Demand learning and firm dynamics: evidence from exporters[*]

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

First version: November 2014

This version: November 26, 2015

Abstract: This paper provides evidence that learning about demand is an important driver of firms' dynamics. We present a model of Bayesian learning in which firms are uncertain about their idiosyncratic demand in each of the markets they serve, and update their beliefs as noisy information arrives. Firms are predicted to update more their beliefs the younger they are. Guided by the model, we use exporter-level data to identify separately the idiosyncratic demand shocks and the firms' beliefs about future demand. The learning process appears stronger for younger firms and weaker in more uncertain environments. Further, accumulated knowledge decays during exit periods. The updating process generates a decline in growth rate with age conditional on size. Firm exit behavior is also consistent with the theory: the exit probability decreases with the firms' beliefs and the demand shocks the firm faces, and demand shocks trigger more exit in younger cohorts.

Keywords: firm growth, belief updating, demand, exports, uncertainty.

JEL classification: D83, F14, L11.

[*]We thank Richard Baldwin, Giuseppe Berlingieri, Patrick Blanchenay, Thomas Chaney, Simon Fuchs, Fabien Gensbittel, Joe Francois, Sacha Kapoor, Erzo Luttmer, Thierry Mayer, Marti Mestieri, Rahul Mukherjee, Franck Portier, Jim Tybout and seminar participants at the Geneva Trade and Development Workshop, Universidad de Navarra, Tilburg University, Banque de France, the VIth Villars Research Workshop on International Trade, LETC, and Compnet conference at ECB for very useful discussions and comments. The authors gratefully acknowledge financial support from Fondation Banque de France. The opinions expressed in this paper are those of the authors and do not reflect the views of the Banque de France. This paper features an online appendix containing additional results and available on the authors' webpages.

[†]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 producers rapidly expand after entry, many others do not survive the first few years. After some time however, those surviving firms account for a large share of sales on both domestic or foreign markets (Haltiwanger *et al.*, 2013; Bernard *et al.*, 2009; Eaton *et al.*, 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 have larger exit rates. For those that survive, the average growth of their sales declines with their age.[2] Similar behaviors have been recently reported for sales on foreign markets.[3] These facts can be rationalized by several theories, relying on different underlying mechanisms, such as stochastic productivity growth, endogenous R&D investment, financial constraints or demand learning. Yet, empirically, 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, using detailed exporter-level data, that it is an important driver of post-entry firm dynamics.

We first present a simple 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 in each period. These signals determine the firms' posterior beliefs about demand, from which they make their quantity decision. It follows that a new signal leads younger firms to update more their beliefs. The first contribution of the paper is to test this core prediction, which is specific to the learning mechanism.

To do so, we derive from the theory a methodology which allows to separately identify the firms' beliefs and the demand shocks (the signals) they face each period, in each of the markets 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 market-specific conditions and firm-specific supply side dynamics (e.g. productivity) from quantities and prices. This is made possible by a unique feature of international trade data, in which we can observe the 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 same product in the same destination allows to control for aggregate market-specific conditions. Second, we use the fact that, in the model, quantity decisions only depend on the firms' beliefs while prices also depend on the realized demand shocks, to

[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 (2015).

separate out the firms' beliefs from the demand signal. Therefore, while requiring few, standard assumptions, our methodology allows to directly test predictions which relate the evolution of firms' beliefs to firm age, in contrast to the literature which has typically looked at the correlation between firm size and firm age.

We find strong support for the core prediction of the model: belief updating following demand shocks is stronger for younger firms, with age being defined at the firm-product-destination level. The learning process appears to be especially strong in the first years after entry on a product-destination market, although even the most experienced firms in our sample still exhibit significant updating. Further, using a variety of indicators of market-specific uncertainty, we find that, as theoretically predicted, the updating process is significantly weakened and thus less dependent on age in more uncertain environments. We provide several robustness exercises to show that these results are not driven by our main modeling assumptions. Our findings survive after controlling for firm size, and are extremely stable across alternative samples and specifications. We also use different definitions of age to account for the fact that exporters enter and exit markets frequently and that the accumulated knowledge about demand might be partially kept even during periods of exit. We show that the bulk of accumulated knowledge is lost during periods of exit exceeding one year.

Consistent with the well documented age dependence of firm growth, our model predicts that, conditional on size, growth rates are higher for younger firms. This comes from the fact that growth rates should be more volatile for young firms as a consequence of their larger belief updating. This is supported by the data: we find that both the absolute value of the mean growth rate of firms' beliefs and its variance within cohorts decrease with age. Combined with firm selection – provided that exit probability does not increase too much with age[5] – this larger variability generates a negative relationship between age and growth, even conditioning on size.

Finally, we show that the exit behavior of the firms in our sample is also consistent with the learning model: the exit rate decreases with firms' beliefs and the demand shocks the firm face, and demand shocks trigger more exit in younger cohorts.

Our paper shows that demand learning is an important characteristic of the micro-dynamics of firms in narrowly defined markets. 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 applied to demand to study exporters' dynamics (Timoshenko, 2015; Fernandes and Tang, 2014; Albornoz *et al.*, 2012; Eaton *et al.*, 2014).

An alternative class of models explains firm and exporter dynamics through supply side mechanisms, including variations in productivity (through stochastic shocks or endogenous decisions) or financial constraints.[6] Both the theories based on demand learning and on supply side dynamics can replicate qualitatively most of stylized facts that we observe in the data. But the literature strikingly lacks direct empirical evidence of the relative relevance of these various al-

[5]As discussed later, the effect of age on survival is theoretically ambiguous. In the data however, exit rates sharply decline with age.

[6]See for instance Hopenhayn (1992), Luttmer (2007), Arkolakis (2013), Impullitti *et al.* (2013) for models with stochastic shocks to productivity, Klette and Kortum (2004) or Rossi-Hansberg and Wright (2007) for theories of endogenous productivity growth, and Clementi and Hopenhayn (2006) for a model with financial constraints.

ternative mechanisms. A major contribution of our paper is to properly isolate the idiosyncratic demand component of firms sales, and therefore to get estimates and results which cannot be driven by standard alternative supply side explanations put forward in the literature to explain firm dynamics.

The literature has also proposed other demand-side mechanisms than learning over a constant demand parameter. Demand could fluctuate over time in an exogenous manner, or could be affected by firm investments, search for new consumers or pricing policy.[7] These elements may be important to explain the dynamics of firms, but they are unlikely to drive our results, as we *de facto* control for all variations in firm-specific expenditures and product-market specific conditions. Moreover, we provide direct support for our interpretation using a test initially proposed by Pakes and Ericson (1998) to discriminate between models of "passive" learning *à la* Jovanovic (1982) and models of "active" learning where firms may engage in specific investments. The idea is to regress current firm beliefs on immediate past beliefs and initial beliefs. Consistent with the passive learning model, we find that initial beliefs are useful to forecast future firms' beliefs throughout their life. Another possibility is to look at the firms' pricing policy. Firms could accumulate customers by setting low prices in their first years. This mechanism has been recently put forward by Foster *et al.* (2013),[8] who find that this process explains a large part of the relationship between firm age and firm size using a panel of US homogenous goods producers. In our data however, once purged from their productivity component, firm-market specific prices are (slightly) decreasing with age, as predicted by our model. Our empirical methodology is close in spirit to Foster *et al.* (2013, 2008), in that they also separate idiosyncratic demand shocks from firms' productivity, but our paper differs in several ways. In particular, we do not need to measure productivity or other firm-specific determinants of sales to identify demand shocks, and we are able to separate the beliefs from the demand shocks to test the learning mechanism.

We assume that the actual sales of a firm in a given product-destination market are the only source of information about demand. In other words, we assume away information spillovers. Firm beliefs in a given market might well be affected by its beliefs on other destinations (Albornoz *et al.*, 2012) or on other products for the same destination (Timoshenko, 2015). These effects might be stronger for similar destinations and products (Morales *et al.*, 2014; Defever *et al.*, 2015; Lawless, 2009). The behavior of other firms serving the same market might also play a role (Fernandes and Tang, 2014). Studying the relative importance of these various potential sources of information is an interesting and vast question in itself, that we indeed plan to study in the future, but which is beyond the scope of this paper. We focus on the way in which firms update their beliefs based solely on their actual sales, and therefore concentrate on their post-entry dynamics, keeping their prior beliefs at the time of entry as exogenous.[9]

The empirical relevance of demand learning has several implications. The first and most direct one is that models trying to explain the dynamics of firm size distribution (within and across industries) based solely on productivity growth would gain at introducing demand learning

[7] See Luttmer (2011), Ericson and Pakes (1995); Pakes and Ericson (1998), Eaton *et al.* (2014) or Foster *et al.* (2013).

[8] See also Gourio and Rudanko (2014).

[9] Our focus therefore differs from 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 of firms belonging to the Chinese ceramic industry, but focuses on entry.

mechanisms. Second, an interesting property of the learning process is that it generates a form of hysteresis. The most experienced firms, having gathered more information about their demand, are less sensitive to demand shocks in terms of sales and exit decisions. This also suggests that aggregate uncertainty shocks should have heterogeneous effects across industries, depending on their age structure. Moreover, our results put forward a new source of irreversibility in the exit decision as accumulated knowledge quickly decays during periods out of the market. Finally, our results tend to justify the implementation of policies supporting start-ups, as recently discussed by Arkolakis *et al.* (2015). In the specific case of trade, it suggests that export promotion would gain at targeting not only entrants but also young exporters.

The paper proceeds as follows. In the next section, we describe our data and provide descriptive evidence on firm's post-entry growth and survival. In section 3 we present our model and its implication for firms' beliefs updating with respect to age. Section 4 describes our identification strategy and section 5 our main results as well as a number of robustness exercises. In section 6 we present additional results on firm growth and survival. The last section concludes.

2 Firm dynamics on foreign markets and export growth

This section describes our data and presents statistics about the dynamics of French firms in their export markets. In particular, we emphasize the role of young firms and new destination markets on aggregate growth, and provide suggestive evidence that demand shocks contribute to a large part of the variance of firms' growth 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 that we will use as a proxy for firm price in the second part of the paper.[10]

A product is defined at the 6-digit level (HS6). We focus on the subset of HS6 product categories that remain stable over the time-period in order to be able to track firms over time on a specific market (destination-and-product).[11] We concentrate on the years 1996-2005 because we use the first two years, 1994 and 1995, to identify entry, as explained in more details in section 4.2. Our final dataset covers exports of 4,183 HS6 product categories to 180 destination markets by 100,690 firms over the period 1996-2005.

2.2 Stylized facts

Contribution to aggregate sales growth. Recent literature has emphasized the essential contribution of young firms to industry dynamics, either in terms of aggregate output, employment

[10] 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.

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

or trade. Haltiwanger *et al.* (2013) show for instance that US start-ups display substantially higher rates of job creation and destruction in their first ten years, and that these firms represent a large share of total employment after a decade of existence. These patterns are also found for other countries (see Criscuolo *et al.*, 2014 for evidence on 18 OECD countries; Lawless, 2014 on Irish firms, Ayyagari *et al.*, 2011 for developing countries). Similar facts characterize trade dynamics: Eaton *et al.* (2007) and Bernard *et al.* (2009) show that exporters start small but that, conditional on survival, they account for large shares of total export growth after a few years.

Our exporter-level data exhibit comparable features. Over the 1996-2005 period, we find that, on average, new firm-destination-product triplets represent only 12.3% of total export value after a year, but their share reaches 53.5% after a decade (27.3% due to new markets served by incumbents and 26.2% by new firms exporting, see Table 1). The contribution of the extensive margin to aggregate exports is determined by three components of firm dynamics: entry, survival and post entry growth on new markets. Since new exporters typically do not survive more than a few years in export markets,[12] firm selection and growth are important drivers of aggregate trade growth over longer horizons, besides the size at entry. Column (2) of Table 1 shows that pure growth after entry accounts for around 40% of the end-of-period share of newly created firm-destination-product triplets. The objective of this paper is precisely to understand how learning about demand can explain this post-entry dynamics.

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

	Average yoy 1996/2005	Overall 1996/2005
New firms	2.4%	26.2%
Initial size	-	*16.5%*
Growth since entry	-	*9.7%*
New product-destination	9.9%	27.3%
Initial size	-	*16.1%*
Growth since entry	-	*11.3%*
Incumbent firm-product-destination	87.7%	46.5%
Total	100%	100%

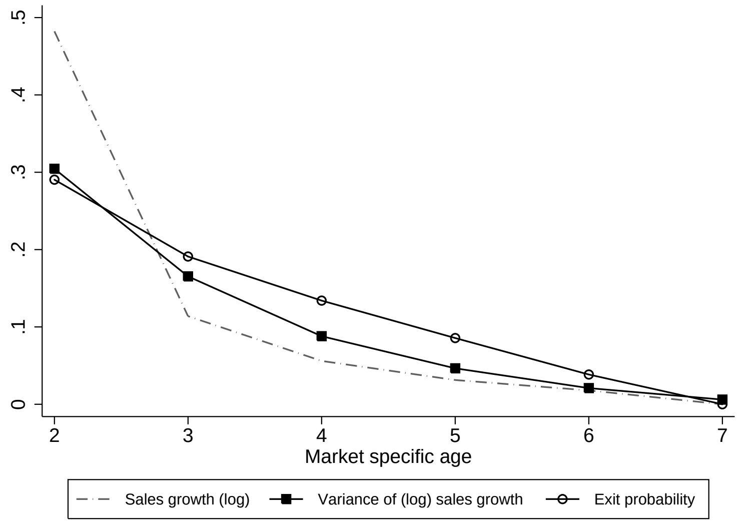
Note: sample of HS6 fixed over time. Source: French Customs. yoy: year-over-year.

Sales growth and age. Consistent with evidence on domestic firms (Evans, 1987; Dunne *et al.*, 1989), the dynamics of firms' sales in each foreign market they serve is characterized by a number of stylized facts: sales growth, the volatility of sales growth and the probability of exit all decrease sharply with age. Such systematic patterns are illustrated in Figure 1, which plots the coefficients obtained from regressing different variables reflecting firm dynamics (either firm-market growth rates, or the variance of growth rates within cohorts, or exit probability) on age dummies, controlling for firm size and time dummies. Note that throughout the paper, we define age as firm-market-specific tenure, i.e. as the number of years since the last entry of a

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

firm in a specific product-destination (we will also show that our results are not sensitive to using alternative definitions – see section 4.2).

Figure 1: Impact of firm-market specific age conditional on size



Note: this figure plots the coefficients obtained from of a regression of the log change of firm sales (respectively variance of firms' sales and exit) on age bins, firm size and year and sector dummies (see table A.2 in online appendix A). All coefficients are relative to the omitted category, age of seven years or more. The variance of firms' sales growth is measured within cohorts of firms on a product-destination market. A similar pattern is obtain when controlling for country-and-sector fixed effects.

The negative relationship between the growth of the value of sales and age *conditional on size* cannot be rationalized in models based on supply side dynamics only – either stochastic shocks to productivity (Hopenhayn, 1992; Luttmer, 2007; Arkolakis, 2013, Impullitti *et al.*, 2013), endogenous productivity variations (Klette and Kortum, 2004; Rossi-Hansberg and Wright, 2007), financial constraints (Clementi and Hopenhayn, 2006) or adjustment costs – as in such models firm size is a sufficient statistic for firm growth. In contrast, such age dependence is in line with different mechanisms of demand driven firm dynamics – passive (Ruhl and Willis, 2014) or active consumer base accumulation (Foster *et al.*, 2013, Eaton *et al.*, 2014). The joint larger volatility of sales growth in younger cohorts illustrated in Figure 1 is however specific to models including a learning mechanism.

Contribution to firm sales variations. Before turning to the model in detail, we provide some additional suggestive evidence that firm-market specific demand-side factors are key drivers of growth by decomposing the variance of post-entry sales growth. We perform an exercise which is similar in spirit to Eaton *et al.* (2011).[13] We first regress firm-market specific sales growth on a set of destination-product-time dummies.[14] The R^2 of such a regression is 0.12: market-specific

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

[14]Table A.1 in the online appendix summarizes the results.

dynamics play a limited role. Adding firm-product-time fixed effects increases the R^2 to 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 on this part of firm dynamics, with the objective of understanding the extent to which it is consistent with firms learning about their demand. Anticipating a bit on our results, we will find that this R^2 jumps to 0.87 when we include our estimates of the growth of firms' beliefs about demand, which we will interpret as suggestive that learning about demand is at least as important as supply side dynamics in explaining the growth of firm sales.

3 A model of firm growth with demand learning

In this section we present a standard model of international trade with Dixit-Stiglitz monopolistic competition and demand learning in the spirit of Jovanovic (1982) (see also Timoshenko, 2015) . 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.

3.1 Economic environment

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

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

with β 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. 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 may also be subject to learning. In that case, the firm would take a quantity decision based on its beliefs about its costs. As we will not back out learning from firms' productivity,[15] 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 – we come back to this assumption at the end of section 4.1.

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

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

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, its prior beliefs about $\overline{a}_{ijk}$ are normally distributed with mean θ_{ijk0} and variance σ^2_{ijk0}. 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^2_{ε}. According to Bayes' rule, the firm's posterior beliefs about $\overline{a}_{ijk}$ after t signals are normally distributed with mean $\widetilde{\theta}_{ijkt}$ and variance $\widetilde{\sigma}^2_{ijkt}$, where:

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

$$\widetilde{\sigma}^2_{ijkt} = \frac{1}{\frac{1}{\sigma^2_{ijk0}} + \frac{t}{\sigma^2_{\epsilon}}} \tag{4}$$

and $\overline{a}_{ijkt}$ is the average signal value, $\overline{a}_{ijkt} = \left(\frac{1}{t}\sum_t a_{ijkt}\right)$. Note that contrary to $\widetilde{\theta}_{ijkt}$, the posterior variance $\widetilde{\sigma}^2_{ijkt}$ does not depend on the realizations of the signals and decreases only with the number of signals (i.e. learning reduces uncertainty). The posterior variance is thus always smaller than the prior variance, $\widetilde{\sigma}^2_{ijkt} < \widetilde{\sigma}^2_{ijkt-1}$. Given that we do not formally model entry, in the rest of the paper we omit the subscripts on $\{\theta_{ijk0}, \sigma^2_{ijk0}, \widetilde{\sigma}^2_{ijkt}\}$ and label these variables $\{\theta_0, \sigma^2_0, \widetilde{\sigma}^2_t\}$ to simplify notations.

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

$$\Delta\widetilde{\theta}_{ijkt} = g_t\left(a_{ijkt} - \widetilde{\theta}_{ijkt-1}\right) \text{ with } g_t = \frac{1}{\frac{\sigma^2_{\epsilon}}{\sigma^2_0} + t}. \tag{5}$$

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

[15] 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.

3.2 Firm size and belief updating

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

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

Here, we assume for simplicity that aggregate market conditions at time t, i.e. $\mu_k Y_{jt}/P_{jkt}^{1-\sigma_k}$, are observed by firms before making their quantity decision. This leads to the following optimal quantities and prices (see appendix A.1):[16]

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

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

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

As firm i makes a quantity decision before observing demand for its product, q^*_{ijkt} depends on expected demand, not on demand realization, contrary to p^*_{ijkt}.

The literature has typically computed correlations between firm age and firm growth rates, and attributed negative ones as potential evidence for a learning mechanism. Indeed, as we formally show in section 6, the fact that younger firms adjust more their beliefs leads growth rate to decrease with age in absolute value. But of course, as is clear from equations (7) and (8), firm size, and therefore firm growth (would it be measured in terms of employment or sales) also depend on the evolution of market-specific conditions and firm productivity, which could be correlated with firm age. Directly testing for the presence of demand learning thus requires either making assumptions about the dynamics of aggregate market conditions and firm productivity or finding a way to account for them. Our methodology follows the second route.

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

[16] Firm size could alternatively be measured by firm sales: $S^*_{ijkt} = q^*_{ijkt} p^*_{ijkt}$. Assessing the impact of firm demand learning on quantities and prices implicitly also provides its impact on sales.

expressions for quantities and prices as:[17]

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

As just underlined, the impact of demand learning is fully included in the Z^q_{ijkt} and Z^p_{ijkt} terms. These terms can be understood as the quantity and price of firm i for product k on market j at time t, purged from firm unit costs and aggregate market conditions, and may be very different from the actual firm size and firm price. From a methodological point of view, we stress that any prediction about firm demand learning should be based on these Z_{ijkt} terms rather than the actual q^*_{ijkt} and p^*_{ijkt}. This also means that we will not look at the dynamics of firm size (at least per se), but directly at the dynamics of the firms' beliefs about demand $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$.

The growth rate of $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$ can be expressed as:

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

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

Prediction # 1 (updating): *A new signal a_{ijkt} leads to a larger updating of the belief $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$, the younger the firm is.*

It is also interesting to note that larger uncertainty (i.e. a higher σ^2_ϵ) reduces the extent of belief updating and the effect of age on belief updating. This is because a signal is less informative when uncertainty is higher (see the proof of proposition 1 in the appendix).

In order to directly test this mechanism of firm updating, we need to identify the demand shock a_{ijkt} and the firm's belief $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$ in each period. In the next section, we provide a methodology to isolate the Z^q_{ijkt} and Z^p_{ijkt} terms, which then allows to distinguish the beliefs from the demand shock components.

4 Identification and measurement

4.1 Identifying beliefs and demand shocks

In order to isolate the Z^q_{ijkt} and Z^p_{ijkt} terms, we need to purge supply side and market specific

[17] Prices do not actually depend on aggregate market conditions, due to our CES assumption. Since other utility functions could make prices depend on market specific conditions (and in particular on market size), we will systematically check the robustness of our results to the inclusion of market-specific conditions in the price equation as well.

factors from actual quantities and prices. This is achieved by estimating the following quantity and price equations in logs:[18]

$$\begin{aligned} \ln q_{ijkt} &= \mathbf{FE}_{ikt} + \mathbf{FE}_{jkt} + \varepsilon^q_{ijkt} && (12) \\ \ln p_{ijkt} &= \mathbf{FE}_{ikt} + \varepsilon^p_{ijkt} && (13) \end{aligned}$$

where k is a 6-digit product and t is a year. $\mathbf{FE}_{ikt}$ and $\mathbf{FE}_{jkt}$ represent respectively firm-product-year and destination-product-year fixed effects. 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}, where S_{ijkt} denote firms sales, to proxy them.

Given that we control for all time-varying, market- and firm-product-specific determinants of quantities and prices, the residuals ε^q_{ijkt} and ε^p_{ijkt} are by construction orthogonal to the standard supply side determinants of firm dynamics (i.e. productivity and market conditions). Our approach could therefore accommodate any underlying dynamic process for the ikt and jkt terms. This include processes driving the evolution of firm productivity, but also any other time-varying, firm-specific factors that might affect firm dynamics such as financial constraints for example.

To be more specific, the estimates of ε^q_{ijkt} and ε^p_{ijkt} are estimates of the Z_{ijkt} terms. Using equations (7) and (8), we get:

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

Testing prediction 1, which is the essence of the learning mechanism, requires getting estimates of both the firms's beliefs about expected demand $\mathbb{E}_{t-1}[e^{\frac{a_{ijkt}}{\sigma_k}}]$ and the demand shock a_{ijkt}. As the firm takes its quantity decision before observing the demand realization, $\ln Z^q_{ijkt}$ depends on the firms' beliefs about demand only, while $\ln Z^p_{ijkt}$ is adjusted for the demand shock.[19] Thus, the residual ε^q_{ijkt} provides a direct estimate of the firms' beliefs. 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 (15) and (14), we get:

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

We estimate (16) by 6-digit product to allow σ_k to differ across products[20] and obtain[21]:

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

[19] We discuss further this assumption and provide a number of robustness checks in section 5.6 and in appendix A.1.2.

[20] 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 (16) is estimated at the 4-digit level.

[21] 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. Note that our results are insensitive to such cleaning of the data.

$$\widehat{\beta} = -\frac{1}{\sigma_k} \quad \text{and} \quad \widehat{v}_{ijkt} = \frac{1}{\sigma_k} a_{ijkt}. \tag{17}$$

This identification strategy is possible to implement because we are able to observe the sales of the same product by the same firm in different destination markets. 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 $\mathbf{FE}_{ikt}$.[22] It may also be of interest to note that our methodology does not make use of the time dimension to identify the firm beliefs. However, it assumes the $\overline{a}_{ijk}$ to be independent across markets, products or firms, i.e. there is no information spillovers.

Following the model, we interpret the residuals from equations (12) and (13) as reflecting the demand-side components of prices and quantities. Our identification assumption is that, within a given firm, costs can differ across products but not across products *and* markets. Note however that we allow variations in costs across markets for a given product. These include in particular trade costs and potential differences in demand for quality and are captured by $\mathbf{FE}_{jkt}$. Yet, we cannot totally exclude the possibility that some product-market specific costs remain. But we consider as unlikely the possibility that firms *learn* about these product-market-specific costs, which would translate into larger revisions of ε^q_{ijkt} for younger firms. This is what we find in the data, and this comforts us in our demand-side interpretation.

4.2 Measuring firm-product-destination specific age

The last variable we need to compute to be able to test our prediction is the market-specific firm age. A major advantage of exporter-level data is that it features a substantial amount of entries and exits, and allows measuring precisely and tracking over time firms' sales in 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 during periods of exits. Our model being 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 in a product-destination. We assume complete depreciation of firm specific knowledge during exit periods and reset the age to zero whenever the firm exits at least one calendar year from a specific product-destination. Age is 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 their knowledge about local demand when they exit, regardless of the number of exit years; this third age variable is simply the number of exporting years since the first entry of the firm. Note that in all the empirical analysis, to ensure

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

the consistency of our measures of age, we 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 in 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.

5 Main results

In this section, we start by providing some descriptive statistics of our final sample, before discussing the results obtained when testing prediction 1. We then turn to additional insights related to the characteristics of the learning process. We finally discuss the sensitivity of our results to our main modeling assumptions and to several measurement issues.

5.1 Sample statistics

Table 2 contains some descriptive statistics about our final sample.

Table 2: Sample statistics

	Obs.	Mean	S.D.	Q1	Median	Q3
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
$\Delta\varepsilon^q_{ijkt}$	2726474	0.03	1.37	-0.74	0.02	0.80
$\Delta\varepsilon^p_{ijkt}$	2726474	-0.00	0.68	-0.24	-0.00	0.24
$\widehat{v}_{ijkt}$	2726474	-0.00	0.58	-0.25	0.00	0.24
σ_k	2675182	11.15	8.07	5.81	8.10	13.94
Age^1_{ijkt}	2726474	3.48	1.78	2	3	4
Age^2_{ijkt}	2726474	3.65	1.84	2	3	5
Age^3_{ijkt}	2726474	3.73	1.84	2	3	5

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 3.5 and 3.7 years depending on the definition (note that since we focus on $\Delta\varepsilon_{ijkt}$ in the following, firms that exit during the first year are dropped and 2 is the minimum value that our age variable can take). This is evidence of the low survival rates observed during the first years a firm serves a particular market, a topic we shall specifically study in the last section of the paper.

Over the period, the firm-market specific beliefs have been characterized by a slightly positive growth,[23] while $\Delta\varepsilon^p_{ijkt}$ exhibit a slightly negative average growth (-0.0002). We will however show in section 5.5 that prices significantly decrease with age. Note that quantitatively, the evolution of beliefs ε^q_{ijkt} is crucial in explaining firms' dynamics. Including the $\Delta\varepsilon^q_{ijkt}$ as an

[23] 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 the second year, because of the potential incompleteness of the first year of export measured over the calendar year. When measuring age by bins as in our estimations, the dummy for year two gets rid of this average bias. Table A.3 of the online appendix shows that our results on prediction 1 are robust to the use of reconstructed years beginning the month of first entry at the firm-product-destination level.

explanatory variable of the growth of sales in the estimation performed in section 2.2 increases the R^2 to 0.87, compared to 0.44 when firm-product-time and product-destination-time fixed effects are included alone. Interpreting $\Delta\varepsilon^q_{ijkt}$ as estimates of beliefs' reflecting mostly demand-side variations, this implies that demand learning contributes at least as much as supply side factors to the explanation of the variance of firms' sales in specific markets.

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 the ones found at similar levels of disaggregation by the literature, using very different methodologies and data. 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.[24] Our estimates of σ_k also follow expected patterns: considering Rauch (1999) classification, the median (resp. mean) across products is 8.5 (resp. 11.1) for differentiated goods, 10.9 (resp. 14.3) for referenced priced goods and 14.7 (resp. 16.8) for goods classified as homogenous.[25] These means and medians of σ_k are statistically different across the three groups.

5.2 Baseline results

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:

$$\Delta\varepsilon^q_{ijkt+1} = \alpha_0 + \alpha_1(\frac{1}{\sigma_k}a_{ijkt}) + u_{ijkt} = \alpha_0 + \alpha_1\widehat{v}_{ijkt} + u_{ijkt} \tag{18}$$

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:

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

where AGE^g_{ijkt} are dummies taking the value 1 for each age category $g = 2, ...7+$ representing the number of years of presence in the export market (e.g. $g = 2$ in the second year of presence). In both cases standard errors are robust to heteroscedasticity and clustered by firm. We expect the α_g to be decreasing with age g. Note that our model predicts that $\alpha_g = g_t = \frac{1}{\sigma^2_\varepsilon/\sigma^2_0+t}$. g_t measures 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 of Table 3 considers the effect of demand

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

[25] Note that these numbers are slightly higher than the means and medians displayed in Table 2 because they are computed across products, while the statistics in Table 2 are based on our final sample, i.e. also reflect the number of French firms selling each product.

Table 3: Prediction 1: demand shocks and beliefs updating

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\Delta\varepsilon^q_{ijkt+1}$							
Age definition	# years since last entry (reset after 1 year of exit)							
$\widehat{v}_{ijkt}$	0.075^a	0.109^a	0.109^a		0.050^a	0.143^a		0.077^a
	(0.009)	(0.009)	(0.004)		(0.012)	(0.016)		(0.020)
Age$_{ijkt}$		-0.040^a	-0.040^a			-0.080^a		
		(0.001)	(0.000)			(0.001)		
$\widehat{v}_{ijkt}\times$Age$_{ijkt}$		-0.009^a	-0.009^a			-0.007^a		
		(0.001)	(0.001)			(0.002)		
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=2$				0.103^a	0.054^a		0.135^a	0.059^a
				(0.009)	(0.009)		(0.014)	(0.013)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=3$				0.066^a	0.016^c		0.141^a	0.064^a
				(0.009)	(0.010)		(0.018)	(0.013)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=4$				0.057^a	0.007		0.111^a	0.035^a
				(0.010)	(0.010)		(0.019)	(0.013)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=5$				0.056^a	0.007		0.102^a	0.025^c
				(0.014)	(0.013)		(0.022)	(0.015)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=6$				0.047^a	-0.003		0.088^a	0.011
				(0.013)	(0.012)		(0.020)	(0.014)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=7+$				0.050^a			0.077^a	
				(0.012)			(0.020)	
Observations	2726474	2726474	2726474	2726474	2726474	2726474	2726474	2726474
Firm×Destination×Product FE	No	No	No	No	No	Yes	Yes	Yes

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3)). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Age dummies included alone in columns (4), (5), (7) and (8) but coefficients not reported. Columns (6) to (8) include firm-product-destination fixed effects. Columns (5) and (8) are the same as columns (4) and (7) except that coefficients are estimated relative to the baseline omitted category, age of seven years or more.

shocks on the adjustment of the firms' beliefs (equation (18)). Columns (2) to (5) study how this effect varies with age (equation (19)). 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 (column (1)). This adjustment is indeed significantly larger when firms are young (columns (2)-(5)). Including age linearly (column (2)) or through bins (columns (4) and (5)) leads to the same conclusion. Bootstrapping the standard errors (column (3)) also leaves the results unaffected. Interestingly, after 7 consecutive years of presence on a market, firms still significantly update their belief to demand shocks, but the extent of belief updating is 50 percent smaller than after entry. Note that the shape of the learning process seems in line with our assumption of normal priors: age has a stronger effect in the early years. In column (5) where coefficients are computed relative to the benchmark category, age of seven years or more, updating is statistically stronger at age 2 and 3 (and the coefficient is significantly larger on age 2 than on age 3), but after four years the coefficients are not statistically different from each other. Note however that in most of our robustness specifications, the coefficients are also statistically different at older ages, as shown later.

Our results might be affected by unobservable characteristics of the firms that impact both their belief updating and their survival probability. The firms remaining in the sample the entire period might have different initial beliefs and better prior knowledge of the market, which would imply that we actually identify a between-firm effect rather than a decline in the extent of belief updating within each firm as predicted by the model. In columns (6) to (8) we thus include firm-product-destination fixed effects and therefore identify the coefficients on the within dimension of our data. The shape of the learning process remains similar (column (7)), and the coefficients of the different age bins in column (8) are now statistically different from each other until age 6.

5.3 Learning and market uncertainty

The model predicts that a higher uncertainty in the market should slowdown the belief updating process: a signal is less informative when uncertainty is higher. It follows that the speed at which firms update their beliefs should decrease with age, but less so when uncertainty is larger (see proof of prediction 1 in the appendix).

We test these predictions using two alternative types of measures of uncertainty obtained from external sources. This is also a way to validate our empirical strategy. The first set of indicators are from Baker and Bloom (2013) and Bloom (2014). We use alternatively the stock-market index volatility and the average exchange rate volatility. Both are time-varying, country-wide measures and, as in Bloom (2014), we standardize them by country. The second type of measure we use is market-specific. We use bilateral trade values by country-pair and 6-digit HS product from the BACI database of the CEPII. We compute the standard deviation of the log of imports by importer and product over the period 1997-2005.

The results are provided in Table 4. We consider sequentially our three measures of uncertainty. Odd numbered columns add to our baseline specification an interaction term between $\widehat{v}_{ijkt}$ and the considered uncertainty variable. Even numbered columns also include a triple interaction term between age, $\widehat{v}_{ijkt}$ and uncertainty. Standard errors are clustered at the dimension of the uncertainty variable (destination-year or destination-product). Although the significance levels vary depending on the measure used, our results are globally supportive of the dampening role of uncertainty on belief updating (odd columns). On the other hand, the coefficient on the interaction term between age and the demand shocks is virtually unaffected. Quantitatively, the role of uncertainty is non negligible. A standard deviation increase from the mean of the level of uncertainty decreases the response of beliefs to demand shocks from 0.089 to 0.078 in column (1), and from 0.134 to 0.118 in column (5).

When uncertainty is large, gaining experience has a lower effect on belief updating, as shown by the coefficient of the triple interaction term in even columns. Another way to represent this result is shown in Figure 2. We divide the foreign markets into high and low uncertainty markets defined according to the sample median of the uncertainty variable and run our baseline specification (column (4) of Table 3) separately on each of the two sub-samples. Figure 2 depicts the extent of belief updating for each sample, by age category, for our market-specific uncertainty

Table 4: Prediction 1: the role of uncertainty

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.			$\Delta\varepsilon^{q}_{ijkt+1}$			
Uncertainty variable	Stock market vol. index		Exch. rate volatility		Imports volatility	
$\widehat{v}_{ijkt}$	0.089^a	0.089^a	0.085^a	0.086^a	0.173^a	0.279^a
	(0.013)	(0.013)	(0.013)	(0.013)	(0.011)	(0.022)
Uncertainty	-0.013^a	0.011	-0.015^a	-0.024	0.021^a	0.057^a
	(0.003)	(0.007)	(0.006)	(0.015)	(0.001)	(0.004)
Age$_{ijkt}$	-0.042^a	-0.043^a	-0.041^a	-0.041^a	-0.040^a	-0.014^a
	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}$	-0.007^a	-0.006^a	-0.006^b	-0.006^a	-0.009^a	-0.043^a
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.006)
$\widehat{v}_{ijkt}\times$Uncertainty	-0.011^c	-0.019	-0.015	-0.049^b	-0.026^a	-0.069^a
	(0.007)	(0.013)	(0.012)	(0.022)	(0.005)	(0.009)
Age$_{ijkt}\times$Uncertainty		-0.006^a		0.003		-0.010^a
		(0.002)		(0.005)		(0.001)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}\times$Uncertainty		0.002		0.011^b		0.014^a
		(0.002)		(0.005)		(0.002)
Observations	1485990	1485990	1438098	1438098	2704301	2704301

Robust standard errors clustered by destination-year in columns (1) to (6), by destination-product in columns (7) and (8) in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Measures of uncertainty: the stock-market index volatility from Bloom (2014) in columns (1) and (2); average cross-sectional firms stock-returns dispersion from Bloom (2014) in columns (3) and (4); average exchange rate volatility from Bloom (2014) in columns (5) and (6); standard deviation of the log of imports by importer and product from BACI in columns (7) and (8).

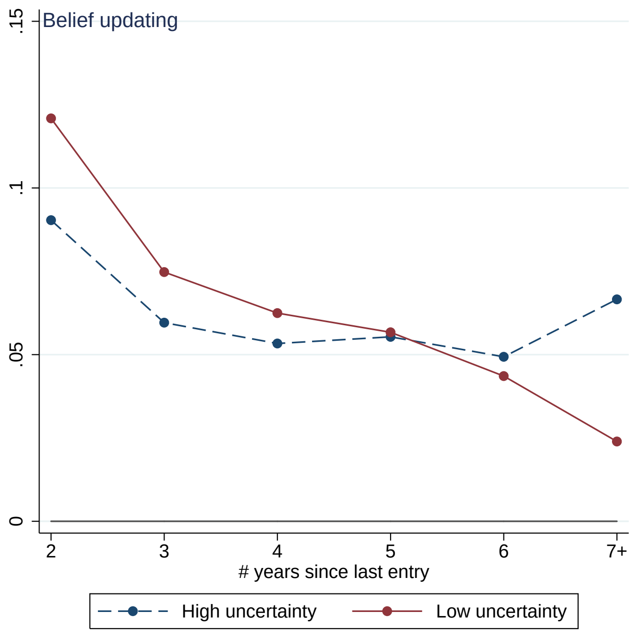
proxy from BACI. Belief updating sharply decreases with age in the least uncertain markets, while the relationship is much flatter on the high uncertainty sample. The complete set of results for all three proxies is relegated to the online appendix section B. Although the estimates are logically less precise in the case of the country-level proxies, a similar pattern emerges.

5.4 Learning and forgetting

How fast accumulated knowledge about demand depreciates when the firm exits the market? So far we have treated each entry into a market as a new one: age was reset to zero in case of exit. Table 12 in the appendix tests the robustness of our results 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 assumption that all experience is kept during exit periods, whatever the length of these periods. The results are qualitatively similar to our baseline estimates, but they differ quantitatively; the effect of age on firms' beliefs updating following demand shocks is slightly lower in Table 12.

While this confirms the robustness of our findings to the measurement of age, we cannot directly infer from them whether and how accumulated knowledge is lost during periods of exit. In order to do so, we directly test whether firms update their beliefs in response to a new signal

Figure 2: Uncertainty and belief updating



This figure is obtained by estimating the specification of column (4) of Table 3 on two sub-samples defined according to the sample median of the uncertainty measure. The market-specific uncertainty measure used here. The figure plots the coefficients of the $\widehat{v}_{ijkt}$ variable for each age category.

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:

$$\Delta\varepsilon^q_{ijkt+1} = \theta_1\widehat{v}_{ijkt} + \sum_{g=2}^{6}\alpha_h(\widehat{v}_{ijkt} \times \text{GAP}^h_{ijkt}) + \sum_{g=1}^{G}\beta_h\text{GAP}^h_{ijkt} + \mathbf{FE}_{ijk} + u_{ijkt} \text{ if } S_{ijk,t-1} = 0 \quad (20)$$

where GAP^h_{ijkt} are dummies for re-entries in 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 5 shows that when re-entering a market after two or more years of exit, firms essentially behave like first time entrants. However, 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 knowledge accumulated by the firm is not necessarily lost when exiting, but it depreciates quickly during periods of exit. After only two years out of the market, firms react as if they had entirely forgotten the information accumulated in the past.

Table 5: Temporary exit and the learning process

Dep. var.:			$\Delta\varepsilon^{q}_{ijkt+1}$			
Gap (years of exit)	1	2	3	4	5	6
$\widehat{v}_{ijkt}\times$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 clustered by firm-product-destination 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.

5.5 Discussion: learning about a constant demand parameter?

In our model, firms learn about their idiosyncratic demand parameter, which is assumed to be constant over time. Alternative demand-side mechanisms could however generate some (stochastic) trend in the demand addressed to firms. Firms could engage in an active and costly search for new buyers, or in specific investments to increase their profitability. They might also price low in their first years to build a consumer base, or demand could simply evolve over time.[26] We show in this section that our results are unlikely to be affected by these elements.

As already mentioned, our identification strategy controls for all the firm-product specific supply side factors – such as firm investment or marketing expenses – which could impact product demand across markets. It also captures all market-specific characteristics, which include in particular trends in product demand that are country-specific, as well as the expenditures of all French firms exporting a given product to a given market. Still, we cannot rule out a priori the possibility that our beliefs estimates also include a time-varying component. Two additional tests provide evidence supporting our mechanism of learning about a constant demand parameter.

The first relates to the way in which firms' 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, the prices of young firms should be lower than those of experienced exporters. In our passive learning model on the other hand, Z^{q}_{ijkt} should increase over time due to selection (firms facing negative demand shocks exit; see section 6.2) while the evolution of Z^{p}_{ijkt} should be of opposite sign, i.e. Z^{p}_{ijkt} should diminish over time, especially in early years (see equations (14) and (15)). The intuition behind this result is simply that survivors have faced on average more positive demand shocks and thus adjusted upwards their prices, leading to prices above their optimal pricing rule in the first years. Because these predictions are generated by composition effects triggered by selection, these age profiles of quantities and prices should become flatter when controlling for firm-destination-product fixed effects. Table 6 shows that the results of regressing ε^{q}_{ijkt} and ε^{p}_{ijkt} on firm age are indeed in line with the model. ε^{p}_{ijkt} decreases with age, although this effect is quantitatively limited. This was expected as changes in the firm beliefs are supposed to affect more ΔZ^{q}_{ijkt} than ΔZ^{p}_{ijkt} (by a factor σ_k). When controlling for fixed effects in columns (3) and (6), the increase

[26] See Eaton *et al.* (2014), Ericson and Pakes (1995), Foster *et al.* (2013), and Luttmer (2011).

Table 6: Dynamics of quantity and prices

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.		ε^{q}_{ijkt}			ε^{p}_{ijkt}	
Age definition			# years since last entry (reset after 1 year of exit)			
Age_{ijkt}	0.103^{a}			-0.008^{a}		
	(0.001)			(0.001)		
$\text{Age}_{ijkt} = 2$		0.202^{a}	0.165^{a}		-0.019^{a}	-0.003^{a}
		(0.002)	(0.003)		(0.001)	(0.001)
$\text{Age}_{ijkt} = 3$		0.318^{a}	0.207^{a}		-0.027^{a}	-0.006^{a}
		(0.004)	(0.004)		(0.001)	(0.002)
$\text{Age}_{ijkt} = 4$		0.402^{a}	0.224^{a}		-0.035^{a}	-0.007^{a}
		(0.005)	(0.005)		(0.002)	(0.002)
$\text{Age}_{ijkt} = 5$		0.464^{a}	0.227^{a}		-0.037^{a}	-0.005^{b}
		(0.006)	(0.007)		(0.002)	(0.002)
$\text{Age}_{ijkt} = 6$		0.514^{a}	0.224^{a}		-0.040^{a}	-0.006^{c}
		(0.008)	(0.009)		(0.003)	(0.003)
$\text{Age}_{ijkt} = 7+$		0.600^{a}	0.221^{a}		-0.045^{a}	-0.006
		(0.009)	(0.011)		(0.004)	(0.005)
Observations	6472999	6472999	6472999	6472999	6472999	6472999
Firm×Destination×Product FE	No	No	Yes	No	No	Yes

Standard errors clustered by firm in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

in ε^{q}_{ijkt} is only observed in the very first years, while ε^{p}_{ijkt} becomes roughly constant over time.

Second, we can directly test the stability of the demand parameter using a methodology proposed by Pakes and Ericson (1998) (see Abbring and Campbell, 2005 for an application). Paraphrasing them, the aim of the test is to discriminate between models with “passive learning” (with a constant parameter, as ours) and models with “active learning”, where firms may invest to increase their sales. Following Pakes and Ericson (1998) we regress current firm beliefs on its immediate past beliefs and its initial, prior beliefs. The passive learning model implies that the firm initial size (more precisely in our case, the firm’s initial beliefs) will be useful to forecast the firms’ beliefs and sales throughout their life, while the active learning model does not. The idea of this test is thus to determine whether $\overline{a}_{ijk}$ can be considered as constant over time, would it be due to firms actions or for other reasons. In Table 7, we regress the firm’s 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.[27] Two results are worth mentioning. First, the 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 assumption on $\overline{a_{ijk}}$.

At this point, we would like to stress that these results do not preclude the existence of

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

Table 7: Passive versus active learning

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	ε^q_{ijkt}					
Age definition	# years since last entry (reset after 1 year of exit)					
Age	3	4	5	6	7	8
ε^q_{ijkt-1}	0.511^a	0.559^a	0.601^a	0.618^a	0.633^a	0.648^a
	(0.006)	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)
ε^q_{ijk0}	0.150^a	0.131^a	0.105^a	0.091^a	0.083^a	0.072^a
	(0.005)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	59425	59425	59425	59425	59425	59425

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

additional processes that could magnify quantity dynamics. Consumer preferences evolving over time and/or firms actions (e.g. investments, consumers search, etc.) may matter for firms demand and thus firm dynamics in general, on top of the learning mechanism put forward here. However, the results presented in this section suggest that these elements are largely accounted for by the various dimensions of fixed effects included in our estimations.

5.6 Further robustness checks

We discuss here the sensitivity of our results to our main modeling hypotheses and perform some robustness exercises. We start with an important assumption that we borrow from Jovanovic (1982): firms make their quantity decision before observing the demand realization. We next show that our assumption of CES preferences has no important quantitative impact on our estimates. We finish by discussing some measurement issues.

Fixed quantities. We have assumed that quantities are set before firms observe their idiosyncratic demand in each market. Prices, on the other hand, fully adjust to take into account the realization of demand shocks. The results of the next section will indeed support this assumption: the growth rates of quantities (and their variance) decrease more with age than the growth rate of prices. Section A.1.2 of the appendix contains a complete discussion of the implications of relaxing our hypothesis of fixed quantities. We summarize it here, and refer the reader to the appendix for more details.

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

Second, we consider in the appendix an intermediate case with partial quantity adjustment in which firms can revise their quantity decision after observing part of the demand shock. In this case, our theoretical predictions still hold, but our identification of the demand shock is affected: ε^q_{ijkt} now also captures part of the demand shock and therefore becomes a noisy measure of the firm's belief. This has the same consequence as a classical measurement error: our demand shocks would be overestimated in absolute terms, which implies a potential downward bias in the estimation of the coefficient of belief updating α_1 from equation (18). Yet, unless this downward bias is correlated with age, our main results that young firms update more their beliefs should not be affected.

One way to gauge the importance of the fixed quantities assumption is to focus on sectors or destinations for which quantities are more likely to be rigid – i.e. those for which the demand shocks are more likely to be correctly estimated – and to compare the results with our baseline estimates of Table 3. We expect less quantity adjustment for complex goods (in which many different relationship-specific inputs are used in the production process) and in destinations characterized by longer time-to-ship. In Table 8 we restrict our sample to sectors or destinations belonging to the top 25% of the sample in terms of input complexity or 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). The updating of the firms' beliefs following a demand shock is similar or slightly stronger quantitatively than in our baseline estimates (columns (1) and (4)), which indeed suggests some limited attenuation bias. However, the coefficient on the interaction term between demand shocks and age (columns (2)-(3) and (5)-(6)), if anything, is slightly more negative than in our baseline estimates. If an attenuation bias was driving our results because of its positive correlation with age, we would expect instead the relationship between age and belief updating to be flattened in these regressions.

An alternative route to evaluate the importance of our assumption of fixed quantities is to exploit the monthly dimension of our data. Indeed, firms may potentially adjust their quantities through multiple infra-annual shipments; French firms export on average 3.4 months per year in each of their markets in our data. We re-run our estimations considering each monthly export observation as a separate shipment. We expect quantities to be more difficult to adjust over these small time intervals. The use of our data at the monthly level however makes our identification strategy more cumbersome and the definition of age less obvious.[28] Results are presented in columns (7) to (9) of Table 8. Looking at these infra-annual shipments yields results on beliefs' updating and its interaction with age that are quantitatively very similar to our baseline.

Altogether, these results strongly suggest that our assumption of fixed quantities is not unrealistic and does not lead our identification strategy to artificially generate our results.

CES demand. With alternative consumer preferences, markups could depend on firm size, which has two implications for our empirical strategy and results. First, prices could now depend on local market conditions, i.e. the price equation (13) should include a set of jkt fixed-effects.

[28] When computing purged prices and quantities (equations 12 and 13), the dimensionality of the data prevents us from controlling for firm-product dynamics at a monthly frequency. We therefore retain firm-product-year fixed effects but allow for country-product-year-month fixed effects. Another concern is the definition of firms' age as both the number of shipments and the time elapsed between any two of them may matter. We abstract from such issue by retaining our definition of age at the yearly level (although results are similar when using the number of months). Note that in this setting, the coefficient on the $Age_{ijkt} = 1$ dummy can be estimated since some firms export several times during their first year.

Table 8: Prediction 1: robustness (fixed quantities)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var.		$\Delta\varepsilon^q_{ijkt+1}$			$\Delta\varepsilon^q_{ijkt+1}$			$\Delta\varepsilon^q_{ijkt+1}$	
Age definition	# years since last entry (reset after 1 year of exit)								
Sample		Complex goods			Long time-to-ship			Monthly frequency	
$\widehat{v}_{ijkt}$	0.091[a]	0.138[a]		0.162[a]	0.231[a]		0.069[a]	0.105[a]	
	(0.011)	(0.011)		(0.015)	(0.012)		(0.006)	(0.005)	
Age$_{ijkt}$		-0.038[a]			-0.035[a]			-0.000	
		(0.001)			(0.001)			(0.000)	
$\widehat{v}_{ijkt}\times$Age$_{ijkt}$		-0.013[a]			-0.022[a]			-0.011[a]	
		(0.003)			(0.003)			(0.001)	
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=1$									0.110[a]
									(0.006)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=2$			0.126[a]			0.198[a]			0.075[a]
			(0.011)			(0.013)			(0.005)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=3$			0.079[a]			0.145[a]			0.056[a]
			(0.013)			(0.015)			(0.006)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=4$			0.066[a]			0.134[a]			0.054[a]
			(0.017)			(0.019)			(0.008)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=5$			0.072[a]			0.096[a]			0.039[a]
			(0.020)			(0.021)			(0.008)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=6$			0.044[b]			0.097[a]			0.035[a]
			(0.019)			(0.025)			(0.009)
$\widehat{v}_{ijkt}\times$Age$_{ijkt}=7+$			0.050[b]			0.093[a]			0.032[a]
			(0.023)			(0.030)			(0.012)
Observations	582450	582450	582450	546586	546586	546586	17540004	17540004	17540004

Robust standard errors clustered by firm in parentheses. [c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Age dummies included alone in columns (3), (6) and (9) but coefficients not reported. Complex goods and large time-to-ship means in the last quartile of the variable.

These can be easily included, and indeed this modification leaves our results largely unchanged (see appendix, Table 13, columns (1) and (2)). Second, our estimates of demand shocks might be partially reflect the firms' mark-ups. Depending on the sign of the link between mark-ups and size, this might bias the results in either direction. Here, we provide evidence that (i) this problem has a very limited impact on our results; (ii) if anything, the bias goes against our findings.

To foster intuition, let us assume that larger firms charge higher markups. In this case, ε^p_{ijkt} will be upward biased while ε^q_{ijkt} will be downward biased for large firms, leading $\widehat{v}$ to be overestimated for large firms. The coefficient on $\widehat{v}$ in Table 3, column (1) would thus be underestimated for large firms, overestimated for small ones. When we further include the interaction term between age and $\widehat{v}$, 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, absent any effect of age on belief updating. Put differently, if $\widehat{v}$ is overestimated for large firms, this should bias the coefficients against our results, i.e. the effect of learning should be underestimated.

In Table 13 in the appendix, we directly control in our estimations for firm size (the log of total quantity sold in market jk by firm i in $t-1$) and its interaction with the demand shock.[29] The results are very similar compared to those found in Table 3. In addition, we find that the interaction term between size and the demand shock displays a positive coefficient (columns (3) to (6)), and that the coefficient on the interaction term between the demand shock and firm age increases slightly in absolute value when we control for size (columns (3) and (5)). These results are consistent with $\widehat{v}$ being overestimated for large firms. Quantitatively the difference between the coefficients on the interaction term in columns (1) and (3) is however extremely limited, suggesting that CES assumption has overall very little impact on our results.

Measurement issues. In Tables A.5 and A.6 of the online appendix we perform two additional robustness checks. First, in Table A.5, columns (1) and (2), we replicate the results with equation (16) 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 might lead to imprecise estimates.

Second, A.5 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.6 of the online appendix we repeat our baseline estimations on the sample of extra EU-15 destinations. 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.

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

6 Implications for firm growth and survival

We revisit in this section some tests that have been used in the literature to support the learning mechanism (see for example Evans, 1987 and Dunne *et al.*, 1989). 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 the distribution of their growth rates and their exit decisions. As discussed in section 2.2, and shown in Table 6, young firms exhibit larger growth rates than incumbents and are thus key contributors to aggregate growth. In our model, this pattern may arise from two different mechanisms: younger firms (i) display larger *unconditional* growth rates and (ii) have more volatile growth rates together with exit rates that are non increasing with age. We concentrate in this section on the latter, i.e. on the implications of firms updating on the volatility of their growth rates and their exit decisions.[30]

[29] The online appendix (Table A.4) provides results with non-linear controls for size, replacing size variables by size bins constructed based on deciles of size computed by HS4-destination-year.

[30] Younger firms also have larger unconditional growth rates in the model, but this is a consequence of the

6.1 Firm growth

We start with the relationship between firm age and firm growth. We show here that the average absolute value of the growth rates as well as their variance within cohorts decline with firm (resp. cohort) age. Compared with previous empirical evidence, we compute the growth rate of firms' beliefs rather than the growth rate of actual firm size which might reflect in particular supply side dynamics.

The growth rates of Z^q_{ijkt} and Z^p_{ijkt} can be expressed as:

$$\begin{aligned}\Delta \ln Z^q_{ijkt+1} &= \sigma_k \Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right] && (21)\\ \Delta \ln Z^p_{ijkt+1} &= \frac{1}{\sigma_k}\Delta a_{ijkt+1} - \Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]. && (22)\end{aligned}$$

Younger firms update more and thus have larger growth rates in absolute value. As firms get older, their beliefs become more accurate, lowering their growth rate in absolute value. It follows immediately that as younger firms update more than older firms, the variance of firm growth decreases with the cohort tenure on a specific market.

As formally shown in the appendix, we get the following prediction which is a direct consequence of firm updating:

Prediction # 2

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

(b) - (variance of growth rate) - *The within cohort variance of growth rates of* Z^q_{ijkt} *and* Z^p_{ijkt} *decrease with cohort age.*

To test prediction 2.a, we estimate:

$$\left|\Delta \varepsilon^X_{ijkt}\right| = \alpha^X + \beta^X \times \text{AGE}_{ijkt} + u_{ijkt} \qquad \forall X = \{q,p\}. \tag{23}$$

Alternatively, we again relax the linearity assumption and replace AGE_{ijkt} by a set of dummy variables as we did for 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.

The results are provided in Table 9. We consider sequentially the growth rate of quantities (columns (1) and (2)) and prices (columns (3) and (4)). Both significantly decrease with firm age. The effect is also quantitatively more pronounced in the case of quantities than prices.

To test prediction 2.b, we estimate:

$$\mathbb{V}\left(\Delta \varepsilon^X_{ijkt}\right) = \delta^X \times \text{AGE}_{cjkt} + \mathbf{FE}_{cjk} + u_{ijkt} \qquad \forall X = \{q,p\} \tag{24}$$

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

specific assumption that firms' beliefs about demand, $\text{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$, are log-normally distributed – see Arkolakis *et al.* (2015) for a formal proof.

Table 9: Prediction 2.a: age and mean growth rates

	(1)	(2)	(3)	(4)
Dep. var.	$\lvert\Delta\varepsilon^{q}_{ijkt}\rvert$		$\lvert\Delta\varepsilon^{p}_{ijkt}\rvert$	
Age definition	# years since last entry (reset after 1 year of exit)			
Age_{ijkt}	-0.040^a		-0.024^a	
	(0.001)		(0.001)	
$\text{Age}_{ijkt} = 3$		-0.077^a		-0.053^a
		(0.002)		(0.001)
$\text{Age}_{ijkt} = 4$		-0.121^a		-0.079^a
		(0.003)		(0.002)
$\text{Age}_{ijkt} = 5$		-0.155^a		-0.097^a
		(0.004)		(0.003)
$\text{Age}_{ijkt} = 6$		-0.187^a		-0.110^a
		(0.005)		(0.003)
$\text{Age}_{ijkt} = 7+$		-0.220^a		-0.131^a
		(0.005)		(0.004)
Observations	2726474	2726474	2726474	2726474

Robust standard errors clustered by firm in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Controlling for year dummies does not affect the results.

expect our coefficient of interest δ^X to be negative: because firms update less their beliefs when they gain experience in a market, their quantities and prices become less volatile.

The results related to the variance of the growth rate of quantities and prices are provided in Table 10. 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. This is still true when controlling for the number of observations in 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.

Robustness. The online appendix, sections 2 and 3, contains robustness checks. We show in particular that the growth and variance of firm sales also significantly decrease with age (see columns (1) and (2) of Tables A.7 and A.10). Our tests of predictions 2.a and 2.b are also robust to: (i) controlling for firm size (Tables A.7 and A.10); (ii) focusing on sectors or destinations with slow quantities adjustment (Table A.8 and A.11); (iii) using alternative definitions of firm age (Tables A.9 and A.12).

Overall, these results strongly support the view that learning about demand is an important driver of faster growth rates for younger firms. To sum up, firm-market growth rates of young firms are larger in absolute value (Table 9) and more volatile (Table 10) due to their larger belief updating. Combined with exit rates that are not increasing with age – exit rates decline with age in all columns of Table 11 in the next section – these results imply that conditional upon survival,

Table 10: Prediction 2.b: age and variance of growth rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\mathbb{V}(\Delta\varepsilon^{q}_{ijkt})$				$\mathbb{V}(\Delta\varepsilon^{p}_{ijkt})$			
Age definition	# years since last entry (reset after 1 year of exit)				# years since last entry (reset after 1 year of exit)			
Sample	All			Permanent exporters[1]	All			Permanent exporters[1]
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.

and for a given size, younger firms have larger growth rates. This is precisely the pattern observed in the data, as shown in section 2.2. Figure 3 illustrates this contribution of belief updating to the decline in firms' sales growth and its variance with firms' age, by introducing the growth of beliefs and its variance – estimated from column (2) in table 6 and column (2) in table 10 – to figure 1.[31]

6.2 Firm survival

This final section provides evidence that the exit behavior of firms on specific markets is also in line with our model of demand learning.

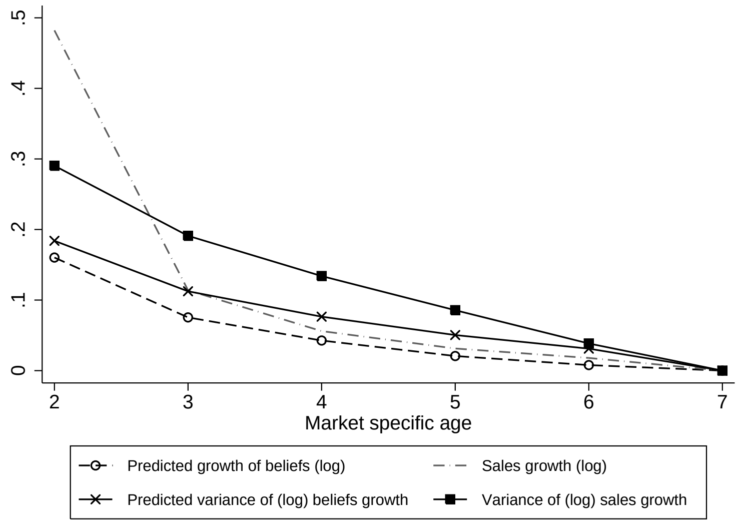
A firm decides to stop exporting a particular product to a given destination whenever the expected value of the profits stream associated with this activity becomes negative. At the beginning of period t (after having received $t-1$ signals), expected profits for period t are given by:

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

Of course, the exit decision also depends on the expected future stream of profits, which

[31]Note that the coefficient of the second year on the growth of sales is biased upward due to the incompleteness of the first year of export for some firms (see footnote 23).

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



Note: this figure plots the growth of firms' beliefs and its variance within cohort over age on each product-destination market, estimated from column (2) in table 6 and column (2) in table 10 respectively, and the corresponding coefficients on the growth and variance of growth of sales from figure 1.

depends on the evolution of C^S_{ikt}, C^S_{jkt}, $\widetilde{\theta}_{ijkt-1}$ and $\widetilde{\sigma}^2_{t-1}$ over time. Our assumption of normal prior beliefs provides the conditional distribution of $\widetilde{\theta}_{ijkt}$ given $\widetilde{\theta}_{ijkt-1}$ while the distribution of $\widetilde{\sigma}^2_{t-1}$ is deterministic. So, the evolution of firms' beliefs can be summarized by $\widetilde{\theta}_{ijkt-1}$ and t. Up to now, we have made no assumption regarding the dynamics of the C^S_{ikt} and C^S_{jkt} terms. Here, to proceed further, we follow Hopenhayn (1992) and introduce some (mild) assumptions on their dynamics. We label $A_{ijkt} \equiv C^S_{ikt}C^S_{jkt}$ 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}.[32]

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

$$V_{ijk}\left(\Omega_{ijkt}\right) = \max\left\{\mathbb{E}\left[\pi_{ijkt}\left(\Omega_{ijkt}\right)\right] + \beta\mathbb{E}\left[V_{ijk}\left(\Omega_{ijkt+1} \mid \Omega_{ijkt}\right)\right], 0\right\} \tag{25}$$

where β is the rate at which firms discount profits and where we have normalized the value of exiting to zero.[33] The value function V_{ijk} is monotonically increasing in A_{ijkt} and $\widetilde{\theta}_{t-1}$.[34] Intuitively, the flow of future expected profits inherits the properties of expected profits at time t. It follows that there exists a threshold value $\underline{\widetilde{\theta}_{ijkt-1}}(A_{ijkt}, t)$ such that a firm exits market jk at time t if $\widetilde{\theta}_{ijkt-1} < \underline{\widetilde{\theta}_{ijkt-1}}(A_{ijkt}, t)$. This implies:

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

[33]Here, we assume that an exiting firm loses all the information accumulated in the past. If the firm enters again market jk in the future, new initial beliefs will be drawn. We thus treat the exit decision as irreversible. Note that this assumption is supported by our results in Table 5.

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

Prediction # 3 (firm exit): *Given A_{ijkt} and t (firm age), (a) the probability to exit decreases with $\widetilde{\theta}_{ijkt-1}$ and (b) negative demand shocks trigger less exit for older firms.*

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

On the other hand, an important and general implication of our demand learning model is that negative demand shocks should trigger less exits for older firms (prediction 3.b). The reason is simply that firms' posterior beliefs $\widetilde{\theta}_{ijkt-1}$ depend less and less on demand shocks as firms age. Thus, the exit rate may not always be decreasing with age, but demand shocks should always have a lower impact on the exit decision in older cohorts, because they imply less updating. Note that this prediction can also be understood as another robustness check for our formulation of a passive learning model: in an active learning model, no matter the age of the firm, demand shocks may trigger new investments. Their impact on future expected profits stream should thus not be weakened for older firms (see Ericson and Pakes, 1995). This prediction is not directly tested in Pakes and Ericson (1998) because they use a much less parametric model than ours which prevents them to back out demand shocks and firms' beliefs. Their test is solely based on actual firm size.

To test prediction 3, note that from equation (5), $\widetilde{\theta}_{ijkt-1}$ depends positively on $\widetilde{\theta}_{ijkt-2}$ and a_{ijkt-1}. We therefore want to test if, conditional on A_{ijkt} and firm age, the probability to exit at the end of period $t-1$ (i.e. beginning of period t) decreases with $\widetilde{\theta}_{ijkt-2}$ and a_{ijkt-1}.

We estimate the following probabilistic model:

$$\begin{aligned}\Pr(S_{ijkt}=0|_{S_{ijkt-1}>0}) &= 1 \text{ if } \alpha\text{AGE}_{ijkt-1} + \beta\widehat{v}_{ijkt-1} + \gamma\varepsilon^q_{ijkt-1} + \delta\widehat{v}_{ijkt-1}\text{AGE}_{ijkt-1} + \mathbf{FE} + u_{ijkt} > 0 \\ &= 0 \text{ otherwise.}\end{aligned}$$

We expect β and γ to be negative, and δ to be positive. **FE** include the two sets of fixed effects $\mathbf{FE}_{ikt}$ and $\mathbf{FE}_{jkt}$, which capture C^S_{ikt} and C^S_{jkt}. We estimate this equation using a linear probability model which does not suffer from incidental parameters problems, an issue that might be important here given the two large dimensions of fixed effects we need to include.

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

Columns (4) and (5) of Table 11 test for prediction 3.b. We simply add to our baseline specification of column (3) an interaction term between age and demand shock in $t-1$.[35] We indeed find that the coefficient on this interaction term is positive: Young firms react more to a given demand shock than mature exporters on the market. In column (5), a negative demand

[35]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 13. This has no importance in columns (1) to (3) as the vector of fixed effects includes $\mathbf{FE}_{jkt}$, but it does in columns (4) and (5) as the coefficient on the interaction between $\widehat{v}_{ijkt-1}$ and age might reflect differences in $\widehat{v}_{ijkt-1}$ along the jkt dimension (as we focus on an interaction term in this case).

Table 11: Firm exit

	(1)	(2)	(3)	(4)	(5)
Dep. var.			$\Pr(S_{ijkt} = 0 \vert S_{ijkt-1} = 1)$		
Age definition			# years since last entry (reset after 1 year of exit)		
ε^q_{ijkt-1}	-0.041^a		-0.041^a		-0.041^a
	(0.000)		(0.000)		(0.000)
Age_{ijkt-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)
$\widehat{v}_{ijkt-1}$		-0.028^a	-0.031^a	-0.030^a	-0.042^a
		(0.000)	(0.000)	(0.000)	(0.000)
$\widehat{v}_{ijkt-1} \times \text{Age}_{ijkt-1}$				0.001^a	0.004^a
				(0.000)	(0.000)
Observations	8786242	8786242	8786242	8786242	8786242

Robust standard errors clustered by firm-product-destination in parentheses. Estimator: LPM. All estimations include jkt and ikt fixed effects. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

shock of 10% increases exit probability by 3.3 percentage points for a young firm (2 years after entry), but by only 1.3 percentage points after 7 years.

7 Conclusion

This paper has assessed the empirical relevance of a model of market-specific firm dynamics incorporating local demand learning. The conclusions of the model are driven by one core prediction: a new signal leads a firm to update more its beliefs, the younger the firm is. One of the implications of this result is that the growth rates of beliefs are more volatile for young firms. Combined with selection, this generates higher growth rates for young firms, even after conditioning for size, a fact that models solely based on supply-side dynamics fail to reproduce. We have also derived additional predictions for firm survival: exit rates decrease with firms' beliefs and the demand shocks they face, and those demand shocks trigger more exits in younger cohorts.

Using detailed exporter-level data containing the values and the quantities sold by French firms in export markets, we have shown how this model can be used to separately infer firm-market specific demand shocks and prior beliefs about demand, and found that its predictions are strongly supported by the data. Importantly, our methodology and therefore our results are consistent with any possible dynamics of firm productivity.

Although the learning process appears to be especially strong in the first years after entry, even the most experienced exporters in our sample still exhibit significant belief updating. Interestingly, the learning process generates a form of hysteresis: The most experienced firms are less sensitive to demand shocks in terms of sales and exit decisions. This in turn suggests that aggregate uncertainty shocks should have heterogeneous effects across industries, depending on their age structure. Our results also put forward a new source of irreversibility in the exit decision as we provided evidence that the accumulated knowledge is quickly lost during exit periods.

We concentrated on post-entry dynamics and assumed away other possible sources of information for firms than their own sales. 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 information gathered from selling other products in the same destination or from shipping the same product to other countries. Information could also be obtained observing rivals selling in the same market. This would allow understanding how information spills 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. (2013), "A unified theory of firm selection and growth", *Quarterly Journal of Economics*, vol. forthcoming.

Arkolakis, C., Papageorgiou, T. and Timoshenko, O. (2015), "Firm learning and growth", unpublished manuscript.

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

Baker, S. and Bloom, N. (2013), "Does uncertainty drive business cycles? Using disasters as a natural experiment", NBER Working Papers 19475, National Bureau of Economic Research, Inc.

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. (2014), "Fluctuations in Uncertainty", *Journal of Economic Perspectives*, vol. 28 n° 2: pp. 153–76.

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 nº 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 nº 4: pp. 1947–1982.

Clementi, G. L. and Hopenhayn, H. A. (2006), "A Theory of Financing Constraints and Firm Dynamics", *The Quarterly Journal of Economics*, vol. 121 nº 1: pp. 229–265.

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

Defever, F., Heid, B. and Larch, M. (2015), "Spatial Exporters", *Journal of International Economics*, vol. 95 nº 1: pp. 145–156.

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 Republica 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", Unpublished 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.

Fernandes, A. and Tang, H. (2014), "Learning from Neighbors' Export Activities: Evidence from Exporters' Survival", *Journal of International Economics*, vol. 94 nº 1: pp. 87–94.

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?", *Economica*, vol. forthcoming.

Gourio, F. and Rudanko, L. (2014), "Customer capital", *Review of Economic Studies*, vol. 81 nº 3: pp. 1102–1136.

Guimaraes, P. and Portugal, P. (2010), "A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects", *Stata Journal*, vol. 10 nº 4: pp. 628–649.

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 n° 2: pp. 347–361.

Hopenhayn, H. A. (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", *Econometrica*, vol. 60 n° 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 n° 1: pp. 75–90.

Jovanovic, B. (1982), "Selection and the Evolution of Industry", *Econometrica*, vol. 50 n° 3: pp. 649–70.

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

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

Lawless, M. (2014), "Age or size? Contributions to job creation", *Small Business Economics*, vol. 42 n° 4: pp. 815–830.

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

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

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

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

Morales, E., Sheu, G. and Zahler, A. (2014), "Extended Gravity", unpublished manuscript, 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° 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° 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 nº 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 nº 5: pp. 1639–1666.

Ruhl, K. J. and Willis, J. L. (2014), "New Exporter Dynamics", Unpublished manuscript.

Timoshenko, O. A. (2015), "Product Switching in a Model of Learning", *Journal of International Economics*, vol. 95 nº 2: pp. 233–249.

A Appendix

A.1 Theory

A.1.1 Detailed derivations and proofs

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

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

The FOC writes:

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

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

And from the constraint, we get:

$$p_{ijkt}^* = \left(\frac{\sigma_k}{\sigma_k - 1}\frac{w_{it}}{\varphi_{ikt}}\right) \left(\frac{e^{\frac{a_{ijkt}}{\sigma_k}}}{E_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]}\right)$$

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

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

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

$$\Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] = \frac{1}{\sigma_k \left(\frac{\sigma_\epsilon^2}{\sigma_0^2} + t \right)} \left(a_{ijkt} - \frac{\left(\theta_0 + \frac{\sigma_0^2}{2\sigma_k} + \overline{a}_{ijkt-1} \frac{\sigma_0^2}{\sigma_\epsilon^2} (t-1) \right)}{\left(1 + \frac{\sigma_0^2}{\sigma_\epsilon^2} (t-1) \right)} \right) \tag{26}$$

Prediction 1. Prediction 1 states that following a new signal, updating is larger for younger firms. Updating is measured directly by $\Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right]$ in (26). We get:

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

The larger the demand shock, the larger the updating. Further, the denominator increases with t: updating is larger for younger firms, which can be directly measured by g_t.

Impact of market uncertainty. Moreover, the updating process is also affected by the level of market uncertainty σ_ϵ^2. Formally, we get:

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

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

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

which is larger (less negative) in more uncertain markets (with larger σ_ϵ^2).

Prediction 2a. Prediction 2a states that expected absolute value of growth rates decrease with age. Growth rates are given by:

$$\Delta \ln Z_{ijkt+1}^q = \sigma_k \Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{27}$$

$$\Delta \ln Z_{ijkt+1}^p = \frac{1}{\sigma_k} \Delta a_{ijkt+1} - \Delta \ln \mathbb{E}_t \left[e^{\frac{a_{ijkt+1}}{\sigma_k}} \right] \tag{28}$$

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

Second, using (26) and the fact that $\mathbb{E}\left[a_{ijkt}\right] = \overline{a}_{ijkt-1}$, the absolute value of the expected growth rate of $\mathbb{E}_{t-1}\left[e^{\frac{a_{ijkt}}{\sigma_k}}\right]$ is given by:

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

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

$$\begin{aligned}\mathbb{E}\left[\left|\Delta \ln Z^q_{ijkt+1}\right|\right] &= \sigma_k\mathbb{E}\left[\left|\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right|\right]\\ \mathbb{E}\left[\left|\Delta \ln Z^p_{ijkt+1}\right|\right] &= \mathbb{E}\left[\left|\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right|\right]\end{aligned}$$

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

Dynamics of prices and quantities. While prediction 2a concentrates on the expected absolute value of growth rates, it is interesting to further note that the model predicts expected growth rates of opposite signs for quantities and prices. This can me thought as an additional discriminative criterion between our learning model and models of consumer base accumulation where firms price low in their first years to attract consumers.

This result can be seen directly from (27) and (28), taken in expected terms. We find:

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

Further, given that firms that decrease in size will on average be more likely to exit, the expected growth rate of quantities must be positive for survivors. It follows that the expected growth rate of prices for these firms should be negative and smaller by a factor $-\frac{1}{\sigma_k}$. Quantitatively, this is very close to what we find in table 6.

Prediction 2b. Prediction 2b states that the variance of growth rates within cohort decrease with cohort age. The variance of these growth rates can be expressed as:

$$\mathbb{V}\left[\Delta \ln Z^q_{ijkt+1}\right] = \sigma_k^2\mathbb{V}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right) \tag{29}$$

$$\begin{aligned}\mathbb{V}\left[\Delta \ln Z^p_{ijkt+1}\right] &= \left(\frac{1}{\sigma_k}\right)^2\mathbb{V}\left(\Delta a_{ijkt+1}\right) + \mathbb{V}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right]\right)\\ &\quad -\frac{2}{\sigma_k}\mathrm{Cov}\left(\Delta \ln \mathbb{E}_t\left[e^{\frac{a_{ijkt+1}}{\sigma_k}}\right], \Delta a_{ijkt+1}\right)\end{aligned} \tag{30}$$

First, a_{ijkt+1} and a_{ijkt} being drawn from the same distribution, $\mathbb{V}\left[\Delta a_{ijkt+1}\right] = 2\sigma_\epsilon^2$.

Second, using (26), it is straightforward to show that:

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

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

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

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

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

Plugging terms into (29) and (30), and after simplification, we get:

$$\mathbb{V}\left[\Delta \ln Z_{ijkt+1}^q\right]=\left(\frac{\sigma_\epsilon}{\left(\frac{\sigma_\epsilon^2}{\sigma_0^2}+t\right)}\right)^2 \tag{31}$$

$$\mathbb{V}\left[\Delta \ln Z_{ijkt+1}^p\right]=\left(\frac{\sigma_\epsilon}{\sigma_k}\right)^2\left(\left(\frac{1}{\left(\frac{\sigma_\epsilon^2}{\sigma_0^2}+t\right)}+1\right)^2+1\right) \tag{32}$$

Both expressions are strictly decreasing with t.

A.1.2 Fixed quantities: discussion

We consider in this section alternative versions of the model where (i) firms choose prices before quantities and (ii) firms partly (but not fully) adjust their quantity decisions to the demand shocks, and discuss their implications for our identification strategy.

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

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

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

From the FOC and the constraint we get:

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

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

$$
\begin{aligned}
\varepsilon_{ijkt}^{q} &= \ln Z_{ijkt}^{q} = a_{ijkt} \\
\varepsilon_{ijkt}^{p} &= \ln Z_{ijkt}^{p} = 0
\end{aligned}
$$

Regressing ε_{ijkt}^{p} on ε_{ijkt}^{q} should therefore generate insignificant $\widehat{\beta}$ coefficients and the absolute value of ε_{ijkt}^{q} should not decrease with age.

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

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

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

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

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

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

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

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

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

Note that if $\overline{a}^1_{ijk} = \overline{a}_{ijk}$ and $\varsigma = 1$, all the demand shock is observed and ε^q_{ijkt} captures the demand shock only while ε^p_{ijkt} does not depend neither on the demand shock, nor on firm beliefs (which are irrelevant in that case). This case is equivalent to the one where prices are set first, as shown above. If on the other hand $\overline{a}^1_{ijk} = \varsigma = 0$, we are back to our baseline assumption of fixed quantities. Importantly, all our theoretical predictions still hold in the intermediate case. In particular, equation (11) still describes the evolution of beliefs, which are now related to the distribution of a^2_{ijkt}. Moreover, $\Delta\varepsilon^q_{ijkt}$ still captures the expected growth rate of beliefs, as $\mathbb{E}\left[\left|\Delta \ln Z^q_{ijk,t+1}\right|\right] = \sigma_k \mathrm{E}\left[\left|\Delta \ln \mathbb{E}_t \left[e^{\frac{a^2_{ijkt+1}}{\sigma_k}}\right]\right|\right]$.

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

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

The fact that ε^q_{ijkt} is measured with noise has the same consequence as a classical measurement error: $\widehat{\beta}$ suffers from attenuation bias, which in turn leads to overestimate in absolute terms the demand shock as $\widehat{v}_{ijkt} = \varepsilon^p_{ijkt} - \widehat{\beta}\varepsilon^q_{ijkt}$. Put differently, due to attenuation bias we would expect $|\widehat{v}_{ijkt}| > |\frac{1}{\sigma_k} a^2_{ijkt}|$.

Are our results on beliefs updating (prediction 1) likely to be affected? Consider our baseline specification, equation (18). There are two distinct issues here. First, $\widehat{\alpha}_1$ – the average extent of belief updating – might be downward biased if our demand shocks are overestimated. Second, and key for our findings, if demand shocks are *more* overestimated for older firms (for instance because they have more capacities to adjust quantities), $\widehat{\alpha}_1$ would decline with age, even absent any belief updating.

As discussed in the main text, a simple way to gauge the importance of this issue is to focus on sectors or destinations for which quantities are more likely to be rigid (those for which $\widehat{v}_{ijkt}$ is more likely to be correctly estimated) and to compare the results with our baseline estimates of Table 3. Alternatively, running our estimations at the monthly level also makes the fixed quantities assumption more likely to hold. If an attenuation bias induced by some partial quantity adjustment was driving our results, we would expect belief updating to decrease less with age in these regressions. Results shown in Table 8 show first that the average extent of belief updating is of similar magnitude or only slightly larger than in our baseline estimates, which suggests that attenuation bias, if any, is limited (note that if present, such attenuation bias would imply that our estimates of σ_k are a bit too high). On the other hand, in Table 8 belief updating declines *more* with age that in our baseline estimates, which clearly suggests that attenuation bias is not correlated with age in a way that would spuriously generate our results. Altogether, our results on the learning process are therefore unlikely to be driven by our assumption of fixed quantities.

A.2 Additional Tables

Table 12: Prediction 1: alternative age definitions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	$\Delta\varepsilon^{q}_{ijkt+1}$				$\Delta\varepsilon^{q}_{ijkt+1}$			
Age definition	# years since last entry (reset after 2 years exit)				# years exporting since first entry			
$\widehat{v}_{ijkt}$	0.075^a	0.106^a	0.106^a		0.075^a	0.101^a	0.101^a	
	(0.009)	(0.008)	(0.003)		(0.009)	(0.007)	(0.004)	
Age$_{ijkt}$		-0.036^a	-0.036^a			-0.034^a	-0.034^a	
		(0.001)	(0.000)			(0.001)	(0.000)	
$\widehat{v}_{ijkt}\times$Age$_{ijkt}$		-0.008^a	-0.008^a			-0.007^a	-0.007^a	
		(0.002)	(0.001)			(0.002)	(0.001)	
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 2$				0.102^a				0.098^a
				(0.008)				(0.007)
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 3$				0.069^a				0.070^a
				(0.009)				(0.009)
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 4$				0.063^a				0.072^a
				(0.011)				(0.012)
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 5$				0.062^a				0.064^a
				(0.014)				(0.013)
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 6$				0.051^a				0.062^a
				(0.012)				(0.013)
$\widehat{v}_{ijkt}\times$Age$_{ijkt} = 7+$				0.051^a				0.051^a
				(0.013)				(0.014)

Robust standard errors clustered by firm in parentheses (bootstrapped in columns (3) and (7)). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Age dummies included alone in columns (4) and (8) but coefficients not reported. 2,726,474 obs. in all estimations.

Table 13: Prediction 1: robustness of the CES assumption

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	$\Delta\varepsilon^q_{ijkt+1}$					
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_{ijkt}		$\overline{\text{Size}}_{ijk,t/t-1}$	
$\widehat{v}_{ijkt}$	0.159^a		0.095^a		0.075^a	
	(0.011)		(0.013)		(0.011)	
Age_{ijkt}	-0.041^a		-0.013^a		-0.044^a	
	(0.001)		(0.001)		(0.001)	
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}$	-0.008^a		-0.009^a		-0.013^a	
	(0.002)		(0.002)		(0.002)	
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=2$		0.160^a		0.088^a		0.065^a
		(0.010)		(0.013)		(0.011)
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=3$		0.118^a		0.048^a		0.013
		(0.010)		(0.014)		(0.011)
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=4$		0.118^a		0.046^a		0.007
		(0.011)		(0.015)		(0.011)
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=5$		0.111^a		0.038^b		-0.004
		(0.014)		(0.017)		(0.014)
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=6$		0.098^a		0.024		-0.020
		(0.014)		(0.018)		(0.016)
$\widehat{v}_{ijkt}\times\text{Age}_{ijkt}=7+$		0.108^a		0.033^b		-0.014
		(0.012)		(0.017)		(0.014)
Size_{ijkt-1}			-0.082^a	-0.081^a	0.010^a	0.011^a
			(0.001)	(0.001)	(0.000)	(0.000)
$\widehat{v}_{ijkt}\times\text{Size}_{ijkt-1}$			0.014^a	0.015^a	0.018^a	0.019^a
			(0.001)	(0.001)	(0.002)	(0.002)
Observations	2739927	2739927	2739927	2739927	2739927	2739927

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