of French exports. We first decompose total export growth into the contributions of the various margins of trade, to emphasize the role of young firms and new destination markets to aggregate growth and the contribution of post-entry growth. We then focus on the importance of firm-destination-product effects in explaining the variance of firms' sales on the markets they serve.

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. The data cover the period from 1994 to 2005, and contains information about both the value and quantity exported by firms, which will allow us to compute firm-market specific unit values which we will use as a proxy for firm price in the second part of the paper.[7]

A product is defined at the 6-digit level (HS6). We focus on HS6 product categories that do not change over the time-period in order to be able to track firms over time on a specific market (destination-and-product).[8] Moreover, we concentrate on the years 1996-2005, as this will be the period considered later in our estimations. The underlying reason is that we use the two first years, 1994 and 1995, to identify entry, as explained in more details in section 4.1. Our final dataset covers exports of 4,193 HS6 product categories to 180 destination markets by 102,943 firms over the period 1996-2005.

2.2 Contribution to the dynamics of aggregate exports

We concentrate in this paper on post-entry firm dynamics, i.e. on the behavior, in terms of sales growth and survival, of the firms-product-destination triplets entering the export markets over our period of study. This section provides descriptive evidence that these new exporters dynamics account for a large part of aggregate French exports growth over the 1996-2005 period.

[7]Two different thresholds apply to the declaration of export transactions, depending on the country of destination. The declaration of extra-EU export flows is mandatory when a flow exceeds 1,000 euros or 1,000 kg. For transactions to EU countries, firms have to report their expeditions when their total exports to all EU countries exceed 150,000 euros over the year. This absence of declaration for small intra-EU flows might introduce noise in our measures of age; we will check that all our results are unchanged when removing EU destinations from the sample.

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

Table 1 performs two exercises. In panel A, we first decompose total export growth into the contributions of firm and products-destinations entry and exit (the net "extensive margin"[9]) and of the pure intensive margin (i.e. the growth of firm-product-destination triplets already present in 1996). We follow the decomposition proposed by Bricongne *et al.* (2012), to which we refer the reader for more details. Column (1) shows the average yearly contribution of each margin, while column (2) concentrates on the contribution to total growth of French exports over the entire time-period. On a yearly basis, the majority of export growth comes from incumbents (column 1, Panel A). Over a decade however, new firms and markets account for almost two third of overall export growth (column 2, Panel A). This is not a French specificity: Eaton *et al.* (2007) and Bernard *et al.* (2009) provided similar evidence on Colombian and US firms, respectively.

Table 1: Contribution to the growth of aggregate exports

	Average yoy 1996/2005	Overall 1996/2005
A. Contributions to export growth		
Net extensive margin	44.2%	63.5%
Net Firm entry	*26.9%*	*39.7%*
Net new product-destination	*17.3%*	*23.8%*
Net intensive margin	55.8%	36.5%
Total	100%	100%
B. Shares in end-of-period exports		
New firms	2.4%	26.2%
Initial size	-	*16.5%*
Growth since entry	-	*9.7%*
New product-destination	6.9%	27.3%
Initial size	-	*16.1%*
Growth since entry	-	*11.3%*
Incumbent firm-product-destination	87.7%	46.5%
Total	100%	100%

Note: sample of HS6 fixed over time. Source: French Customs.

Over time, the contribution of new cohorts of exporters by product-and-destination depends on three components: entry on new markets, survival and post entry growth on these new markets. Since new exporters typically do not survive more than one year on export markets[10], these entries and exits inflate year-on-year gross margins but do not contribute to the growth of aggregate exports over longer time horizon. Another way to see the contribution of the new cohorts of exporters is provided in Panel B of Table 1, in which we compute the share of entrants (new firms or new destinations-products) after one year (column (1)) and at the end of

[9]Table A.1 of the online appendix further decomposes the extensive margins into the contribution of entries and exits.

[10]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-years horizon.

our sample period (column (2)). These new firms and market represent only 12% of exports in their first year, but account for 53.5% of total exports after a decade (27.3% due to new markets served by incumbents, 26.2 by new firms). How much of this share can be attributed to pure growth since entry, and how much to initial firm size? Many papers have shown that exports are highly concentrated in the hand of large exporters (see e.g. Mayer and Ottaviano, 2008). Are these firms already large at the time of entry? It is clear from column (2) that initial size is only part of the story: around 40% of the end-of-period share of newly created flows comes from their growth since entry.

These descriptive statistics underline the importance of post-entry dynamics (of both new firms and incumbent exporters) to explain aggregate export dynamics, besides the process of entry into foreign markets. The objective of this paper is precisely to understand the sources of these dynamics – including export growth and survival –, emphasizing in particular the role of demand learning.

2.3 Contribution to sales variations

Having shown that a large portion of export growth is due to post-entry dynamics of newly created flows, we now focus more specifically on the determinants of post-entry sales. Using firm-destination data, Eaton *et al.* (2011) showed that firm-specific effects explain well the probability of serving a market (57%), but less so sales variations conditional on selling on a market (39%).[11] In the spirit of Eaton *et al.* (2011), we perform a number of estimations in which we regress firm-market specific export growth on various sets of fixed effects.

First, we regress firm-market specific sales growth on a set of destination-product-time dummies.[12] The R^2 of such a regression is 0.12; market-specific dynamics play a limited role. Adding firm-product-time fixed effects, we get an R^2 of 0.44, 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 precisely on this part of firm dynamics, with the objective to understand to which extent it can be explained by learning about demand. Anticipating a bit on our results, we will indeed find that the R^2 increases to 0.87 when we include our estimates of the growth of firms beliefs about future demand in the regression of column (2), which suggests that learning on demand is at least as important as supply side dynamics to explain the growth of firm sales.

3 A simple model of firm age and firm growth

In this section we present a standard model of international trade with Dixit-Stiglitz monopolistic competition and demand learning. It will be at the basis of our identification of the effect of demand learning on firm growth and survival. We index firms by i, destination markets by j, products by k and time by t.

[11] 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 on a product-destination market. Lawless and Whelan (2008) find an adjusted pseudo-R2 of 45% on a sample of Irish exporters.

[12] Table A.2 in the online appendix contains the results.

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:

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

with ρ the discount factor, Ω_{kt} the set of varieties of product k available at time t, 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:

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

where σ_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 ε_{ijkt} a white noise. $\overline{a_{ijk}}$ is an idiosyncratic constant parameter and is unknown to the firm.

Production. At 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). Therefore, the quantity decision is a static decision.

We do not make any assumption on the evolution of firm productivity at the product level over time. Our results will be consistent with virtually any possible dynamics of firms unit costs at the product level. Productivity might be driven by various mechanisms, as proposed by the literature: It could simply be fixed over time, or evolve according to a Markov process (Hopenhayn, 1992), or be affected by firm's investments decisions (Klette and Kortum, 2004; Rossi-Hansberg and Wright, 2007).[13] Additionally, as originally proposed by Jovanovic (1982), it may also be subject to learning. In that case, the firm would take a quantity decision based on its belief about its costs. As we will not back out learning from firms' productivity[14], we do not add expectation terms here to save on notations. The only key assumption here is that firms unit costs at the firm-product level are *not* destination specific.

Per period profits in market j from product k are thus given by:

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

[13] We do not want to make any assumption about the evolution of firm productivity, mainly because, as underlined for instance by Luttmer (2011), "Deciding on the relative importance of these alternative interpretations poses difficult identification problems [...] Much remains to be done to sort out the relative importance of each of these aspects of firm heterogeneity."

[14] We come back to this point in section 4. We concentrate on demand learning because identifying firm idiosyncratic demand requires few assumptions, while identifying learning on firm productivity – and more generally computing firms unit costs – comes at the expense of making more heroic hypotheses.

where w_{it} is the wage rate in the origin country, φ_{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, the firm's prior beliefs about $\overline{a_{ijk}}$ are normally distributed with mean θ_0 and variance σ_0^2. The firm observes t independent signals about $\overline{a_{ijk}}$: $a_{ijkt} = \overline{a_{ijk}} + \varepsilon_{ijkt}$, where each ε_{ijkt} is normal with (known) mean 0 and variance σ_ε^2. According to Bayes' rule, the firm's posterior beliefs about $\overline{a_{ijk}}$ after t signals are normally distributed with mean $\widetilde{\theta}_t$ and variance $\widetilde{\sigma}_t^2$, where:[15]

$$\widetilde{\theta}_t = \theta_0 \frac{\frac{1}{\sigma_0^2}}{\frac{1}{\sigma_0^2} + \frac{t}{\sigma_\epsilon^2}} + \overline{a}_t \frac{\frac{t}{\sigma_\epsilon^2}}{\frac{1}{\sigma_0^2} + \frac{t}{\sigma_\epsilon^2}} \tag{3}$$

$$\widetilde{\sigma}_t^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{t}{\sigma_\epsilon^2}} \tag{4}$$

and $\overline{a}$ is the average signal value, $\overline{a}_t = \left(\frac{1}{t}\sum_t a_{ijkt}\right)$. Note that contrary to $\widetilde{\theta}_t$, the posterior variance $\widetilde{\sigma}_t^2$ does not depend on the realizations of the signals and decreases only with the number of signals (i.e. learning reduces uncertainty). The posterior variance is thus always smaller than the prior variance, $\widetilde{\sigma}_t^2 < \widetilde{\sigma}_{t-1}^2$. In the following, it will be useful to formulate the Bayesian updating recursively. Denoting $\Delta\widetilde{\theta}_t = \widetilde{\theta}_t - \widetilde{\theta}_{t-1}$, we have:

$$\Delta\widetilde{\theta}_t = g_t \left(a_{ijkt} - \widetilde{\theta}_{t-1}\right) \text{ with } g_t = \frac{1}{\frac{\sigma_\epsilon^2}{\sigma_0^2} + t} \tag{5}$$

Intuitively, observing a higher-than-expected signal, $a_{ijkt} > \widetilde{\theta}_{t-1}$ leads the agent to revise the expectation upward, $\widetilde{\theta}_t > \widetilde{\theta}_{t-1}$, 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".

Firm size. 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. $\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}$, are observed by firms before making their quantity decision.[16] This leads to the following optimal

[15] To ease the reading and with some notational abuse, we drop the subscript ijk from the terms θ, $\overline{a}$ and σ, despite the fact that those are firm-market-specific.

[16] As for firm productivity, we could have assumed alternatively that market conditions are not observable at time t before the quantity decision is made. In that case, we would have to assume that all firms selling varieties of product k have the same belief about aggregate market conditions in destination j at time t.

quantities and prices:[17]

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

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

$$S_{ijkt} = q^*_{ijkt}p^*_{ijkt} = \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right)^{1-\sigma_k} \left(\frac{\mu_k Y_{jt}}{P_{jkt}^{1-\sigma_k}}\right) \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]^{\sigma_k - 1} e^{\frac{a_{ijkt}}{\sigma_k}} \tag{9}$$

with $\mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right] = \int e^{\frac{a_{ijkt}}{\sigma_k}} dG_{t-1}(a_{ijkt})$.

As the firm make quantities decisions before observing demand for their product, q^*_{ijkt} depends on the expected demand, not on the demand realization, contrary to prices and therefore sales.[18]

The literature has typically computed correlations between firm age and firm growth rates, and attributed negative ones as potential evidence of demand learning. Indeed, as we formally show in the next subsection, the fact that younger firms adjust more their belief leads growth rate to decrease with age in absolute value. But of course, as is clear from equations (7) to (9), firm sales – and therefore firm growth – also depend on the evolution of market-specific conditions and firm productivity, which might be correlated with firm age. Directly testing for the presence of demand learning thus requires either making assumptions about aggregate market conditions and firm productivity or finding a way to account for them. Our methodology follows the second path.

Let us decompose optimal quantities, prices, and sales into three components. They first depend on unit costs, which are a function of wages in country i and firm-product specific productivity φ_{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[19], 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 belief 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

[17]See the appendix for details.

[18]This is one reason why previous empirical studies testing for firm learning based on Jovanovic (1982) have concentrated on employment (and not sales) as a measure of firm size (see Dunne *et al.*, 1989, Evans, 1987 among others).

[19]Prices do not actually depend on aggregate market conditions. This a consequence of constant marginal costs implied by the CES assumption. As other utility functions – e.g., quasi-linear, translog – would generate variable markups and make price depend on market specific conditions (and in particular on market size), our empirical strategy however systematically checks the robustness of our results to the inclusion of market-specific conditions in the price equation as well.

rewrite the above expressions for sales, quantities and prices as:

$$S^*_{ijkt} = C^S_{ikt} C^S_{jkt} Z^S_{ijkt} \tag{10}$$
$$q^*_{ijkt} = C^q_{ikt} C^q_{jkt} Z^q_{ijkt} \tag{11}$$
$$p^*_{ijkt} = C^p_{ikt} Z^p_{ijkt} \tag{12}$$

As just underlined, the impact of firm demand learning should be fully and only included in the Z^X_{ijkt}, $X = \{S, q, p\}$ terms. These terms can be understood as the quantities, price and sales of the firm, purged from firm unit costs and from aggregate market conditions. As such, they may potentially be very different than the actual firm size and firm price. From a methodological point of view, we stress that any prediction about firm demand learning should be based on the Z^X_{ijkt} terms rather than q^*_{ijkt}, p^*_{ijkt} and/or S^*_{ijkt}. This is one contribution of this paper: we provide a simple methodology to isolate the Z^X_{ijkt} terms, which allows to distinguish the beliefs from the demand shock components. Our approach is consistent with any underlying dynamic process for the ikt and jkt terms. Finally, this also means that we do not look at the dynamics of firm size (at least per se), but directly at the evolution of the firms' beliefs. Before describing our identification strategy, we detail the testable predictions of the model that are specific to firms learning about demand.

3.2 Firm growth and firm age: predictions

At least since Jovanovic (1982), firm learning has been put forward as a mechanism able to explain important stylized facts about the dynamics of firms, and more specifically about their growth rate, and their entry and exit decisions. We start with our predictions about firm growth. The following predictions are usually associated in the literature with learning and are all present in Jovanovic (1982). A major difference with the subsequent empirical literature is that we test these predictions directly on firm beliefs about expected demand and not on firm size.[20]

The growth rates of the Z^X_{ijkt} can be expressed as:

$$\Delta \ln Z^q_{ijk,t+1} = \sigma_k \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{13}$$
$$\Delta \ln Z^p_{ijk,t+1} = \frac{1}{\sigma_k} \Delta a_{ijkt+1} - \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{14}$$
$$\Delta \ln Z^S_{ijk,t+1} = (\sigma_k - 1) \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] + \frac{1}{\sigma_k} \Delta a_{ijkt+1} \tag{15}$$

with $\Delta a_{ijkt+1} = a_{ijkt+1} - a_{ijkt}$ and $\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] = \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] - \ln \mathrm{E}_{t-1} \left[e^{\frac{a_{ijkt}}{\sigma_k}} \right]$.

The growth of firm's beliefs about expected demand can be expressed as:[21]

$$\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] = \frac{1}{\sigma_k} \left(\Delta \widetilde{\theta}_t + \frac{\widetilde{\sigma}^2_t - \widetilde{\sigma}^2_{t-1}}{2\sigma_k} \right) \tag{16}$$

[20] Our predictions highlight a direct impact of firm age on firm dynamics, even after controlling for size. This is an important difference with the literature on (stochastic) productivity dynamics. We show that our results survive when we control explicitly for firm size in the estimations.

[21] The details of the computation, as well as proofs for the following 3 predictions are relegated in the appendix.

At the beginning of period t, firms make quantity decisions based on their belief about local demand for their product. Then, demand is realized and firms update their belief. A higher than expected demand, induced by $a_{ijkt} > \widetilde{\theta}_{t-1}$, leads the firm to update upwards its belief. As a consequence, the expected growth rate of the belief between period t and $t+1$, will be positive. The opposite is true for a lower than expected demand. Importantly, as clear from equation (16), this upward or downward updating is larger for younger firms. It follows our first prediction, that directly illustrates the updating process:

Prediction # 1 (updating): *A new signal a_{ijkt} leads to a larger updating of the belief, the younger the firm is.*

Proof. See appendix.

In order to test this prediction, we need to identify the demand shock a_{ijkt} as well as the growth of firm's beliefs about expected demand as expressed in (16), which is only driven by firm's belief and firm age. It may be also interesting to note that one consequence of this prediction is that, in the absence of any dynamics of the ikt and jkt terms, we should observe a reversion to the mean size after any demand shock. We however want to get closer to the model testing directly for the evolution of the belief and thus allowing for any dynamics of the ikt and jkt terms.

The next two predictions are also closely related to the evolution of $\Delta \ln \mathrm{E}_t[e^{\frac{a_{ijkt+1}}{\sigma_k}}]$. We test for the impact of firm updating on the expected growth rate of Z^q_{ijkt} and Z^p_{ijkt} and the variance of these growth rates within age cohorts. These predictions, originally regarding firm size, were understood as an indication that firm learning can explain deviations from the Gibrat's law.

Prediction # 2 (growth rate): *The expected absolute value of growth rates of Z^q_{ijkt} and Z^p_{ijkt} decrease with firm age.*

Proof. See appendix.

This prediction is also a direct consequence of firm updating. Younger firms update more and thus have higher growth rates in absolute value. As firms get older, their beliefs become more accurate, and growth rate are only affected by the random demand shocks a_{ijkt}. Note that this prediction implies that the expected absolute value of the growth rate of Z^S_{ijkt} should decrease with age as well.

The third prediction follows immediately: as younger firms update more than older firms, the variance of firm growth decreases with the cohort tenure on a specific market.

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

Proof. See appendix.

Prediction 2 also holds for Z^S_{ijkt} provided that the negative covariance between $\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right]$ and Δa_{ijkt+1} is not too strong.[22]

4 Identification

To test predictions 2 and 3, we only need to isolate the Z^X_{ijkt} terms, i.e. we need to purge the quantities, prices and sales from supply side and market specific factors. This is achieved by estimating the following quantities, price and sales equations in logs:[23]

$$\begin{aligned} \ln q_{ijkt} &= \mathbf{FE}_{ikt} + \mathbf{FE}_{jkt} + \varepsilon^q_{ijkt} && (17) \\ \ln p_{ijkt} &= \mathbf{FE}_{ikt} + \varepsilon^p_{ijkt} && (18) \\ \ln S_{ijkt} &= \mathbf{FE}_{ikt} + \mathbf{FE}_{jkt} + \varepsilon^S_{ijkt} && (19) \end{aligned}$$

where q 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. 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. We however systematically check that relaxing this assumption by including jkt fixed effects does not affect the results. Note that we do not have direct price data, so we rely on unit values, defined as S_{ijkt}/q_{ijkt}, to proxy them.

Given that we control for all time-varying, market- and firm-product-specific determinants of quantities, prices and sales, the residuals $\{\varepsilon^q_{ijkt}, \varepsilon^p_{ijkt}, \varepsilon^S_{ijkt}\}$ are by construction orthogonal to the standard determinants of firm dynamics (i.e. productivity and market conditions). This is an important contribution of the paper: our methodology would survive to the inclusion of any process underlying the evolution of firm productivity – including Markov processes, imitation, R&D investments, or even learning –, provided that productivity is the same across destination markets for a given firm-product. Importantly, the ikt fixed effects also control for any other time-varying, firm-specific factors that might affect growth rates. These include in particular financial constraints which have been suggested as being an important determinant of firm dynamics (Cooley and Quadrini, 2001; Cabral and Mata, 2003).

To be more specific, the residuals $\{\varepsilon^q_{ijkt}, \varepsilon^p_{ijkt}, \varepsilon^S_{ijkt}\}$ provide estimates of the Z^X_{ijkt} terms. Using equations (7), (8), (9) and (10), we get:

$$\begin{aligned} \varepsilon^q_{ijkt} &= \ln Z^q_{ijkt} = \sigma_k \ln \mathrm{E}_{t-1} \left[e^{\frac{a_{ijkt}}{\sigma_k}} \right] && (20) \\ \varepsilon^p_{ijkt} &= \ln Z^p_{ijkt} = \frac{1}{\sigma_k} a_{ijkt} - \ln \mathrm{E}_{t-1} \left[e^{\frac{a_{ijkt}}{\sigma_k}} \right] && (21) \\ \varepsilon^S_{ijkt} &= \ln Z^S_{ijkt} = (\sigma_k - 1) \ln \mathrm{E}_{t-1} \left[e^{\frac{a_{ijkt}}{\sigma_k}} \right] + \frac{1}{\sigma_k} a_{ijkt} && (22) \end{aligned}$$

With these residuals at hand, we can directly compute the growth rates of the Z^X_{ijkt} terms, allowing to test for predictions 2 and 3. Note that this identification strategy is possible to

[22] Formally, this will be the case if $\sigma_k > 1 + \frac{\sigma^2_\varepsilon}{\sigma^2_0} + t$. See appendix for details.

[23] We use the Stata routine `reg2hdfe` developed by Guimaraes and Portugal (2010).

implement because we are able to observe the sales of the same product on different destination markets by the same firm. The use of firm-level export data is therefore key as it allows to purge market-specific firm dynamics from the evolution of firm productivity through the inclusion of $\mathbf{FE}_{ikt}$.[24]

Testing prediction 1, which is the very core of the learning mechanism, is slightly more complicated as it implies getting estimates of both the firm beliefs about expected demand $\mathrm{E}_{t-1}[e^{\frac{a_{ijkt}}{\sigma_k}}]$ and the demand shock a_{ijkt}.

We use the assumption that the firm takes its quantity decision before observing the demand realization. It follows that $\ln Z^q_{ijkt}$ only depends on the firm's belief about future demand, while $\ln Z^p_{ijkt}$ is adjusted for the demand shock. The residual ε^q_{ijkt} thus provides us with a direct estimate of the firm belief. We only need to correct for σ_k.

In order to back out the demand shock and get an estimate of σ_k, we regress ε^p_{ijkt} on ε^q_{ijkt}. Using (21) and (20), we get:

$$\left(\frac{1}{\sigma_k}a_{ijkt} - \ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right) = \beta\left(\sigma_k \ln \mathrm{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]\right) + v_{ijkt} \tag{23}$$

We estimate (23) by 6-digit product to allow σ_k to differ across products.[25] We obtain[26] :

$$\widehat{\beta} = -\frac{1}{\sigma_k} \quad \text{and} \quad \widehat{v}_{ijkt} = \frac{1}{\sigma_k}a_{ijkt} \tag{24}$$

Discussion. Before turning to the results, let us come back to the identifying assumptions required by our empirical strategy. A first hypothesis we made is that demand is CES, which implies that markups are constant and do not depend on market conditions (in particular on market size). Section 5.4 discusses the implications of this assumption, in particular the fact that with non-CES demand, the price equation should include jkt fixed-effects, and the estimations should control for firm size (to capture the position of the firm on the demand curve and therefore the elasticity of demand it faces).[27] Importantly, we will also show that our results can be interpreted as evidenc e of demand learning even after relaxing the CES assumption.

The second assumption is that firms cannot fully adjust their quantities produced after observing the realization of the demand shock. If it were not the case, prices and quantities would not depend on the firm's belief and therefore on its age; none of the predictions would hold. A related assumption – commonly made in the literature – is that firms can adjust their prices after demand is realized, but not their quantities. This assumption is important

[24]The reason why we do not model learning about productivity appears more clearly in equations (20) to (22). Identifying demand variations is possible because we are able to control for productivity through the inclusion of ijk fixed effects. On the other hand, we cannot distinguish productivity variations from global demand shocks faced by firms in all the markets, as these will be mixed with unit costs in the $\mathbf{FE}_{ikt}$.

[25]k is defined throughout our analysis as a 6-digit product. One potential issue here is that running estimations at such level of disaggregation implies getting too few observations for some products. We therefore perform a robustness check where equation (23) is estimated at the 4-digit level.

[26]Whenever our estimates of β are statistically insignificant or imply values of σ_k which are lower than 1, we replace $\widehat{v}$ by a missing value and do not consider the observation in the estimations.

[27]Note that in models with stochastic productivity growth (e.g. Luttmer, 2007), age and growth are related only through firm size. The fact that we find a significant relationship between the Z_{ijkt} and firm age on a market also supports the learning mechanism.

to identify separately the demand shock from the belief and therefore to test prediction 1, but bears no impact on the other predictions.[28] Precisely, in our model we only need quantities to adjust *less* than prices for the predictions to hold. We believe this is a realistic assumption, especially given that we look at international trade flows. Empirically, we perform a number of robustness checks related to this assumption. In particular, in section 5.4 we concentrate on sectors and destinations for which it is more likely that production is fixed *ex-ante* (sectors in which adjustment costs are higher).

The last assumption we make is that firms learn their demand "passively". Empirically, this implies in particular that the Z^X_{ijkt} terms do not capture marketing expenses or other expenses aiming at building a consumer base. We discuss this issue specifically in section 5.4, and perform a test initially proposed by Pakes and Ericson (1998) which supports our interpretation of the results. For now, we note that the controls that we include in the estimations from which we back up the Z^X_{ijkt} terms make unlikely the possibility that our results reflect active learning: controlling for $\mathbf{FE}_{ikt}$ purges the Z^X_{ijkt} from all the firm-product specific supply side factors – such as investment or marketing expenses – which might impact the firm demand; controlling for $\mathbf{FE}_{jkt}$ purges Z^X_{ijkt} from all the destination characteristics which might impact the expenses of all French firms exporting a given product to a given market. These include market size or competition, but also regulations, i.e. technical norms, adaptation costs, marketing requirements, or costs of translation.

4.1 Measuring age

The last variable we need to compute to be able to test our three predictions is (firm-market specific) age. A major advantage of exporter-level data is that it allows measuring precisely and following over time firms' sales on each specific destination market. We use the time variation in the product-destination markets served by the firm to measure its market-specific experience. Given that firms enter and exit markets frequently, measuring age requires making assumptions about the learning process and about how information over local demand depreciates over time during periods of exits. Given that our model is silent on this issue, we compute three different variables.

Our baseline measure of age is the number of years since last entry of a firm on a product-destination. We assume complete depreciation of firm specific information upon exit and reset the age to zero whenever the firm exits at least one calendar year from a specific product-destination. Age is be either defined as a single discrete variable or as a set of dummies, to allow the learning processes to be non-linear.

To check robustness, we also define two alternative measures of age. We first assume 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. Second, we assume that firms keep entirely the information 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.

Note that in all the empirical analysis, to ensure the consistency of our measures of age, we

[28] We only need that quantities do not adjust perfectly after the firm has observed the realization of demand.

drop firm-product-destination triplets already served in 1994 and 1995, as these years are used to define entry.

Finally, we define a cohort of new exporters on a product-destination market as 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.[29]

5 Main results

After showing some descriptive statistics on our final sample, we provide our baseline results regarding our three predictions, before discussing their sensitivity to the relaxation of a number of modeling assumptions. Finally, we check the robustness of the results to various measurement issues and provide additional insights on the depreciation of learning capital.

5.1 Sample statistics

Table 2 contains some descriptive statistics about our final sample. Over the period, firm-market specific exports have grown on average by 10%[30]; the firm-market specific beliefs have also been characterized by a slightly positive growth.

Table 2: Sample statistics

	Obs.	Mean	S.D.	Q1	Median	Q3
ln S_{ijkt}	6472999	8.31	2.39	6.91	8.29	9.83
ln q_{ijkt}	6472999	5.28	3.05	3.04	5.06	7.27
ln p_{ijkt}	6472999	3.03	1.87	1.82	3.00	4.19
Δ ln S_{ijkt}	3006343	0.10	1.53	-0.68	0.07	0.86
Δ Prior$_{ijkt}$	2795979	0.03	1.37	-0.73	0.02	0.79
σ_k	2675182	11.14	8.07	5.81	8.09	13.93
Age$^1_{ijkt}$	6472999	2.13	1.71	1	1	3
Age$^2_{ijkt}$	6472999	2.30	1.80	1	2	3
Age$^3_{ijkt}$	6472999	2.42	1.82	1	2	3

Source: Authors computations from French Customs data. Age$^1_{ijkt}$: reset after 1 year of exit; Age$^2_{ijkt}$: reset after 2 years of exit; Age$^3_{ijkt}$: years of exporting.

Firms in our sample are typically young in the markets they serve: the average age is comprised between 2.1 and 2.4 years depending on the definition. This is evidence of the low survival rates observed during the first years a firm serves a particular market, a topic we shall specifically study in the last section of the paper.

Interestingly, our methodology generates reasonable estimates of σ_k: After cleaning the top and bottom percentile of these estimates, we get a median value of 8.1 and an average of 11.1 in our final sample. These numbers are high yet comparable to ones found at similar levels of

[29]For our two alternative definitions of age, cohorts cannot be defined by their year of entry so we consider product-destination specific cohorts depending on age only.

[30]Note that the 'calendar year effect' pointed out by Berthou and Vicard (2014) and Bernard *et al.* (2014) is likely to bias upwards the growth rate between the first and second years, because of the potential incompleteness of the first year of export measured over entire the calendar year. When measuring age by bins as in our estimations, the dummy for year two gets rid of the average bias.

disaggregation by the literature, using very different methodologies. For instance Broda and Weinstein (2006) report average elasticities in the range of 12-17 when estimated at the 7-10 digits level. In Romalis (2007), elasticities are estimated at the HS6-level and are generally comprised between 6 and 11. Imbs and Mejean (2014) provide a detailed literature review, and show that lower estimates are typically obtained when using more aggregated data.[31] Our estimates of σ_k also follow expected patterns: considering Rauch (1999) classification, the median (resp. mean) across products is 8.6 (resp. 11.1) for differentiated goods, 9.9 (resp. 13.6) for referenced priced goods and 13.9 (resp. 16.1) for goods classified as homogenous. These means and medians of σ_k are statistically different across the three groups.

5.2 Baseline results

Prediction 1. Prediction 1 states that following a new signal, updating is larger for younger firms. Put differently, we want to know how the demand shock a_{ijkt} affects the firms beliefs. We estimate:

$$\Delta\varepsilon^q_{ijkt} = \alpha_0 + \alpha_1(\frac{1}{\sigma_k}a_{ijk,t-1}) + u_{ijkt} = \alpha_0 + \alpha_1\widehat{v}_{ijkt} + u_{ijkt} \tag{25}$$

and we expect α_1 to be positive. It should also be lower for older firms, a prediction that we capture by adding interaction terms between firm age and the shock:

$$\Delta\varepsilon^q_{ijkt} = \sum_{g=2}^{G}\alpha_g(\widehat{v}_{ijk,t-1} \times \text{AGE}^g_{ijkt}) + \sum_{g=1}^{G}\beta_g\text{AGE}^g_{ijkt} + u_{ijkt} \tag{26}$$

where AGE^g_{ijkt} are dummies taking the value 1 for each age category $g = 2, ...7+$. Experience is defined by firm-destination-product. In our baseline estimations, it is the number of years since the *last entry* of firm i, product k in destination j; age is reset to zero after one year of exit (the next subsections discusses the sensitivity of the results to the use of alternative age definitions). We expect the α_g to be decreasing with age g.

Note that, as formally shown in the appendix, our model predicts that $\alpha_g = g_t = \frac{1}{\sigma^2_\varepsilon/\sigma^2_0+t}$. g_t is the speed of learning; its specific shape is due to our parametric assumption of Normally distributed priors. Looking at the way in which the α_g coefficients evolve with firm age is useful to understand how firms learn about their demand parameter, and also because it allows to discuss the relevance of the normality assumption used to infer the firms' beliefs using Bayes' rule.

The results are provided in Table 3. The first column considers the effect of demand shocks on the adjustment of the firm' beliefs (equation (25)). Columns (2) to (4) study how this effect varies with age (equation (26)). Column (3) is the same as column (2) except that standard errors are bootstrapped 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 positive demand shock (columns (1)). This adjustment is indeed significantly larger when firms are young (columns (2)-(4)). Including age linearly (column (2)) or as bins (column (4)) leads to the same conclusion.

[31] See Broda and Weinstein (2006), Table IV; Romalis (2007), Tables 3a and 3b; Imbs and Mejean (2014), section 3.2.

Table 3: Prediction 1: demand shocks and beliefs updating

	(1)	(2)	(3)	(4)
Dep. var.	$\Delta\varepsilon^q_{ijkt}$			
Age definition	# years since last entry (reset after 1 year of exit)			
$\widehat{v}$	0.075^a	0.109^a	0.109^a	
	(0.002)	(0.004)	(0.004)	
Age_{ijkt}		-0.040^a	-0.040^a	
		(0.000)	(0.000)	
$\widehat{v}\times\text{Age}_{ijkt}$		-0.009^a	-0.009^a	
		(0.001)	(0.001)	
$\widehat{v}\times\text{Age}_{ijkt} = 2$				0.103^a
				(0.003)
$\widehat{v}\times\text{Age}_{ijkt} = 3$				0.066^a
				(0.004)
$\widehat{v}\times\text{Age}_{ijkt} = 4$				0.057^a
				(0.005)
$\widehat{v}\times\text{Age}_{ijkt} = 5$				0.056^a
				(0.006)
$\widehat{v}\times\text{Age}_{ijkt} = 6$				0.047^a
				(0.008)
$\widehat{v}\times\text{Age}_{ijkt} = 7+$				0.050^a
				(0.007)
Observations	2726474	2726474	2726474	2726474

Robust standard errors 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.

Similarly, bootstrapping the standard errors leaves the results unaffected. Note that the shape of the learning process seems consistent with our assumption of normal priors: age has a stronger effect in the early years. After 7 consecutive years of presence on a market, the extent of belief updating is 50 percent smaller than after entry. Interestingly, our results suggest that even for the most experience exporters, firms still learn about the market, as beliefs still significantly adjusts to demand shocks.

Quantitatively, the evolution of firms' beliefs is crucial in explaining firms' dynamics. Including the growth of beliefs as an explanatory variable of the growth of export performed in section 2 increases the R^2 to 0.87, compared to 0.44 when firm-product-time and product-destination-time fixed effects were included alone. Said differently, our mechanism of demand learning therefore contributes at least as much as supply side factors to the explanation of the variance of firms' sales on specific markets.

Prediction 2. Proposition 2 states that the expected absolute values of the growth rates of both quantities and prices decrease with age. We estimate:

$$\left|\Delta\varepsilon_{ijkt}^{X}\right| = \alpha^X + \beta^X \times \text{AGE}_{ijkt} + u_{ijkt} \qquad \forall X = \{q, p\} \tag{27}$$

Alternatively, we will again relax the linearity assumption and replace AGE_{ijkt} by a set of categorical variables as we did in prediction 1. We expect β^X to be negative. The model also predicts that $|\beta^q| > |\beta^p|$: the growth rate of quantities should decrease relatively faster with age than the growth rate of prices.

Table 4: Prediction 2: age and mean growth rates

	(1)	(2)	(3)	(4)
Dep. var.	$\Delta\varepsilon^q_{ijkt}$		$\Delta\varepsilon^p_{ijkt}$	
Age definition	# years since last entry (reset after 1 year of exit)			
Age_{ijkt}	-0.040[a]		-0.024[a]	
	(0.000)		(0.000)	
$\text{Age}_{ijkt} = 3$		-0.076[a]		-0.053[a]
		(0.001)		(0.001)
$\text{Age}_{ijkt} = 4$		-0.119[a]		-0.079[a]
		(0.002)		(0.001)
$\text{Age}_{ijkt} = 5$		-0.152[a]		-0.096[a]
		(0.002)		(0.001)
$\text{Age}_{ijkt} = 6$		-0.184[a]		-0.109[a]
		(0.002)		(0.001)
$\text{Age}_{ijkt} = 7+$		-0.216[a]		-0.129[a]
		(0.002)		(0.001)
Observations	2795979	2795979	2795979	2795979

Robust standard errors in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Controlling for year dummies does not affect the results.

The results are provided in Table 4. We consider sequentially the growth rate of quantities (columns (1) and (2)) and prices (columns (3) and (4)). Both significantly decrease with firm age.[32] The effect is quantitatively more pronounced in the case of quantities than prices, as predicted by the theory.

Prediction 3. Our last prediction relates the variance of growth rates within cohorts to the age of the cohort. We estimate:

$$\text{Var}\left(\Delta\varepsilon_{ijkt}^{X}\right) = \delta^X \times \text{AGE}_{cjkt} + \mathbf{FE}_{cjk} + u_{ijkt} \qquad \forall X = \{q, p\} \tag{28}$$

where $\mathbf{FE}_{cjk}$ represent cohort fixed effects. As mentioned earlier, we define a cohort of new exporters on a product-destination market as all firms starting exporting in year t. We again expect our coefficient of interest δ^X to be negative: because firms update less their beliefs when

[32]Columns (1) and (2) of Table 11 in the appendix show that this is also the case of firm sales.

they gain experience on a market, their quantities and prices become less volatile.

Table 5: Prediction 3: age and variance of growth rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.		$\text{Var}(\Delta\varepsilon^{q}_{ijkt})$				$\text{Var}(\Delta\varepsilon^{p}_{ijkt})$		
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]
Age_{cjkt}	-0.067^{a}		-0.060^{a}	-0.043^{a}	-0.033^{a}		-0.029^{a}	-0.014^{a}
	(0.001)		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)
$\text{Age}_{cjkt} = 3$		-0.130^{a}				-0.072^{a}		
		(0.003)				(0.002)		
$\text{Age}_{cjkt} = 4$		-0.208^{a}				-0.108^{a}		
		(0.004)				(0.002)		
$\text{Age}_{cjkt} = 5$		-0.271^{a}				-0.134^{a}		
		(0.005)				(0.003)		
$\text{Age}_{cjkt} = 6$		-0.314^{a}				-0.153^{a}		
		(0.006)				(0.003)		
$\text{Age}_{cjkt} = 7+$		-0.375^{a}				-0.184^{a}		
		(0.006)				(0.003)		
# observations			0.007^{a}	0.015^{a}			0.003^{a}	0.003^{c}
			(0.001)	(0.004)			(0.000)	(0.002)
Observations	598821	598821	598821	262849	598821	598821	598821	262849
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered by cohort in parentheses. 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.

The results related to the variance of the growth rate of quantities and prices are provided in Table 5. Columns (1) to (4) consider quantities, columns (5) to (8) use prices as a dependent variable. Within cohort, the variance of the growth rate of both quantities and prices sharply decreases with age in all columns.[33] This is still true when controlling for the number of observations of the cohort (columns (3)-(4) and (7)-(8)). Note that our results are not due to attrition: concentrating on the firms which survive over the entire period in columns (4) and (8) leads to similar conclusions.

5.3 Age definition and the learning process

How fast does demand learning depreciate when the firm exits the market? So far we have treated each entry of firms into a market as a new one: age was reset to zero in case of exit. Table 6 tests the robustness of our results on prediction 1 to alternative definitions of firms' age. Columns (1) to (4) assumes that experience is kept if the firm exits only during one year (but is lost if it does not sell for two years or more). In columns (5) to (8) we make the more extreme

[33]Columns (1) and (2) of Table 12 in the appendix show that the variance of firm sales also significantly decreases with age.

assumption that all experience is kept during exit periods, whatever the length of these periods. Tables A.5 and A.6 in the online appendix contain the equivalent sensitivity exercises applied to predictions 2 and 3.

Table 6: Prediction 1: alternative age definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\Delta\varepsilon^q_{ijkt}$				$\Delta\varepsilon^q_{ijkt}$			
Age definition	# years since last entry (reset after 2 years exit)				# years exporting since first entry			
$\widehat{v}$	0.075^a	0.106^a	0.106^a		0.075^a	0.101^a	0.101^a	
	(0.002)	(0.004)	(0.004)		(0.002)	(0.004)	(0.004)	
Age$_{ijkt}$		-0.036^a	-0.036^a			-0.034^a	-0.034^a	
		(0.000)	(0.000)			(0.000)	(0.000)	
$\widehat{v}\times$Age$_{ijkt}$		-0.008^a	-0.008^a			-0.007^a	-0.007^a	
		(0.001)	(0.001)			(0.001)	(0.001)	
$\widehat{v}\times$Age$_{ijkt} = 2$				0.102^a				0.098^a
				(0.003)				(0.003)
$\widehat{v}\times$Age$_{ijkt} = 3$				0.069^a				0.070^a
				(0.004)				(0.004)
$\widehat{v}\times$Age$_{ijkt} = 4$				0.063^a				0.072^a
				(0.005)				(0.005)
$\widehat{v}\times$Age$_{ijkt} = 5$				0.062^a				0.064^a
				(0.006)				(0.006)
$\widehat{v}\times$Age$_{ijkt} = 6$				0.051^a				0.062^a
				(0.007)				(0.007)
$\widehat{v}\times$Age$_{ijkt} = 7+$				0.051^a				0.051^a
				(0.006)				(0.006)
Observations	2726474	2726474	2726474	2726474	2726474	2726474	2726474	2726474

Robust standard errors 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.

The results are qualitatively similar to our baseline estimates, but they differ quantitatively; the effect of age on firm belief's updating following demand shocks is slightly lower in Table 6. Similar results are found in the case of predictions 2 and 3 (Tables A.5 and A.6 in the online appendix).

While these results confirm the robustness of our findings to the measurement of age, we cannot directly infer from them whether and how accumulated learning is lost during periods of exit. In order to do so, we directly test whether firms update their belief in response to a new signal similarly after their first entry and subsequent re-entries on a given market, depending on the time elapsed since last exit. We expect a lower response of beliefs during re-entries whenever the firm keeps some stock of knowledge of its demand in the market.

We estimate:

$$\Delta\varepsilon^q_{ijkt} = \theta_1 \widehat{v}_{ijk,t-1} + \sum_{g=2}^{6} \alpha_h (\widehat{v}_{ijk,t-1} \times \text{GAP}^h_{ijkt}) + \sum_{g=1}^{G} \beta_h \text{GAP}^h_{ijkt} + \mathbf{FE}_{ijk} + u_{ijkt} \text{ if } S_{ijk,t-2} = 0 \tag{29}$$

where GAP^h_{ijkt} are dummies for re-entries on a market by number of years since last exit. We only focus on entrants, i.e. on firms which did not serve a particular market two years before (as we need to observe the demand shock in $t-1$). Put differently, we compare the responsiveness to demand signals of firms which re-enter after a period of x years to the responsiveness of first time entrants.

Table 7: Temporary exit and the learning process

Dep. var.:	$\Delta\varepsilon^q_{ijkt}$					
Gap (years of exit)	1	2	3	4	5	6
$\widehat{v}\times$Gap	-0.079[a]	-0.023	0.000	-0.011	0.177	0.452
	(0.022)	(0.036)	(0.053)	(0.093)	(0.153)	(0.280)

Robust standard errors in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Dummies by number of years since last exit and $\widehat{v}$ included alone but coefficient not reported. Observations: 133,776.

Table 7 shows that when re-entering a market after two or more years of exit, firms essentially behave like first time entrants. On the other hand, when their exit lasted only one year, the level of updating of re-entrants is lower (around 40% lower given that the unreported coefficient on the non-interacted $\widehat{v}$ is 0.21), suggesting that learning capital has not been completely lost In other words, the learning accumulated by the firm is not necessarily lost when exiting, but it depreciates extremely quickly during periods of exit. After only two years out of the market, firms react as if they had entirely forgotten their accumulated learning.

5.4 Discussion and robustness

This section discusses the sensitivity of our results to various modeling hypotheses and performs a number of additional robustness exercises.

CES demand. With alternative demand structures – e.g. quadratic (as in Melitz and Ottaviano, 2008), translog (Feenstra, 2003), or nested CES (combined with Cournot competition as in Atkeson and Burstein, 2008) –, markups become variable, which has two implications for our empirical strategy and results. First, prices now depend on local market conditions, i.e. the price equation (18) should feature a set of jkt fixed-effects. These can be easily included, and we will indeed show that this modification leaves our results largely unchanged.

Second, optimal quantities q_{ijkt} now depend on the firm's beliefs about both its future demand and its future markup. This "expected markup" is itself a function of local market conditions, firm-specific parameters such as productivity, and firm-market-specific variables such as size or market share. This means that, even after controlling for local market conditions and firm-product characteristics, the residuals ε^q_{ijkt} might not only reflect the firm's beliefs about demand but also the firm's expected markup on market jk. Similarly, ε^p_{ijkt} will confound demand factors (the firms belief and demand shock) and firms markups.

How does this affect the interpretation of our results? Let us assume that age has no effect on belief updating, contrary to what our model and in particular prediction 1 suggests. With most demand systems, the elasticity of demand is decreasing in market-specific firm size (or equivalently in productivity or market share) and large firms charge higher markups. In our case, this implies that ε^p_{ijkt} (respectively ε^q_{ijkt}) will be upward (resp. downward) biased for large firms. As a consequence, the demand shock $\widehat{v}$ obtained by regressing ε^p_{ijkt} on ε^q_{ijkt} (equation (23)) will be upward biased for large firms (we need "too large" $\widehat{v}$ to compensate "too low" ε^q_{ijkt}). Now, when regressing $\Delta\varepsilon^q_{ijkt}$ on $\widehat{v}$ (Table 3, column (1)), we impose that the coefficient on $\widehat{v}$ is the same across firms with different sizes: this coefficient will be underestimated for large firms, overestimated for small ones. When we further include the interaction term between age and $\widehat{v}$ (Table 3, column (2)), and given that age and size are positively correlated, we partly correct for this bias and expect accordingly the coefficient on the interaction term to be positive – the opposite of our result. Put differently, assuming a CES utility should bias the results against us, i.e. the coefficient on the interaction term between the demand shock and age should be upward biased, and the effect of learning underestimated.

Table 8 assesses the sensitivity of our results on prediction 1 to relaxing the CES assumption. Tables 11 and 12 in the appendix contain similar exercises applied to predictions 2 and 3. In columns (1) and (2) we simply re-estimate ε^p_{ijkt} and $\widehat{v}$, including $\mathbf{FE}_{jkt}$ in the price equation. Our results are very robust: firms positively update their beliefs following a positive demand shock, but less so as they gain experience. The coefficient on the interaction term between age and the demand shock is quantitatively close to our baseline estimates. In columns (3) and (4) we directly control in our estimations for lagged firm size (the log of total quantity[34] sold in market jk by firm i in $t-1$) and its interaction with the demand shock. As explained above, this aims at capturing directly the firms' position along the demand curve, and to correct for the fact that our estimated demand shocks might be correlated with size. In columns (5) and (6), we use the average quantity sold in market jk by firm i between years t and $t-1$ as an alternative measure of size.[35] As expected, the interaction term between size and the demand shock displays a positive coefficient. At the same time, the effect of age on belief updating gets slightly reinforced (columns (3) and (5)). This suggests that the CES assumption has a limited effect on our results.[36]

Finally, the estimations of Table 8 implicitly assume that the way in which markups vary with size is linear. We relax this assumption in Table 13 in the appendix, in which we replace our two size variables by size bins. These bins are constructed based on deciles of each size

[34] Controlling for the value of sales instead of quantity yields similar results.

[35] The idea is to control for the fact that age and size might be "too" positively correlated in the first two years because of the potential incompleteness of the first year of export measured over the calendar year (Berthou and Vicard, 2014; Bernard *et al.*, 2014).

[36] Note that the fact that our coefficients of interest are robust to controlling for firm size also makes it clear that our results are not caused by stochastic productivity growth. Indeed, in models with stochastic productivity (e.g. Luttmer, 2007), firm growth and age are uncorrelated if one conditions on firm size.

Table 8: Prediction 1: relaxing the CES assumption

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	$\Delta\varepsilon^{q}_{ijkt}$					
Age definition	# years since last entry (reset after 1 year of exit)					
Robustness	Controlling for FE$_{jkt}$ in prices		Controlling for FE$_{jkt}$ in prices and size			
			Size$_{ijk,t-1}$		$\overline{\text{Size}}_{ijk,t/t-1}$	
$\widehat{v}$	0.159^{a}		0.095^{a}		0.075^{a}	
	(0.005)		(0.005)		(0.005)	
Age$_{ijkt}$	-0.041^{a}		-0.013^{a}		-0.044^{a}	
	(0.000)		(0.000)		(0.000)	
$\widehat{v}\times$Age$_{ijkt}$	-0.008^{a}		-0.009^{a}		-0.013^{a}	
	(0.001)		(0.001)		(0.001)	
$\widehat{v}\times$Age$_{ijkt} = 2$		0.160^{a}		0.088^{a}		0.065^{a}
		(0.003)		(0.004)		(0.004)
$\widehat{v}\times$Age$_{ijkt} = 3$		0.118^{a}		0.048^{a}		0.013^{b}
		(0.004)		(0.006)		(0.006)
$\widehat{v}\times$Age$_{ijkt} = 4$		0.118^{a}		0.046^{a}		0.007
		(0.006)		(0.007)		(0.007)
$\widehat{v}\times$Age$_{ijkt} = 5$		0.111^{a}		0.038^{a}		-0.004
		(0.007)		(0.008)		(0.008)
$\widehat{v}\times$Age$_{ijkt} = 6$		0.098^{a}		0.024^{b}		-0.020^{b}
		(0.008)		(0.009)		(0.009)
$\widehat{v}\times$Age$_{ijkt} = 7+$		0.108^{a}		0.033^{a}		-0.014^{c}
		(0.007)		(0.008)		(0.008)
Size$_{ijk,t-1}$			-0.082^{a}	-0.081^{a}	0.010^{a}	0.011^{a}
			(0.000)	(0.000)	(0.000)	(0.000)
$\widehat{v}\times$Size$_{ijk,t-1}$			0.014^{a}	0.015^{a}	0.018^{a}	0.019^{a}
			(0.001)	(0.001)	(0.001)	(0.001)
Observations	2739927	2739927	2739927	2739927	2739927	2739927

Robust standard errors in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Size$_{ijk,t-1}$ is the log of the total quantity exported by firm i in product k, destination j in year $t-1$, and $\overline{\text{Size}}_{ijk,t/t-1}$ is the average quantity exported by firm i in market jk between t and $t-1$. Age dummies included alone in columns (2), (4) and (6) but coefficients not reported.

variables computed by HS4-destination-year[37] We use the same indicators of size as in Table 8. We lose a number of observations in the computation of the size bins but the effect of age on belief updating remains significant (columns (4) and (6)), especially when we control for size (columns (1) and (3)).

Fixed quantities. A second assumption we made is that quantities are fixed ex-ante, before the firm observes its idiosyncratic demand on each market. Prices, on the other hand,

[37]We chose to compute these deciles by 4-digit products rather than 6-digit for two reasons. First, because we need enough observations in each product-destination-year to be able to compute the deciles (using HS6 products indeed drops 40% of the observations). Second, because this allows firms to compete within an HS4 even if they produce different HS6.

take into account the realization of demand shocks. For our theoretical predictions to hold we only need quantities to adjust *less* than prices. Again, this hypothesis is important to identify separately the demand shock from the belief and therefore to test prediction 1, but bears no impact on the other predictions. The results shown in the previous section support this: the growth rates of quantities (and their variance) indeed decrease more with age than the growth rate of prices.

To gauge the importance of this assumption, we have re-run our estimations on sectors and destinations for which it is more likely that production is fixed *ex-ante*. We expect adjustment costs to be higher 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. 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). We restrict our sample to sectors or destinations with high adjustment costs, i.e. sectors/destinations belonging to the top 25% of the sample in terms of input complexity or time-top-ship. The results for prediction 1 are provided in Table 14 in the appendix (see also Tables A.3 and A.4 in the online appendix for results related to predictions 2 and 3). The adjustment of the firm's belief following a demand shock is quantitatively stronger than in our baseline estimates (columns (1) and (5)), as is the coefficient on the interaction term between demand shocks and age (columns (2)-(4) and (6)-(8)).

Passive versus active learning. We assume in our model – and therefore in the interpretation of our results – that learning is passive. Firms do not engage expenses in their early years to build a consumer base. How can we ensure that our results are indeed driven by passive learning of demand? First, as mentioned earlier, our methodology controls for all the firm-product specific supply side factors – such as investment or marketing expenses – which might impact the firm demand, as well as for all the destination characteristics which might impact the expenses of all French firms exporting a given product to a given market.[38]

Second, we can directly test for the presence of passive learning using a methodology initially proposed by Pakes and Ericson (1998) (see Abbring and Campbell, 2005 for an application). The general idea of the test is to discriminate between models with passive learning (as ours) and models with active learning in which firms can invest to accumulate demand (as in Ericson and Pakes, 1995) by regressing current firm size on its immediate past size and its initial size. The passive learning model imply that the firm initial size (more precisely, the firm's initial belief) will be useful to forecast the firms' belief and sales throughout its life, while the active learning model does not.

In Table 9, we regress the firm belief after x years, $x = 3, ..., 8$, on the 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.[39] Two results are worth mentioning. First, the

[38]Another argument in favor of our interpretation is related to the way in which prices vary with age. A way to accumulate demand is to price low in the first years in order to increase demand in the long-run (Foster *et al.*, 2013). This would imply that, purged from productivity and local demand conditions, prices of young firms are lower than prices of experienced exporters. We do not find evidence of this relationship when regressing ε_{ijkt} on firm age (results are available upon request).

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

Table 9: Passive versus active learning

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Prior_{ijkt}					
Age definition	# years since last entry (reset after 1 year of exit)					
Age	3	4	5	6	7	8
$\text{Belief}_{ijk,t-1}$	0.511^a	0.559^a	0.601^a	0.618^a	0.633^a	0.648^a
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
$\text{Belief}_{ijk,0}$	0.150^a	0.131^a	0.105^a	0.091^a	0.083^a	0.072^a
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	59425	59425	59425	59425	59425	59425

Robust standard errors in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

initial belief has 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. Both results are consistent with our passive learning model.

At this point, it is important to note that our results do not preclude the possibility that active learning is an important determinant of firm dynamics in general; they only suggest that we can interpret our results as strong evidence in favor of passive demand learning, and that, if active learning were important, it is largely accounted for by the various dimensions of fixed effects included in our estimations.

Measurement issues. In Tables A.7 and A.8 of the online appendix we perform two additional robustness checks. First, in Table A.7, columns (1) and (2), we replicate the results with equation (23) being estimated at the 4-digit (HS4) instead of 6-digit level. This in particular accounts for the fact that, due to the large number of 6-digits products, many categories contain very few observations, which leads to imprecise estimates.

Second, in Table A.7, columns (3) and (4) we check that our results are robust to the inclusion of an additional interaction term between firm age and our estimates of σ_k. This is to ensure that our results are not driven by heterogenous learning processes across sectors with different elasticities (as $\widehat{v}$ contains σ_k). In all cases, the results are extremely close to our baseline estimates shown in Table 3.

Finally, in Table A.8 of the online appendix we replicate our baseline estimates on extra EU-15 countries. We do so because small intra-EU transactions are potentially not recorded in the customs data, which might introduce noise in our measures of age and therefore lead to attenuation bias. Indeed, the estimated coefficients we obtain are quantitatively larger when we restrict our sample to extra-EU countries.

6 Firm survival

So far, we have only considered the dynamics of firm growth in the export markets conditional on firm survival. In this section, we extend our analysis to the exit decision for each firm-product-destination pair, in order to cover the entire post-entry firm dynamics.

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:[40]

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

Obviously, the exit decision does not only depend on this value but also on the expected future stream of profits, which depends on the evolution of C_{ikt}^{S}, C_{jkt}^{S}, $\widetilde{\theta}_{t-1}$ and $\widetilde{\sigma}_{t-1}^{2}$ over time. Our assumption of normal prior beliefs already provides the conditional distribution of $\widetilde{\theta}_t$ given $\widetilde{\theta}_{t-1}$, while the evolution of $\widetilde{\sigma}_{t-1}^{2}$ is deterministic: it monotonically decreases over time (see equations 4 and 5). The evolution of firm beliefs can therefore be summarized by $\widetilde{\theta}_{t-1}$ and t, the age of the firm. 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 need to introduce some (mild) assumptions on their dynamics. Using the fact that profits only depend on $C_{ikt}^{S} C_{jkt}^{S}$, 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}.[41]

With these additional assumptions, we can now assess the impact of firm beliefs updating on its exit decision. Note that the set of firm state variables at time t can be summarized by $\Omega_{ijkt} = \left\{A_{ijkt}, \widetilde{\theta}_{t-1}, t\right\}$. Thus, the Bellman equation is given by:

$$V_{ijk}\left(\Omega_{ijkt}\right) = \max\left\{\mathrm{E}\left[\pi_{ijkt}\left(\Omega_{ijkt}\right)\right] + \beta \mathrm{E}\left[V_{ijk}\left(\Omega_{ijkt+1} \mid \Omega_{ijkt}\right)\right], 0\right\} \tag{30}$$

where β is the rate at which firms discount profits and where we have normalized the value of exiting to zero.[42]

Under the assumptions made above, this Bellman equation has a unique solution V_{ijk}. Moreover, this solution is monotonically increasing in A_{ijkt} and $\widetilde{\theta}_{t-1}$.[43] Intuitively, the flow of future expected profits inherits the properties of expected profits at time t. It follows that there exists a threshold value $\underline{\widetilde{\theta}_{t-1}}(A_{ijkt}, t)$ such that a firm exits market jk at time t if $\widetilde{\theta}_{t-1} < \underline{\widetilde{\theta}_{t-1}}(A_{ijkt}, t)$. This implies the following prediction:

[40] See equations (2), (9) and (10) and the expression of beliefs after $t-1$ signals derived in the appendix.

[41] While not very demanding, these assumptions restrict the set of possible dynamics for firm productivity. In particular, the Markov assumption implies that we have to assume away a learning process behind the C_{ikt}^{S} and C_{jkt}^{S} terms. In that sense, our results on firm exit decision are somewhat weaker than those about firm growth, which are consistent with any dynamics of firm productivity.

[42] Here, we assume that an exiting firm loses all the information (learning) accumulated in the past. If the firm enters again market jk in the future, new initial beliefs will be drawn. We thus treat the exit decision as irreversible. Note that this assumption is supported by our results in Table 7.

[43] The proof of these two statements is almost identical to the proof of proposition 1 in Hopenhayn (1992) and of theorem 1 in Jovanovic (1982). However, our problem is slightly different than the one in Jovanovic (1982), as firm profits are not bounded. We thus need to show first that $\lim_{T\to\infty} \mathrm{E}\left[\sum_{t=1}^{T} \beta^t \mathrm{E}\left[\pi_t\right]\right]$ has at most an exponential growth. See appendix for details.

Prediction # 4 (firm exit): *Given A_{ijkt} and t (firm age), the probability to exit decreases with $\widetilde{\theta}_{t-1}$.*

To test this prediction, note that from equation (5), $\widetilde{\theta}_{t-1}$ can be expressed as:

$$\widetilde{\theta}_{t-1} = \left(\frac{\widetilde{\sigma}^2_{t-1}}{\sigma^2_\epsilon}\right) a_{ijkt-1} + \left(1 - \frac{\widetilde{\sigma}^2_{t-1}}{\sigma^2_\epsilon}\right) \widetilde{\theta}_{t-2} \tag{31}$$

where we used the fact that $g_{t-1} = \frac{\widetilde{\sigma}^2_{t-1}}{\sigma^2_\epsilon}$. $\widetilde{\theta}_{t-1}$ thus increases with $\widetilde{\theta}_{t-2}$ and a_{ijkt-1}. We therefore want to test if, conditional on A_{ijkt} and firm age, the probability to exit decreases with $\widetilde{\theta}_{t-2}$ and a_{ijkt-1}. While prediction 4 has been traditionally associated with learning in the literature, it has usually been tested showing that exit rates decline with firm size, sometimes conditional on age. We mainly depart from these papers because our identification strategy provides us with estimates of $\widetilde{\theta}_{t-2}$ and a_{ijkt-1}, thus allowing to test directly the impact of beliefs updating on the firm exit decision.

More formally, to test prediction 4 we estimate the following probabilistic model:

$$\begin{aligned} \Pr(S_{ijkt} > 0 | S_{ijk,t-1} = 1) &= 1 \text{ if } \alpha_1 \text{AGE}_{ijkt-1} + \alpha_2 \widehat{v}_{ijk,t} + \alpha_3 \varepsilon^q_{ijkt-1} + \mathbf{FE} + u_{ijkt} > 0 \\ &= 0 \text{ otherwise.} \end{aligned}$$

We expect α_2 and α_3 to be negative. $\mathbf{FE}$ 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, which might be important here given the two large dimensions of fixed effects we need to include.

The results are shown in Table 10, columns (1) to (3). These are largely consistent with the model's predictions: conditional on age, exit probability significantly decreases with positive demand shocks $\widehat{v}$ and with the firm's belief (columns (1) to (3)).

Interestingly, the literature has also usually associated learning with exit rate declining with age, and we indeed find this to be the case in our estimations. However, as discussed in Pakes and Ericson (1998), this prediction is not robust to the learning mechanism we put forward. Indeed, the decision to exit not only depends on the extent of firm updating (which indeed declines with age) but also on how $\underline{\widetilde{\theta}_{t-1}}(A_{ijkt}, t)$ evolves through time. If this threshold increases very rapidly for some t, the exit rate could actually be higher for older firms.

On the other hand, a clear prediction of our passive learning model is that negative demand shocks should trigger less exits for older firms. The reason is apparent in equation (31): firm posterior beliefs $\widetilde{\theta}_{t-1}$ depend less and less on demand shocks as firms age. Thus, the exit rate may not be decreasing with age, but demand shocks should 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 thus not be weakened for older firms (see Ericson and Pakes, 1995). This (discriminant) prediction is not directly tested in Pakes and Ericson (1998) because they use a much less parametric model than ours that

Table 10: Firm exit

	(1)	(2)	(3)	(4)	(5)
Dep. var.			$\Pr(S_{ijkt} > 0 \mid S_{ijk,t-1} = 1)$		
Age definition			# years since last entry (reset after 1 year of exit)		
$\text{Prior}_{ijk,t-1}$	-0.041^a		-0.041^a		-0.041^a
	(0.000)		(0.000)		(0.000)
$\text{Age}_{ijk,t-1}$	-0.034^a	-0.045^a	-0.033^a	-0.045^a	-0.033^a
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\widehat{v_{ijk,t-1}}$		-0.028^a	-0.031^a	-0.030^a	-0.042^a
		(0.000)	(0.000)	(0.000)	(0.000)
$\widehat{v_{ijk,t-1}} \times \text{Age}_{ijk,t-1}$				0.001^a	0.004^a
				(0.000)	(0.000)
Observations	8786242	8786242	8786242	8786242	8786242

Robust standard errors in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

prevent them to back out demand shocks and firm beliefs. Their test is solely based on actual firm size.

Columns (4) and (5) of Table 10 test for this prediction. We simply add to our baseline specification of column (3) an interaction term between age and demand shock in $t-1$.[44] We expect the coefficient on this interaction term to be positive: negative demand shocks lead to more exits in younger cohorts. Our results are consistent with the model: young firms react more to a given demand shocks than mature exporters on the market. In column (5), a 10% negative demand shock 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. Therefore, both the growth of beliefs and the lower responsiveness of older firms to demand shocks explain the decline of exit rates with age.

Overall, these results support the model predictions about firm exit. They also provide further evidence that we isolate a passive learning mechanism with our identification of the demand shocks and firm beliefs.

7 Conclusion

This paper has structurally assessed the empirical relevance of a model of export dynamics featuring local learning of demand, in the spirit of Jovanovic (1982). This model has three main predictions: (i) a new signal leads a firm to update more its belief, the younger the firm is; (ii) controlling for aggregate market conditions and firm productivity, the absolute values of firm-market specific growth rates of sales, quantities and prices decrease with age; (iii) within cohort, the variance of growth rates decreases with age. Using detailed exporter-level data

[44]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 8. 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 the coefficient on the interaction between $\widehat{v}_{ijk,t-1}$ and age might reflect differences in $\widehat{v}_{ijk,t-1}$ along the jkt dimension (as we focus on an interaction term in this case).

containing the prices and the quantities sold by French firms on export markets, we have shown that this model can be used to estimate firm-market specific demand shocks and prior beliefs about demand, and that its three predictions are strongly supported by the data. Importantly, our methodology and therefore our results are consistent with any possible dynamics of firm productivity.

Overall, the learning mechanism we uncover is quantitatively important: the growth of beliefs explains a larger part of the variance in the firm-market specific growth rates than supply side dynamics. Although the learning process appears to be especially strong in the first years after entry, even the most experienced firms in our sample still exhibit significant belief updating. Interestingly, we also provide evidence that the accumulated learning is quickly lost during exit periods: after exiting the market two years or more, firms essentially behave like a first-time entrant. A direct extension of our work would be to consider the – market, sector or firm-specific – determinants of learning speed.

Finally, we have considered the predictions of our model in terms of firm survival. When firm productivity follows a Markov process, the model predicts that given age, the probability to exit decreases with the firm belief and with the idiosyncratic demand shocks it faces. Further, a demand shock leads to more exit in younger cohorts than in older ones. Our empirical results again support these predictions.

The empirical relevance of firm learning has implications for the modeling of firm (and industry) dynamics in general. In particular, it underlines that firms' age is important to understand firms reaction to idiosyncratic demand shocks. Beyond idiosyncratic shocks, it also means that firms of different ages do not face the same amount of uncertainty, leading to a heterogeneous impact of firm responses to aggregate uncertainty shocks. This could refine the analysis of uncertainty shocks on aggregate outcomes, as for example developed in Bloom *et al.* (2012).

We concentrated on post-entry dynamics, leaving for future research the study of the impact of learning on entry decisions. The next step is to use our methodology to investigate how the differences in firms' initial size when entering a market can be explained by the firms' beliefs on other products they sell in the same market, on the same product they sell in other destinations or by other firms' beliefs serving the same market. This would allow to see how information spread over products, markets and firms.

References

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

Albornoz, F., Calvo Pardo, H. F., Corcos, G. and Ornelas, E. (2012), "Sequential exporting", *Journal of International Economics*, vol. 88 n° 1: pp. 17–31.

Arkolakis, C. (2010), "Market Penetration Costs and the New Consumers Margin in International Trade", *Journal of Political Economy*, vol. 118 n° 6: pp. 1151 – 1199.

Arkolakis, C. (2013), "A unified theory of firm selection and growth", manuscript, Yale University.

Atkeson, A. and Burstein, A. (2008), "Pricing to Market, Trade Costs, and International Relative Prices", *American Economic Review*, vol. Forthcoming.

Berman, N., de Sousa, J., Martin, P. and Mayer, T. (2013), "Time to Ship during Financial Crises", *NBER International Seminar on Macroeconomics*, vol. 9 n° 1: pp. 225 – 260.

Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K. (2009), "The Margins of US Trade", *American Economic Review*, vol. 99 n° 2: pp. 487–93.

Bernard, A. B., Massari, R., Reyes, J.-D. and Taglioni, D. (2014), "Exporter Dynamics, Firm Size and Growth, and Partial Year Effects", NBER Working Papers 19865, National Bureau of Economic Research, Inc.

Berthou, A. and Vicard, V. (2014), "Firms' export dynamics: experience vs. size", *World Economy*, vol. forthcoming.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I. and Terry, S. J. (2012), "Really Uncertain Business Cycles", NBER Working Papers 18245.

Bricongne, J.-C., Fontagné, 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 n° 1: pp. 134–146.

Broda, C. and Weinstein, D. (2006), "Globalization and the Gains from Variety", *The Quarterly Journal of Economics*, vol. 121 n° 2: pp. 541–585.

Cabral, L. and Mata, J. (2003), "On the Evolution of the Firm Size Distribution: Facts and Theory", *American Economic Review*, vol. 93 n° 4: pp. 1075–1090.

Caves, R. E. (1998), "Industrial Organization and New Findings on the Turnover and Mobility of Firms", *Journal of Economic Literature*, vol. 36 n° 4: pp. 1947–1982.

Cooley, T. F. and Quadrini, V. (2001), "Financial Markets and Firm Dynamics", *American Economic Review*, vol. 91 n° 5: pp. 1286–1310.

Defever, F., Heid, B. and Larch, M. (2011), "Spatial Exporters", CEP Discussion Papers, Centre for Economic Performance, LSE dp1100, Centre for Economic Performance, LSE.

Dunne, T., Roberts, M. J. and Samuelson, L. (1989), "The Growth and Failure of U.S. Manufacturing Plants", *The Quarterly Journal of Economics*, vol. 104 nº 4: pp. 671–98.

Eaton, J., Eslava, M., Kugler, M. and Tybout, J. (2007), "Export Dynamics in Colombia: Firm-Level Evidence", BANCO DE LA REPÚBLICA WP N.446.

Eaton, J., Kortum, S. and Kramarz, F. (2011), "An Anatomy of International Trade: Evidence from French Firms", *Econometrica*, vol. 79 nº 5: pp. 1453–1498.

Eaton, J., Eslava, M., Jinkins, D., Krizan, C. J. and Tybout, J. (2014), "A search and learning model of export dynamics", Manuscript.

Ericson, R. and Pakes, A. (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *Review of Economic Studies*, vol. 62 nº 1: pp. 53–82.

Evans, D. (1987), "Test of alternative theories of firm growth", *Journal of Policial Economy*, vol. 95 nº 4: pp. 657–74.

Feenstra, R. C. (2003), "A homothetic utility function for monopolistic competition models, without constant price elasticity", *Economics Letters*, vol. 78 nº 1: pp. 79–86.

Fernandes, A. and Tang, H. (2014), "Learning from Neighbors' Export Activities: Evidence from Exporters' Survival", forthcoming in the *Journal of International Economics*.

Foster, L., Haltiwanger, J. and Syverson, C. (2008), "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?", *American Economic Review*, vol. 98 nº 1: pp. 394–425.

Foster, L., Haltiwanger, J. and Syverson, C. (2013), "The slow growth of new plants: learning about demand?", Manuscript.

Haltiwanger, J., Jarmin, R. S. and Miranda, J. (2013), "Who Creates Jobs? Small versus Large versus Young", *The Review of Economics and Statistics*, vol. 95 nº 2: pp. 347–361.

Hopenhayn, H. A. (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", *Econometrica*, vol. 60 nº 5: pp. 1127–50.

Imbs, J. and Mejean, I. (2014), "Elasticity optimism", *American Economic Journal: Macroeconomics*, vol. forthcoming.

Impullitti, G., Irarrazabal, A. A. and Opromolla, L. D. (2013), "A theory of entry into and exit from export markets", *Journal of International Economics*, vol. 90 nº 1: pp. 75–90.

Jovanovic, B. (1982), "Selection and the Evolution of Industry", *Econometrica*, vol. 50 nº 3: pp. 649–70.

Klette, T. J. and Kortum, S. (2004), "Innovating Firms and Aggregate Innovation", *Journal of Political Economy*, vol. 112 n^{o} 5: pp. 986–1018.

Lawless, M. (2009), "Firm export dynamics and the geography of trade", *Journal of International Economics*, vol. 77 n^{o} 2: pp. 245–254.

Lawless, M. and Whelan, K. (2008), "Where Do Firms Export, How Much, and Why?", Working Papers, School Of Economics, University College Dublin 200821, School Of Economics, University College Dublin.

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

Luttmer, E. G. J. (2007), "Selection, Growth, and the Size Distribution of Firms", *The Quarterly Journal of Economics*, vol. 122 n^{o} 3: pp. 1103–1144.

Luttmer, E. G. J. (2011), "On the Mechanics of Firm Growth", *Review of Economic Studies*, vol. 78 n^{o} 3: pp. 1042–1068.

Mayer, T. and Ottaviano, G. (2008), "The Happy Few: The Internationalisation of European Firms", *Intereconomics: Review of European Economic Policy*, vol. 43 n^{o} 3: pp. 135–148.

Melitz, M. and Ottaviano, G. (2008), "Market Size, Trade, and Productivity", *Review of Economic Studies*, vol. 75 n^{o} 1: pp. 295–316.

Morales, E., Sheu, G. and Zahler, A. (2014), "Extended Gravity", mimeo, Princeton University.

Munch, J. R. and Nguyen, D. (2014), "Decomposing Firm-Level Sales Variation", *Journal of Economic Behavior and Organization*, vol. 106: pp. 317–334.

Nunn, N. (2007), "Relationship-Specificity, Incomplete Contracts, and the Pattern of Trade", *The Quarterly Journal of Economics*, vol. 122 n^{o} 2: pp. 569–600.

Pakes, A. and Ericson, R. (1998), "Empirical Implications of Alternative Models of Firm Dynamics", *Journal of Economic Theory*, vol. 79 n^{o} 1: pp. 1–45.

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

Romalis, J. (2007), "NAFTA's and CUSFTA's Impact on International Trade", *The Review of Economics and Statistics*, vol. 89 n^{o} 3: pp. 416–435.

Rossi-Hansberg, E. and Wright, M. L. J. (2007), "Establishment Size Dynamics in the Aggregate Economy", *American Economic Review*, vol. 97 n^{o} 5: pp. 1639–1666.

Timoshenko, O. A. (2012), "Product Switching in a Model of Learning", Working Papers 2012-10, The George Washington University, Institute for International Economic Policy.

A Appendix

A.1 Theory

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

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

The FOC writes:

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

And from the constraint, we get

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

Growth of firm's beliefs about expected demand (prior). First note that firm i has a prior about the demand shock given by $a_{ijkt} \sim N(\widetilde{\theta}_{t-1}, \widetilde{\sigma}_{t-1}^2 + \sigma_\varepsilon^2)$ and thus $e^{\frac{a_{ijkt}}{\sigma_k}} \sim LN(\frac{\widetilde{\theta}_{t-1}}{\sigma_k}, \frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{\sigma_k^2})$. It follows that $\int \left(e^{\frac{a_{ijkt}}{\sigma_k}}\right) dG_{t-1}(a_{ijkt}) = e^{\frac{1}{\sigma_k}\left(\widetilde{\theta}_{t-1}+\frac{\widetilde{\sigma}_{t-1}^2+\sigma_\varepsilon^2}{2\sigma_k}\right)}$. We get the expression in the text:

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

Using the definition of $\Delta\widetilde{\theta}_t$, $\widetilde{\sigma}_{t-1}^2$ and $\widetilde{\sigma}_t^2$ (see (3) and (4)), we further get:

$$\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] = \frac{1}{\sigma_k\left(\frac{\sigma_\epsilon^2}{\sigma_0^2}+t\right)}\left(a_{ijkt} - \frac{\left(\theta_0 + \frac{\sigma_0^2}{2\sigma_k} + \overline{a}_{t-1}\frac{\sigma_0^2}{\sigma_\epsilon^2}(t-1)\right)}{\left(1+\frac{\sigma_0^2}{\sigma_\epsilon^2}(t-1)\right)}\right) \tag{32}$$

Proof of proposition 1. Prediction 1 states that following a new signal, updating is larger for younger firms. Updating is measured directly by $\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right]$ in (32). We get:

$$\frac{\partial \left(\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \right)}{\partial a_{ijkt}} = \frac{1}{\sigma_k \left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t \right)} \equiv \frac{g_t}{\sigma_k} > 0$$

The larger the demand shock, the larger the updating. However, the denominator increases with t: updating is larger for younger firms. This higher updating can be directly measured by g_t. It may also be of interest to note that updating decreases with uncertainty, i.e. σ_ϵ^2, as the signal is less informative when uncertainty is higher.

Proof of proposition 2. Proposition 2 states that expected absolute value of growth rates decrease with age. Growth rates are given by:

$$\Delta \ln Z^q_{ijk,t+1} = \sigma_k \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{33}$$

$$\Delta \ln Z^p_{ijk,t+1} = \frac{1}{\sigma_k} \Delta a_{ijkt+1} - \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{34}$$

$$\Delta \ln Z^S_{ijk,t+1} = (\sigma_k - 1) \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] + \frac{1}{\sigma_k} \Delta a_{ijkt+1} \tag{35}$$

First, note that a_{ijkt+1} and a_{ijkt} being drawn from the same distribution, $\mathrm{E}\left[\Delta a_{ijkt+1}\right] = 0$. The growth rates thus only depend on $\Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right]$.

Second, using (32) and the fact that $\mathrm{E}\left[a_{ijkt}\right] = \overline{a}_{t-1}$, the absolute value of the expected growth rate of $\mathrm{E}_{t-1} \left[e^{\frac{a_{ijkt}}{\sigma_k}} \right]$ is given by:

$$\mathrm{E}\left[\left| \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \right| \right] = \frac{\left| \left(\overline{a}_{t-1} - \theta_0 - \frac{\sigma_0^2}{2\sigma_k} \right) \right|}{\sigma_k \left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t \right) \left(1 + \frac{\sigma_0^2}{\sigma_\epsilon^2} (t-1) \right)}$$

The numerator, in absolute value, is necessarily positive and independent of age. The denominator is positive and strictly decreasing in age. And we have:

$$\begin{aligned}
\mathrm{E}\left[\left| \Delta \ln Z^q_{ijk,t+1} \right| \right] &= \sigma_k \mathrm{E}\left[\left| \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \right| \right] \\
\mathrm{E}\left[\left| \Delta \ln Z^p_{ijk,t+1} \right| \right] &= \mathrm{E}\left[\left| \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \right| \right] \\
\mathrm{E}\left[\left| \Delta \ln Z^S_{ijk,t+1} \right| \right] &= (\sigma_k - 1) \mathrm{E}\left[\left| \Delta \ln \mathrm{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \right| \right]
\end{aligned}$$

Which completes the proof. Note that the growth rates of quantities should decrease relatively faster than the one of prices.

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

$$
\begin{aligned}
\mathrm{V}\left[\Delta \ln Z^q_{ijk,t+1}\right] &= \sigma_k^2 \mathrm{V}\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) && (36)\\
\mathrm{V}\left[\Delta \ln Z^p_{ijk,t+1}\right] &= \left(\frac{1}{\sigma_k}\right)^2 \mathrm{V}\left(\Delta a_{ijkt+1}\right) + \mathrm{V}\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) \\
&\quad - \frac{2}{\sigma_k} cov\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right) && (37)\\
\mathrm{V}\left[\Delta \ln Z^S_{ijk,t+1}\right] &= (\sigma_k - 1)^2 \mathrm{V}\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) + \left(\frac{1}{\sigma_k}\right)^2 \mathrm{V}\left(\Delta a_{ijkt+1}\right) && (38)\\
&\quad + \frac{2(\sigma_k - 1)}{\sigma_k} cov\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right) && (39)
\end{aligned}
$$

First, a_{ijkt+1} and a_{ijkt} being drawn from the same distribution, $\mathrm{V}\left[\Delta a_{ijkt+1}\right] = 2\sigma_\epsilon^2$. Second, using (32), it is straightforward to show that:

$$
\mathrm{V}\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) = \left(\frac{\sigma_\epsilon}{\sigma_k\left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t\right)}\right)^2
$$

Finally, using the fact that $\mathrm{E}\left[\Delta a_{ijkt+1}\right] = 0$, we have:

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

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

$$
cov\left(\Delta \ln \mathrm{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right) = -\frac{\sigma_\epsilon^2}{\sigma_k\left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t\right)}
$$

Plugging terms into (**??**), and after simplification, we get:

$$
\begin{aligned}
\mathrm{V}\left[\Delta \ln Z^q_{ijk,t+1}\right] &= \left(\frac{\sigma_\epsilon}{\left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t\right)}\right)^2 && (40)\\
\mathrm{V}\left[\Delta \ln Z^p_{ijk,t+1}\right] &= \left(\frac{\sigma_\epsilon}{\sigma_k}\right)^2 \left(\left(\frac{1}{\left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t\right)} + 1\right)^2 + 1\right) && (41)\\
\mathrm{V}\left[\Delta \ln Z^S_{ijk,t+1}\right] &= \left(\frac{\sigma_\epsilon}{\sigma_k}\right)^2 \left(\left(\frac{(\sigma_k - 1)}{\left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t\right)} - 1\right)^2 + 1\right) && (42)
\end{aligned}
$$

The variance of quantities and prices strictly decrease with cohort age. It is also the case for sales if $\sigma_k > 1 + \frac{\sigma_\epsilon^2}{\sigma_0^2} + t$.

Existence, uniqueness and variation of V_{ijk}.

TBC

A.2 Additional results

Table 11: Prediction 2: robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\Delta\varepsilon^S_{ijkt}$		$\Delta\varepsilon^p_{ijkt}$		$\Delta\varepsilon^q_{ijkt}$		$\Delta\varepsilon^p_{ijkt}$	
Age definition	# years since last entry (reset after 1 year of exit)							
Robustness	Export sales		Controlling for FE_{jkt} in prices		Control for size		Controlling for FE_{jkt} in prices and size	
Age_{ijkt}	-0.037^a		-0.024^a		-0.032^a		-0.014^a	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\text{Age}_{ijkt} = 3$		-0.068^a		-0.053^a		-0.058^a		-0.030^a
		(0.001)		(0.001)		(0.001)		(0.001)
$\text{Age}_{ijkt} = 3$		-0.111^a		-0.079^a		-0.094^a		-0.047^a
		(0.002)		(0.001)		(0.002)		(0.001)
$\text{Age}_{ijkt} = 5$		-0.141^a		-0.096^a		-0.121^a		-0.057^a
		(0.002)		(0.001)		(0.002)		(0.001)
$\text{Age}_{ijkt} = 6$		-0.167^a		-0.109^a		-0.149^a		-0.065^a
		(0.002)		(0.001)		(0.002)		(0.001)
$\text{Age}_{ijkt} = 7+$		-0.201^a		-0.129^a		-0.175^a		-0.078^a
		(0.002)		(0.001)		(0.002)		(0.001)
$\text{Size}_{ijk,t-1}$					-0.023^a	-0.023^a	-0.029^a	-0.028^a
					(0.000)	(0.000)	(0.000)	(0.000)
Observations	2795979	2795979	2795979	2795979	2795979	2795979	2795979	2795979

Robust standard errors in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Controlling for year dummies does not affect the results. $\text{Size}_{ijk,t-1}$ is the log of the total quantity exported by firm i in product k, destination j in year $t-1$.

Table 12: Prediction 3: robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\text{Var}(\Delta\varepsilon^{S}_{ijkt})$		$\text{Var}(\Delta\varepsilon^{p}_{ijkt})$		$\text{Var}(\Delta\varepsilon^{q}_{ijkt})$		$\text{Var}(\Delta\varepsilon^{p}_{ijkt})$	
Age definition	# years since last entry (reset after 1 year of exit)							
Robustness	Export sales		Controlling for FE_{jkt} in prices		Control for size		Controlling for FE_{jkt} in prices and size	
Age_{cjkt}	-0.064^{a}		-0.032^{a}		-0.065^{a}		-0.031^{a}	
	(0.001)		(0.001)		(0.001)		(0.001)	
$\text{Age}_{cjkt} = 3$		-0.121^{a}		-0.069^{a}		-0.123^{a}		-0.065^{a}
		(0.003)		(0.002)		(0.004)		(0.002)
$\text{Age}_{cjkt} = 4$		-0.195^{a}		-0.104^{a}		-0.200^{a}		-0.099^{a}
		(0.004)		(0.002)		(0.004)		(0.002)
$\text{Age}_{cjkt} = 5$		-0.256^{a}		-0.130^{a}		-0.262^{a}		-0.125^{a}
		(0.005)		(0.003)		(0.005)		(0.003)
$\text{Age}_{cjkt} = 6$		-0.300^{a}		-0.149^{a}		-0.305^{a}		-0.143^{a}
		(0.005)		(0.003)		(0.006)		(0.003)
$\text{Age}_{cjkt} = 7+$		-0.357^{a}		-0.180^{a}		-0.366^{a}		-0.174^{a}
		(0.005)		(0.003)		(0.006)		(0.003)
$\text{Size}_{cjk,t-1}$					-0.020^{a}	-0.014^{a}	-0.012^{a}	-0.008^{a}
					(0.002)	(0.002)	(0.001)	(0.001)
Observations	598821	598821	598821	598821	598821	598821	598821	598821
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered by cohort in parentheses. Cohort fixed effects included in all estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. [1] firms present all years. $\text{Size}_{cjk,t-1}$ is the log of the average total quantity exported by the firms in cohort cjk in year $t-1$.

Table 13: Prediction 1: controlling for size bins

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	$\Delta\varepsilon^q_{ijkt}$					
Age definition	# years since last entry (reset after 1 year of exit)					
Size variable	$\text{Size}_{ijk,t-1}$			$\overline{\text{Size}}_{ijk,t/t-1}$		
Size dummies	Yes	Yes	No	Yes	Yes	No
Age_{ijkt}	-0.005^a		-0.042^a	-0.052^a		-0.053^a
	(0.000)		(0.000)	(0.000)		(0.000)
$\widehat{v}\times\text{Age}_{ijkt}$	-0.005^a		-0.004^a	-0.011^a		-0.006^a
	(0.001)		(0.001)	(0.001)		(0.001)
$\widehat{v}\times\text{Age}_{ijkt}=2$		0.365^a			0.296^a	
		(0.008)			(0.008)	
$\widehat{v}\times\text{Age}_{ijkt}=3$		0.339^a			0.247^a	
		(0.008)			(0.009)	
$\widehat{v}\times\text{Age}_{ijkt}=4$		0.340^a			0.237^a	
		(0.009)			(0.009)	
$\widehat{v}\times\text{Age}_{ijkt}=5$		0.330^a			0.231^a	
		(0.010)			(0.010)	
$\widehat{v}\times\text{Age}_{ijkt}=6$		0.318^a			0.218^a	
		(0.011)			(0.012)	
$\widehat{v}\times\text{Age}_{ijkt}=7+$		0.335^a			0.229^a	
		(0.009)			(0.010)	
Observations	2327572	2327572	2327572	1951476	1951476	1951476

Robust standard errors in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. $\text{Size}_{ijk,t-1}$ is the log of the total quantity exported by firm i in product k, destination j in year $t-1$, and $\overline{\text{Size}}_{ijk,t/t-1}$ is the average quantity exported by firm i in market jk between t and $t-1$. Estimations (1), (2), (4) and (5) include size dummies (and their interactions with $\widehat{v}$) constructed according to deciles of the variable, deciles being computed by HS4-product-destination-year. Age dummies included alone in columns (2), (4) and (6) but coefficients not reported.

Table 14: Prediction 1: robustness (high production adjustment costs)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.		$\Delta\varepsilon^q_{ijkt}$				$\Delta\varepsilon^q_{ijkt}$		
Age definition	# years since last entry (reset after 1 year of exit)							
Sample	Complex goods				Large time-to-ship			
$\widehat{v}$	0.091^a	0.138^a	0.138^a		0.162^a	0.231^a	0.231^a	
	(0.004)	(0.008)	(0.008)		(0.004)	(0.008)	(0.008)	
Age$_{ijkt}$		-0.038^a	-0.038^a			-0.035^a	-0.035^a	
		(0.001)	(0.001)			(0.001)	(0.001)	
$\widehat{v}\times$Age$_{ijkt}$		-0.013^a	-0.013^a			-0.022^a	-0.022^a	
		(0.002)	(0.002)			(0.002)	(0.002)	
$\widehat{v}\times$Age$_{ijkt} = 2$				0.126^a				0.198^a
				(0.006)				(0.005)
$\widehat{v}\times$Age$_{ijkt} = 3$				0.079^a				0.145^a
				(0.008)				(0.008)
$\widehat{v}\times$Age$_{ijkt} = 4$				0.066^a				0.134^a
				(0.011)				(0.010)
$\widehat{v}\times$Age$_{ijkt} = 5$				0.072^a				0.096^a
				(0.013)				(0.013)
$\widehat{v}\times$Age$_{ijkt} = 6$				0.044^a				0.097^a
				(0.016)				(0.017)
$\widehat{v}\times$Age$_{ijkt} = 7+$				0.050^a				0.093^a
				(0.014)				(0.016)
Observations	582450	582450	582450	582450	546586	546586	546586	546586

Robust standard errors 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) and (8) but coefficients not reported. Complex goods and large time-to-ship means in the last quartile of the variable.